

Lauri Laaksonen

SPECTRAL RETINAL IMAGE PROCESSING AND ANALYSIS FOR OPHTHALMOLOGY

Thesis for the degree of Doctor of Science (Technology) to be presented with due permission for public examination and criticism in the Auditorium 4301 at Lappeenranta University of Technology, Lappeenranta, Finland on the 27th of May, 2016, at noon.

Acta Universitatis
Lappeenrantaensis 699

Supervisor Professor Lasse Lensu
LUT School of Engineering Science
Lappeenranta University of Technology
Finland

Reviewers Professor Alain Trémeau
Laboratory of Computer Graphics and Image Vision
Department of Physics
University of Jean Monnet
France

Professor Emanuele Trucco
School of Computing
University of Dundee
Scotland, UK

Opponents Professor Ulla Ruotsalainen
Department of Signal Processing
Tampere University of Technology
Finland

Professor Emanuele Trucco
School of Computing
University of Dundee
Scotland, UK

ISBN 978-952-265-956-9
ISBN 978-952-265-957-6 (PDF)
ISSN-L 1456-4491
ISSN 1456-4491

Lappeenrannan teknillinen yliopisto
Yliopistopaino 2016

To my friends and family.

Preface

This thesis consists of work performed during project ReVision, a collaboration of the Machine Vision and Pattern Recognition Laboratory in Lappeenranta University of Technology, Department of Ophthalmology in University of Tampere and Color Research Laboratory in University of Eastern Finland (Joensuu). A number of people participated, directly or indirectly, in the work presented in this thesis, and I want to express my gratitude.

Firstly, I want to thank my supervisor Professor Lasse Lensu for the constant guidance, support and ideas, both during the PhD work and earlier. The balanced combination of direction and individual responsibility made me a considerably better researcher. I also want to thank Professor Ela Claridge for supervising my work in Birmingham and helping with numerous practical matters during my stay, and valued collaboration in the project.

I want to thank my co-workers and co-authors Professors Markku Hauta-Kasari, Professor Hannu Uusitalo, Dr. Pauli Fält, Dr. Pasi Ylitepsa, Dr. Kati Adjers, Joni Herttuainen and Antti Hannuksela for rewarding collaboration, and new insights and points of view. I also want to extend my gratitude to Professor Alain Trémeau of the University of Jean Monnet and Professor Emanuele Trucco of the University of Dundee for reviewing the thesis and suggesting many improvements that lead to a significantly better version of the thesis. Furthermore, Professor Trucco and Professor Ulla Ruotsalainen of Tampere University of Technology have my thanks for agreeing to be the opponents in my thesis defence.

I wish to thank the whole MVPR laboratory for creating a relaxed, enjoyable and yet professional work environment. Doctors Jukka Lankinen, Ekaterina Ryabchenko and Nataliya Strokina deserve a special mention. I am thankful for your friendship and our numerous conversations that were always interesting, often peculiar, and occasionally related to research. I want to thank my parents and my family for teaching me to value education and hard work, and for your unwavering support. It was always appreciated, even though that appreciation was too rarely expressed.

I would like to thank the Academy of Finland for the financial support of the ReVision project (No. 259560), and Lappeenranta teknillisen yliopiston tukisäätiö for the financial support to initiate the thesis work. I also wish to thank Prof. Majid Mirmehdi from the University of Bristol, U.K. for assistance with the comparative experiments.

Lappeenranta, May 2016

Lauri Laaksonen

Abstract

Lauri Laaksonen

Spectral retinal image processing and analysis for ophthalmology

Lappeenranta, 2016

159 p.

Acta Universitatis Lappeenrantaensis 699

Diss. Lappeenranta University of Technology

ISBN 978-952-265-956-9

ISBN 978-952-265-957-6 (PDF)

ISSN-L 1456-4491

ISSN 1456-4491

Diabetic retinopathy, age-related macular degeneration and glaucoma are the leading causes of blindness worldwide. Automatic methods for diagnosis exist, but their performance is limited by the quality of the data. Spectral retinal images provide a significantly better representation of the colour information than common grayscale or red-green-blue retinal imaging, having the potential to improve the performance of automatic diagnosis methods.

This work studies the image processing techniques required for composing spectral retinal images with accurate reflection spectra, including wavelength channel image registration, spectral and spatial calibration, illumination correction, and the estimation of depth information from image disparities. The composition of a spectral retinal image database of patients with diabetic retinopathy is described. The database includes gold standards for a number of pathologies and retinal structures, marked by two expert ophthalmologists. The diagnostic applications of the reflectance spectra are studied using supervised classifiers for lesion detection. In addition, inversion of a model of light transport is used to estimate histological parameters from the reflectance spectra.

Experimental results suggest that the methods for composing, calibrating and post-processing spectral images presented in this work can be used to improve the quality of the spectral data. The experiments on the direct and indirect use of the data show the diagnostic potential of spectral retinal data over standard retinal images. The use of spectral data could improve automatic and semi-automated diagnostics for the screening of retinal diseases, for the quantitative detection of retinal changes for follow-up, clinically relevant end-points for clinical studies and development of new therapeutic modalities.

Keywords: Image processing, spectral imaging, retinal imaging, diabetic retinopathy

SYMBOLS AND ABBREVIATIONS

AMD	age-related macular degeneration
AUC	area under the curve
BDT	binary decision tree
BRISK	binary Robust Invariant Scalable Keypoints
BristolDB	Bristol retinal image data set
CC	correlation coefficient
CCD	charge-coupled device
CD2	similarity measure by Myronenko et al.
CLAHE	contrast limited adaptive histogram equalisation
CT	computed tomography
DiaRetDB1	DiaRetDB1 diabetic retinopathy database
DiaRetDB2	DiaRetDB2 diabetic retinopathy database
DIV	difference in variation
DR	diabetic retinopathy
ED-DB-ICP	edge-driven dual-bootstrap iterative closest point
FA	fluorescein angiogram
FCM	fuzzy c-means clustering
FNR	false-negative rate
FOV	field-of-view
FP	false positive
FPR	false-positive rate
FREAK	fast retina keypoint
GDB-ICP	generalized dual-bootstrap iterative closest point
GLCM	graylevel co-occurrence matrix
GMM	Gaussian mixture model
Graph	graph-cuts
GrowCut	GrowCut algorithm
HMA	haemorrhage and microaneurysm
ICP	iterative closest point
IRMA	intra-retinal microvascular abnormalities
ISOS	inner segment/outer segment
KDE	kernel density estimate
kNN	k-nearest neighbour

Lab	CIE Lab
LBP	local binary pattern
LBPHF	local binary pattern histogram Fourier feature
LED	light emitting diode
LSO	laser scanning ophthalmoscopy
MAP	maximum a posteriori
MC	Monte Carlo
MCMC	Markov-chain Monte Carlo
MI	mutual information
ML	maximum likelihood
MP	macular pigment
MRI	magnetic resonance imaging
MS	similarity measure by Cohen and Dinstein
MSER	maximally stable extremal regions
MSRM	maximal similarity region merging
NB	Bayesian probability regions
NCC	normalised cross-correlation
NN	neural network
NPV	negative predictive value
OCT	optical coherence tomography
PCA	principal component analysis
PDF	probability density function
PET	positron emission tomography
PPV	positive predictive value
QI	quality index
RANSAC	random sample consensus
RC	minimisation of residual complexity
RF	random forest
RGB	red-green-blue
RMSE	root-mean-square error
ROC	receiver operating characteristic
ROI	region of interest
RPE	retinal pigment epithelium
RRGS	recursive region-growing segmentation
SAD	sum of absolute differences
SAM	spectral angle measure

SCM	spectral correlation measure
SD-OCT	spectral domain optical coherence tomography
SH	systemic hypertension
SID	spectral information divergence
SIFT	scale invariant feature transform
SN	sensitivity
SNAKE	active contour
SP	specificity
SSD	sum of squared differences
SURF	speeded-up robust feature
SVM	support vector machine
SWFCM	spatially weighted fuzzy c-means clustering
TH5	Otsu thresholding
TN	true negative
TNR	true-negative rate
TP	true positive
TPR	true-positive rate
\mathbf{x}	a vector
\mathbf{A}	a matrix
\mathbf{A}^T	the transpose of \mathbf{A}
\mathbf{A}^{-1}	matrix inversion of \mathbf{A}
\mathbf{I}	identity matrix
s	photon propagation step size
ϕ	photon propagation direction
$\log(x)$	natural logarithm of x
μ_t	tissue interaction coefficient
μ_a	tissue absorption coefficient
μ_s	tissue scattering coefficient
ξ	random number from uniform distribution between $[0, 1]$
E	photon energy
Φ	photon scattering azimuthal angle
θ	photon scattering deflection angle
g	tissue anisotropy factor
d_b	photon distance to tissue boundary
α_i	angle of incidence (boundary reflection)

n_i	refractive index of incident medium
n_t	refractive index of transmission medium
$R(\alpha_i)$	likelihood of internal reflection
α_t	photon angle of transmission
s_u	camera scale factor
d_u	pixel width
d_v	pixel height
$\delta u^{(r)}$	radial distortion term
$\delta u^{(t)}$	tangential distortion term
p_d	geometric distortion model parameters
$p(x)$	probability of x
μ	mean
\bar{x}	sample mean of x
σ	standard deviation
σ_{xy}	covariance of x and y
Σ	covariance matrix
\sum_i	sum over i
$\sum_{i=n}^m$	sum over i from n to m
\mathbf{W}	weight matrix
\mathbf{T}	transmittance matrix
η	noise term
t_λ	exposure time of wavelength channel λ
f_{ill}	estimated illumination field
\mathbf{v}_0	vector of reflected intensities
\mathbf{v}	vector of image intensity values
Y	intermediate template image
d_U	image width
α_{fov}	horizontal field-of-view angle of camera
\mathbf{R}	homogeneous rotation matrix
φ_x	angle of rotation around x-axis
β	vector of illumination field parameters
f	camera focal length
α_0	radial vignetting factor
γ	angle of camera tilt
$A \cup B$	union of A and B

$A \cap B$	intersection of A and B
ω	haemoglobin absorption based weight factor
res	vector of residual values
P	projection matrix
ϵ	reprojection error
$\log_n(x)$	base n logarithm of x
$p(x y)$	probability of x given y
C	matrix of principal component vectors

1	Introduction	17
1.1	Objectives	18
1.2	Contribution and publications	19
1.3	Outline of the thesis	20
2	Spectral fundus imaging and spectral image composition	21
2.1	Introduction	21
2.2	Related work	22
2.3	Spectral fundus image formation	25
2.3.1	Structure of the eye	27
2.3.2	Modelling of light interaction with retinal tissue	28
2.4	Spectral fundus image acquisition	33
2.4.1	30-channel spectral fundus camera	33
2.4.2	Six-channel spectral fundus camera	35
2.5	Spectral camera calibration	36
2.5.1	Related work	38
2.5.2	Methods	38
2.5.3	Experiments and results	43
2.5.4	Discussion	46
2.5.5	Summary	46
2.6	Spectral image composition	47
2.6.1	Related work	47
2.6.2	Methods	49
2.6.3	Registration strategy	50
2.6.4	Experiments	51
2.6.5	Results	60
2.6.6	Discussion	69
2.6.7	Summary	71
2.7	Illumination correction in spectral images	72
2.7.1	Illumination field estimation using the image spectra	73
2.7.2	Experiments and results	75
2.7.3	Discussion	77
2.7.4	Summary	77
2.8	3D-reconstruction of the retina from spectral images	78
2.8.1	Methods	80
2.8.2	Experiments and results	83
2.8.3	Discussion	84
2.8.4	Summary	86
3	Spectral image database of diabetic retinopathy patients	89
3.1	Introduction	89
3.2	Public fundus image databases	90
3.3	DiaRetDB2 spectral retinal image database with gold standard	91
3.3.1	Human subjects and ethical considerations	91

3.3.2	Imaging procedure	91
3.3.3	Gold standard annotations	91
3.3.4	Summary	93
3.4	Effect of ground truth inaccuracy on lesion classification	93
3.4.1	Related work	94
3.4.2	Methods	95
3.4.3	Experiments and results	97
3.4.4	Discussion	100
3.4.5	Summary	100
3.5	Annotation refinement	101
3.5.1	Related work	101
3.5.2	Region refinement through stable probability regions	104
3.5.3	Experiments and results	104
3.5.4	Discussion	108
3.5.5	Summary	108
4	Medical applications of spectral fundus data	110
4.1	Introduction	110
4.2	Lesion detection by supervised classification	112
4.2.1	SVM	112
4.2.2	Gaussian Mixture Models	112
4.2.3	Neural Networks	113
4.2.4	Random Forests	113
4.2.5	Evaluation	113
4.3	Histological parameter maps from spectral images	114
4.3.1	Model generation	114
4.3.2	Alignment of model and image data	114
4.4	Experiments and results	116
4.4.1	Lesion detection by supervised classification	116
4.4.2	Histological parameter maps from spectral images	119
4.5	Discussion	124
4.6	Summary	126
5	Discussion	134
5.1	Main contributions	134
5.2	Limitations of the study	135
5.3	Future work	135
6	Conclusion	137
	Bibliography	139
	Appendix	
I	Spectral image composition	161
I.1	Registration method parameters	161
I.2	Transformation parameter distributions used for sampling	162

Introduction

Diabetic retinopathy (DR) , age-related macular degeneration (AMD) and glaucoma are among the leading causes of blindness world wide [78, 109, 186]. In addition to the personal impact on quality of life due to impairment or loss of vision, the aforementioned conditions form a significant financial cost to the society in the form of disability benefits, medical care and early retirement. [65, 104]

Prolonged high blood glucose levels associated with diabetes damage the capillaries and disrupt the circulation of blood in the retina. As the delivery of oxygen and nutrients is disrupted, the growth of new retinal blood vessels accelerates as the retina tries to circumvent the disrupted circulation. The increased growth rate of vessels can cause a dilation of small blood vessels (intra-retinal microvascular abnormalities (IRMA)) and a formation of new vessels. IRMA and neovascularisation lead to a high risk of haemorrhage, and with the tendency of the new vessels to form over the retina, haemorrhages may block light entering the photoreceptors and lead to sudden loss of vision (this process is known as proliferative diabetic retinopathy). [219] The most common cause of visual impairment in diabetic patients is macular edema, a condition where the increase vascular permeability causes exudation and swelling of the macular structures.

AMD is the leading cause of blindness in the elderly [99]. With the ageing of the eye fundus, the metabolism of the retina may begin to slowly deteriorate. Problems with the metabolism may lead to an accumulation of extracellular material, forming yellow or grey spots called drusen. The appearance of large drusen in significant numbers has been associated with the development of exudative form of AMD that is the most common cause of AMD-related visual impairment.

Despite the severity of the diseases, a number of treatments exist that can delay or stop the progression of the pathologies and prevent the loss of vision (e.g., [33, 61, 77, 79, 168, 193]). Therefore, early detection of pathology is crucial for effective and cost-effective treatment, and preserving the vision of the patient. Diabetic retinopathy (DR) and AMD are typically diagnosed from colour or grayscale fundus images. Fundus imaging offers a non-invasive view of the human retina, but due to the small aperture (the pupil),

the curvature of the fundus, and the optical system of the eye, specialised optics are required to acquire an in-focus image of the curved fundus on a flat digital camera sensor array.

Eye disease screening programs have been implemented [41, 142, 199] to bring patients in early stages of the disease (who have not yet exhibited symptoms) into the treatment program. Widening the screening programs, however, means a significant increase in the workload of the ophthalmologists responsible for performing diagnosis based on the images.

To enable automatic diagnostics and support the screening programs, a significant body of work on automatic detection of lesions related to DR and AMD exists (e.g., [23, 72, 138, 161, 162, 184, 198, 227]). However, automated methods are limited by the available data. Early pathological changes in the retina may be difficult or impossible to automatically detect from red-green-blue (RGB) or grayscale fundus images.

Various imaging modalities have been developed to acquire more representative views of different features of the eye fundus. These modalities include angiography, retinal optical coherence tomography (OCT), retinal magnetic resonance imaging (MRI) and laser scanning ophthalmoscopy (LSO). Products providing multiple modalities in a single device have also become available (e.g., [171, 222]).

Among the promising relatively recent imaging modalities is spectral fundus imaging. Spectral images combine the benefits of spectroscopy with the field-of-view (FOV) of traditional retinal imaging. As the spectra are measured simultaneously over the whole FOV, the analysis of the spectra is not limited to a set of point-wise measurements. The spectra can be used to better discriminate between different retinal tissues and structures than standard RGB colour information, potentially improving segmentation and contrast of the structures.

However, spectral imaging has a number of additional challenges compared to the acquisition of traditional grayscale or RGB fundus images. Depending on the approach to spectral fundus imaging, several steps are required to compose a spectral image with correct spectral content from the individual channel images. Depending on the system, these steps may include image registration, correcting geometric distortions, correcting bias due to uneven illumination fields in the channel images or due to spectral aberrations, and dealing with artifacts caused by dust and dirt in the optics.

1.1 Objectives

Spectral fundus image data has the potential to significantly improve the automatic diagnosis of retinal pathologies. The goal of the work in this thesis was to study two available spectral image acquisition systems and to study and develop methods for composing and post-processing the spectral channel images acquired by the systems. The challenges of the acquisition of accurate retinal spectra are addressed in this work by the study of registration, calibration and illumination correction of the retinal spectral images.

One of the main goals of the work was the composition of a database of spectral fundus images with gold standard of the location of lesions of multiple types, provided by two

expert ophthalmologists. Another important goal was to provide examples and considerations on the use of the spectral data.

The scope of this work was limited to the acquisition, processing and use of spectral fundus image data in the context of automatic detection of retinal pathologies. The study of automatic detection concentrated on intensity, colour and spectral features of the fundus images. In-depth studies into other automatic diagnostic approaches, and the medical and biological study of retinal pathologies were considered out of the scope.

1.2 Contribution and publications

During the thesis work, an evaluation of the performance of a number of image registration methods on spectral fundus image data was performed, and the results were reported in [129]. An extended study on the registration has been performed and a manuscript of the study has been submitted for review [128]. The author was responsible for the performing and reporting of both the initial and the extended studies.

This thesis introduces a method for spectral retinal image illumination correction that considers the consistency of the image spectra. The author has been responsible for the reporting, partially implementing, and planning the implementation of the method.

In addition, an extension of the method by Lin and Medioni [139] for estimating the 3D-structure of the retina from the disparities between retinal images was implemented. The author was responsible for a part of the implementation of the original method, planning the implementation of the extensions and the reporting of the extended method.

As a part of the ReVision consortium project, methods for visualising spectral images with the visual contrast of lesions or retinal structures optimised were developed. The methods were published in [57]. The author contributed to the quantitative evaluation of the methods and participated in the reporting of the study.

Among the major contributions of the work in this thesis is the gathering of the gold standard annotations for the spectral fundus image database DiaRetDB2 diabetic retinopathy database (DiaRetDB2). The author implemented the software tool used for the annotations, guided and supported the annotation work, and was responsible for the development, implementation and evaluation of the annotation post-processing and the baseline lesion detection methods.

As a part of the annotation gathering, the effect of ground truth inaccuracy on different image features and the possibility of refinement by post-processing of the annotations was studied. A paper on the annotation refinement work has been published in [131]. The annotation refinement manuscript has been extended with results related to the relevance of the annotation accuracy, and the extended manuscript has been submitted for review [130]. The author was responsible for the development and implementation of the post-processing methods, and for the experiments, evaluation and reporting of the results.

During a research visit in the University of Birmingham, the author participated in the extension of the inverse modelling of light transport in the retina. The author was responsible for improving the original model (Styles et al. [216]), developing approaches

for aligning the model with the image data, and applying the model on data from a different spectral retinal imaging system.

The main contributions of the thesis work can be summarised as

- Quantitative evaluation of registration methods for channel image registration.
- Implementation, improvement and evaluation of the method of Lin and Medioni [139] for estimating depth information for retinal images.
- Method for correcting uneven illumination in spectral images.
- Software tool and support for gathering gold standard annotations for DiaRetDB2.
- Study on the effect of ground truth inaccuracy on the performance of supervised classifiers.
- A method for post-processing coarse manual annotations.
- Improvement of the light interaction model by Styles et al. [216], and the alignment of the model with spectral image data.

1.3 Outline of the thesis

The rest of the thesis is structured as follows:

Chapter 2, Spectral fundus imaging and spectral image composition, presents a theory of spectral fundus image formation and its modelling, detailed descriptions of the imaging equipment utilised in this thesis, considerations and approaches to challenges in spectral fundus camera calibration, and the composition of spectral images from individual channel images. Spectral image composition includes the introduction and quantitative evaluation of image registration methods and strategies. In addition, certain unique features (and their use) of spectral images, such as channel-wise independent illumination fields and their correction, and stereo reconstruction from the disparity due to inter-channel eye movement, are discussed.

Chapter 3, Spectral image database of diabetic retinopathy patients, details the acquisition and composition of a new publicly available spectral fundus image database with ground truth markings by an expert ophthalmologist, DiaRetDB2. The importance of the level of spatial accuracy of the ground truth in lesion detection is quantitatively evaluated for a number of different image features. Methods and quantitative evaluation for the post-processing of the expert annotations are presented.

Chapter 4, Medical applications of spectral fundus data, discusses the use of the spectral image data in medical applications. A method for automatic diagnosis based on the classification of spectral colour features is presented and evaluated. Another application using inverse modelling of light interaction in retinal tissue to generate histological parameter maps of the retina is presented.

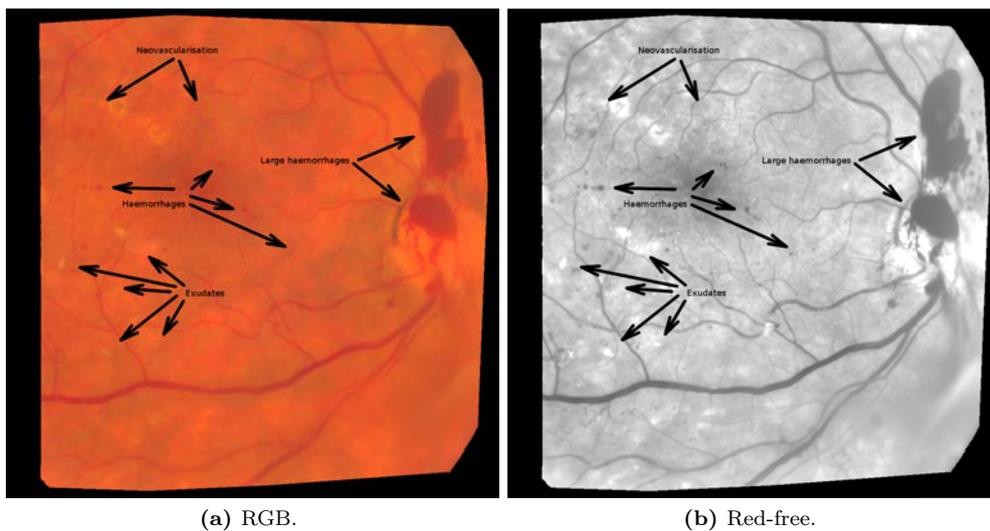
Chapter 5, Discussion, presents the implications of the results, and the future work related to the content of the thesis.

Chapter 6, Conclusion, summarises the goals, methods, experiments and results, and concludes the thesis.

Spectral fundus imaging and spectral image composition

2.1 Introduction

Fundus imaging offers a non-invasive view to the eye fundus. A typical modern fundus camera consists of a light source, a digital camera, and microscope optics for both projecting the illumination onto the eye fundus and guiding the light reflected from the eye fundus into the camera. Due to its ease of use and relative inexpensiveness, digital fundus imaging remains the standard method for diagnosing diseases of the eye fundus, such as DR, AMD and glaucoma. Typically either RGB images or grayscale images taken with a red-free filter are used (see Figure 2.1).



(a) RGB.

(b) Red-free.

Figure 2.1: Example fundus images from DiaRetDB2.

Other modalities such as OCT, MRI and angiography are available for cases where traditional retinal images are not sufficient for diagnosis or treatment planning, MRI and OCT providing a non-invasive view to the inner structure of the retina, and angiography providing a view of the vasculature and retinal blood flow. These, however, require specialised, and often expensive, equipment which are not available in all diagnosis centres.

Spectral fundus cameras capture images of the retina with a significantly higher wavelength resolution than traditional RGB images. Spectral retinal images are, in short, 3-dimensional matrices with two spatial and one spectral dimension. Instead of one intensity channel of a grayscale, or the three colour channels of an RGB image, the spectral dimension of a spectral image consists of several, tens or even hundreds of channels, depending on the imaging system. They can provide information on the eye fundus beyond that of the traditional fundus cameras with low additional cost or requirements for the operator.

The analysis of fundus reflectance spectra has been used in a number of medical applications. Retinal reflectance spectroscopy has been used to evaluate the oxygen saturation of retinal blood (e.g., [87, 202]), for the estimation of the concentrations of xanthophyll, melanin and haemoglobin in the retina and choroid (e.g., [84]), and for determining the optical density of macular pigment (e.g., [20]).

Spectral images have been used to produce improved visualisations of clinically interesting structures and pathologies. Fält et al. [55] suggest directly modifying the illumination spectra to optimise the contrast between various retinal structures or lesions, and the fundus background. Another approach to enhanced visualisation of retinal structures is presented in [57], where the contrast between the structures and the background is optimised by assigning different weights to the individual channels of the spectral images.

Due to the significant increase in the colour resolution, spectral retinal images offer a richer feature space for automatic detection, classification and diagnosis. Thus, the use of spectral fundus images has the potential to significantly increase the performance of automatic diagnostics.

2.2 Related work

The eye has been studied extensively, and a significant body of work on imaging, measuring and modelling the eye fundus exists. Various approaches to imaging and quantifying the structures, both *in vivo* and *in vitro* (usually from animals), and or functional parameters (e.g., blood oxygen saturation) have been developed.

Berendschot et al. [19] review historical and modern instruments for the measurement and applications of fundus reflectance. A number of reflectometers, fundus imaging and video systems, and scanning laser ophthalmoscopes are presented. The paper includes a review of reflectance models for various parts of the eye derived from the measurements of the different systems, as well as the evaluation of the retinal microstructure such as macular pigment (MP) density, melanin content and retinal blood oxygenation based on the reflectance measurements.

Historical and present approaches to the MRI of the retina are presented by Duong [53]. While the emphasis is on MRI performed on animals, some studies on MRI of the human retina are presented. A more recent review of techniques and instruments used in

ophthalmology is presented by Keane and Sadda [119]. In addition to retinal imaging, techniques such as adaptive optics, angiography, and spectral imaging are presented. A number of methods for retinal OCT are also introduced.

Retinal reflectometry has been used to obtain measurements of various retinal absorbers and to study the reflectance of retinal structures. Delori and Pflibsen [51] used a reflectometer based on a modified Carl Zeiss fundus camera to capture the reflectance spectra at nasal fundus, perifovea and fovea of ten healthy subjects. A fundus reflectance model including ocular media, inner limiting membrane, photoreceptor and retinal pigment epithelium (RPE) layers, Bruch's membrane, choriocapillaris, choroidal stroma and sclera was derived from the reflectance measurements.

Kaya et al. [118] used fundus reflectance to compare the optical density of MP between patients with AMD and healthy subjects. The fundus reflectance was measured and the optical density estimated using the system and model in [251]. The optical density of MP was found to be reduced for patients with AMD.

Berendschot et al. [21] measured the fovea of 435 subjects of age 55 and older using a fundus reflectometer to determine whether age-related maculopathy affected the optical density of MP and/or melanin. No differences were found between healthy subjects and subjects with any stage of age-related maculopathy.

Van de Kraats et al. [229] studied the interaction between light and photoreceptor layer of the eye to derive a model of the spectral, directional and bleaching properties of the fovea using the retinal densitometer described in [232]. The model was validated by comparing the visual pigment density estimated using the model with results from psychophysical experiments.

To measure the reflectance spectrum over a specific region of the retina, a number of instruments for retinal spectroscopy have been developed. Schweitzer et al. [200] presented a method for measuring the oxygen saturation retinal reflectance spectra. Using a Carl Zeiss CS 250 adapted with a Jobin Yvon CP 200 spectrograph, reflection spectra from line scans over retinal vessels can be acquired. A model based on the transmission of oxygenated and deoxygenated blood was used to estimate the retinal blood oxygenation levels from 30 eyes. The mean oxygen saturation was found to be 92.2% for arteries and 57.9% for veins.

Delori [49] presented a spectrophotometer capable of both inducing fluorescence and capturing the reflected or fluorescent light from the fundus. Utilising a motorised filter wheel placed after a 150-W xenon-arc lamp, the system is capable of producing excitation at wavelengths between 430 nm and 550 nm. A neutral filter is included for traditional reflectance measurements. The same optical setup is used for capturing the reflected or fluorescence light.

Zagers et al. [251] described an apparatus for simultaneously measuring the spectral reflectance of the fovea, and the directionality of cone-photoreceptors. A least-squares fit of the model described in [229] to the measured spectra was performed for the purpose of evaluating the densities of photostable ocular absorbers.

Retinal spectroscopy has been used to acquire measurements of various retinal structures. Delori and Burns [50] measured the absorption of the crystalline lens of the human eye *in*

vivo on 148 eyes of varying age and retinal health, using a fundus spectrometer. The spectra acquired by the spectrometer were corrected for lens back-scatter and fluorescence, and instrument noise using an additional baseline measurement with the illumination field in a different position on the retina. Lens density was estimated from the measured spectra.

Savage et al. [197] compared different non-invasive measurements of the optical density of the ocular media of 41 healthy subjects. An objective measurement of the spectral transmission of the lens is gained by comparing the intensity of the reflectance from the posterior surface of the lens to an external reference on eight wavelengths. The results of the objective measurement were compared with those from a psychophysical procedure with low-light condition brightness-matching of the halves of a bipartite field after 15 min dark-adaptation. The two approaches were found to correlate well for the shortest measured wavelength, but not at longer wavelengths.

Bone et al. [29] measured the distributions of macular pigment, photopigments and melanin in the retina. They used a Topcon TRC NW5SF non-mydratic retinal camera with the original exciter filter replaced with two multiband interference filters to acquire reflectance maps at wavelengths, where the density of the pigments can be estimated based on the amount of light they absorb.

Salzer et al. [194] studied the diffuse spectral reflectance of the fundus using a Spectralon reflectance target inserted into the eye of domestic swine. The target placed in the eye under the retina was image *in vivo*. Spectral images of the fundus with reflectance target were acquired using narrow-band illumination at a number of different central wavelengths.

A number of systems for acquiring spectral retinal images can be found in the literature. Fawzi et al. [60] presented an instrument for fast hyperspectral retinal imaging. The system uses computed tomography to reconstruct images from spectra acquired by an imaging spectrometer attached to a fundus camera. The acquired spectra were used to recover MP optical density using spectra measured *in vitro* as a prior.

Retinal blood oximetry has been presented as either the motivation for or the example use case of many of the spectral retinal imaging systems. Beach et al. [17] described a modified fundus camera with optics dividing the light reflected from the retina to two separate band-pass filters to acquire simultaneous dual-wavelength images. The dual-wavelength images, where one filter is centred at a wavelength where the difference between the spectra of oxygenated and deoxygenated blood is significant, and the other where the difference is minimal, were used in retinal oximetry.

In [88], Harvey et al. propose a spectral imaging system capable of acquiring a multi-spectral image in a single exposure. An optical system of polarising beam splitters and waveplates (a plate that alters the polarisation state of the transmitted light) is used to separate the desired wavelengths and to guide them to different parts of a sensor array. As the system projects the wavelength channels to different locations on the same sensor array, spectral resolution of the acquired spectral image comes at the cost of the FOV of the system. The system has been used to study the effect of acute mild hypoxia on retinal oxygen saturation [42].

Hirohara et al. [95] validated their spectral fundus imaging system via oxygen saturation analysis. The imaging system consisted of a Topcon TRC-50LX fundus camera fitted

with a VariSpec liquid crystal tunable filter. The system is capable of acquiring images in the range 500 nm to 720 nm with 10 nm steps. The validation was performed comparing the spectra from imaging to the spectra measured from artificial capillaries with known blood oxygenation levels.

Ramella-Roman et al. [182] presented a multiaperture system for acquiring spectral fundus images for estimating the oxygen saturation of the retinal blood. A lenslet array is utilised to project the light passing through an array of narrow-band filters to specific locations of a charge-coupled device (CCD) array. The system is capable of simultaneous acquisition of fundus images at six different wavelengths.

Mordant et al. [149] use a spectral imaging system based on a liquid crystal tunable filter for retinal blood oximetry. By nonlinear fitting of the acquired image spectra to a model of (wavelength-dependent) optical density of oxygenated and deoxygenated haemoglobin, the ratio of blood oxygenation is estimated at each point of the spectral image corresponding to a blood vessel. In [150], Mordant et al. validate the performance of their approach to blood oximetry. The validation was performed by placing samples of human blood, with reference oxygen saturations measured with a CO oximeter, into quartz tubes placed inside a model eye. The mean difference between the measured reference and the estimated oxygenation was found to be approximately 5%.

Rodmell et al. [191] study light propagation through the retina using Monte Carlo simulation. The paper concludes that illumination at the edges of the vessel, and detection directly above the vessel result in the capture of light that has made only a single pass through the vessel. This has relevance in retinal oximetry where light interaction with other retinal tissue can affect the reflected spectrum and influence the estimated oxygen saturation values.

Based on the reviewed literature, the properties of retinal structures and molecules have been studied largely using retinal reflectometry and spectroscopy. An emphasis on the measurement of the properties of ocular media and retinal absorbers can be found. Various approaches to spectral retinal imaging have been proposed, with an emphasis on retinal blood oximetry. Partly due to the multitude of approaches for image acquisition, general calibration and signal processing to acquire accurate spectra have received limited attention. Calibration and image processing are typically specific to an individual image acquisition system or measurement. A table summarising the presented literature is shown in Table 2.1.

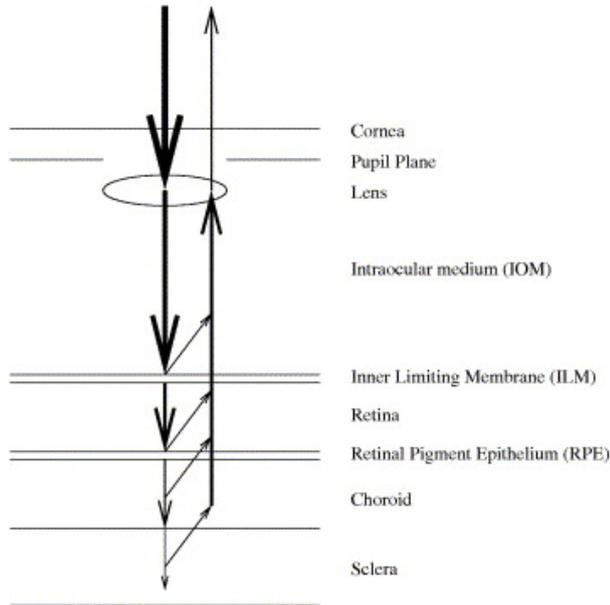
2.3 Spectral fundus image formation

While relevant clinical knowledge on the eye fundus and its pathologies may be sufficient to analyse traditional fundus images, to properly understand the characteristics of the spectral fundus image data, insight into the process of spectral fundus image formation is required. As the spectrum of the light reflected from the fundus is affected by the interaction with retinal tissue, structural changes in the retina due to pathologies change the measured spectrum. However, visual inspection of the spectra is generally not useful diagnostically, and any individual channel is unlikely to be sufficient to identify a specific change in structure.

Table 2.1: Summary of literature review.

Category	Modality	Molecule/microstructure	Property	Subjects	Year	Reference(s)
Review	Various				2003	[19]
Review	Various				2014	[119]
Review	MRI				2011	[53]
Instrument	Spectroscopy	Vasculature	Oxygenation		1999	[17, 200]
Instrument	Spectral retinal images	Vasculature	Oxygenation		2005-2011	[88, 95, 149, 150, 182]
Instrument	Spectroscopy, fluorescence				1994	[49]
Instrument	Spectral retinal images	Macular pigment	Optical density		2011	[60]
Instrument	Spectroscopy	Ocular absorbers	Density		2002	[251]
Measurement	Reflectometry	Macular pigment	Optical density	181	2012	[118]
Measurement	Reflectometry	Macular pigment, melanin	Optical density	435	2002	[21]
Measurement	Reflectometry	Various	Reflectance	10	1989	[51]
Measurement	Spectroscopy	Various	Distribution	22	2007	[29]
Measurement	Spectroscopy	Fundus	Reflectance		2008	[194]
Measurement	Densitometry	Fovea	Various	10	1996	[229]
Measurement	Spectroscopy	Lens	Absorption	148	1996	[50]
Measurement	Spectroscopy	Ocular media	Optical density	41	2001	[197]

The structures of the eye have a complex effect on the spectrum of the light that is reflected from the eye fundus (see Figure 2.2). Longer wavelengths penetrate deeper into the fundus, resulting in different tissue interactions than shorter wavelengths. Retinal tissues have significantly different optical properties, with various degrees of absorption, scatter and refraction.

**Figure 2.2:** Light paths in the retina. [216]

The paths the photons take through the retinal tissue before being reflected back to the detector have a significant, non-linear effect on the resulting reflectance spectrum. As some of the photons are reflected from the interfaces and inside the tissue layers, the contribution of a single layer on the emitted spectrum is difficult to determine.

To the knowledge of the authors, no comprehensive physical model of light interaction

with the eye exists. As the reflectance spectrum is the result of reflection, absorption and back-scatter from multiple different layers with various optical properties, and accurate reference measurements are difficult to obtain, the interactions become difficult to model properly. However, computational models of the light interaction in retinal tissue have been proposed (e.g., [48, 81, 181, 231]).

2.3.1 Structure of the eye

The human eye is a complex organ, both functionally and structurally. The eye is composed of various tissues and media, with significant differences in how they interact with light entering the eye. This section provides a short description of the different parts of the eye, their function and optical properties.

CORNEA AND OCULAR MEDIA

Cornea is the transparent outermost part of the eye. It helps protect the eye from external, often harmful forces, and refracts light to provide a larger field of vision. Behind the cornea are the (near-) transparent parts of the eye that allow the light entering the eye to be transmitted onto the retina. The transparent ocular media located between the cornea and the eye fundus can be divided into aqueous humor, lens and vitreous humor.

While mostly transparent in the longer wavelengths, the lens absorbs strongly in the near-ultraviolet and short wavelengths. Furthermore, the absorption of the lens changes with time, the lens becoming more yellow as the person becomes older. [179]

RETINA

The retina consists of several layers with different structures and functions (see Figure 2.3). The main functionality related to the sensing of light is located at the retina.

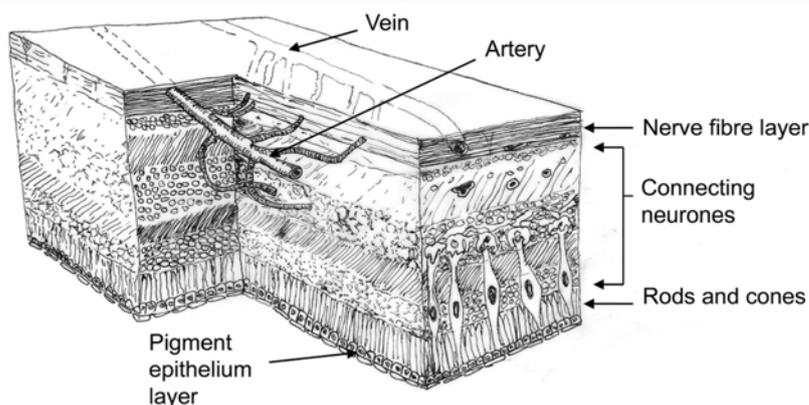


Figure 2.3: Retinal layers. [219]

Neural retina is the outermost layer of the eye fundus, located between the ocular media and the retinal RPE layer. The neural retina contains the photoreceptors that are

responsible for converting the photons striking the retina into neural responses to be processed by the visual system.

The inner segment/outer segment (ISOS) junction is a structure inside the neural retina that is assumed to originate from the boundary between the inner and the outer segment of the photoreceptor [243]. While a separate functional part of the retina, the ISOS junction interacts with photons passing through the neural retina.

The RPE layer, located below the neural retina, is a pigmented layer that absorbs a large portion of the scattered light in the retina, reducing false photoreceptor activations. The RPE also protects the retina from photo-oxidation and subsequent oxidative damage, and take part in many essential processes such as metabolism of the visual pigments, phagocytosis of the photoreceptor outer segments, formation of the blood-retinal barrier and homeostasis of the retinal micro-environment by producing growth factors regulating the vital functions like angiogenesis and vascular bed maturation.

CHOROID AND SCLERA

The choroidal layer contains connective tissue and vasculature. The choroid is responsible for the blood supply of the outer parts of the retina. While not directly a part of the formation of visual stimuli, the choroid is vital for healthy vision as it provides parts of the retina with nutrients and oxygen. [90].

The sclera is the white matter of the eye. It forms and maintains the shape of the eyeball. The sclera connects the optical system it surrounds to the muscles responsible for the movement of the eye.

2.3.2 Modelling of light interaction with retinal tissue

The model of light interaction in retinal tissue described in this section extends the model by Styles et al. [216]. While general structure of the model remains the same, a layer modelling the cornea is added, and the transmittance values of ocular media are altered.

The retinal interaction model discussed in this work is built upon the Monte Carlo (MC) model of light transport in multilayered tissue by Wang et al. [239]. The model simulates the transport of an infinitely narrow photon beam in a multilayered tissue of infinite width, with the beam perpendicular to the tissue surface. Each layer of the tissue is characterised by its thickness, refractive index, absorption and scattering coefficients, and anisotropy factor. A flowchart of the simulation process is shown in Figure 2.4.

At each iteration, the photon takes a step of the size s in the propagation direction ϕ (initially perpendicular to the tissue layer) before tissue interaction. The step size is defined

$$s = \frac{-\log \xi}{\mu_t}, \quad (2.1)$$

where ξ is a random number between $[0, 1]$. μ_t is a tissue interaction coefficient defined as $\mu_t = \mu_a + \mu_s$, where μ_a and μ_s are the tissue absorption and scattering coefficients. The photon position is updated by

$$\begin{aligned} \hat{x} &= x + \phi_x s \\ \hat{y} &= y + \phi_y s \\ \hat{z} &= z + \phi_z s \end{aligned} \quad (2.2)$$

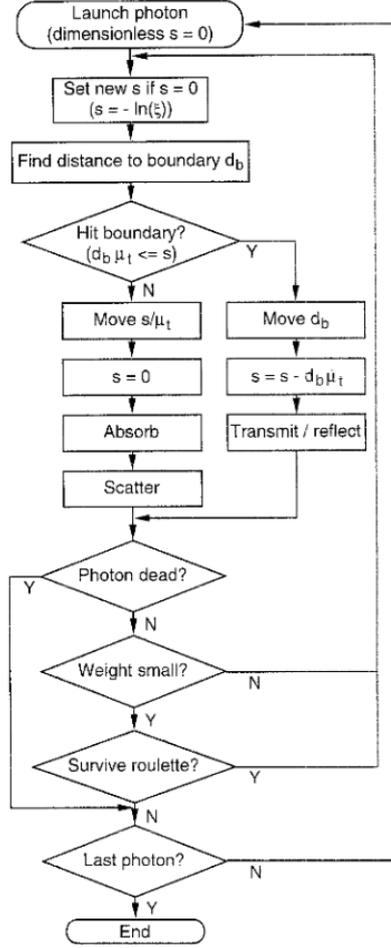


Figure 2.4: MC modelling flowchart. [239]; ξ is a random number (uniform distribution) between $[0, 1]$, s is the path length the photon can travel before tissue interaction, μ_t is the tissue interaction coefficient, d_b is the distance between the photon and the boundary of the current tissue layer along the direction the photon is travelling.

after which photon interaction with the tissue is simulated.

The photon interacts with the tissue by undergoing absorption and scattering. Absorption reduces the energy of the photon, E , by

$$\dot{E} = E - \frac{\mu_a}{\mu_t} E, \quad (2.3)$$

where μ_a is the tissue absorption coefficient and μ_t is the interaction coefficient of the tissue. After absorption the photon undergoes scattering, affecting the direction of the

photon propagation. The new propagation direction after scattering becomes

$$\begin{aligned}\dot{\phi}_x &= \frac{\sin \theta (\phi_x \phi_z \cos \Phi - \phi_y \sin \Phi)}{\sqrt{1 - \phi_z^2 + \phi_x \cos \theta}} \\ \dot{\phi}_y &= \frac{\sin \theta (\phi_y \phi_z \cos \Phi + \phi_x \sin \Phi)}{\sqrt{1 - \phi_z^2 + \phi_y \cos \theta}} \\ \dot{\phi}_z &= \sin \theta \cos \Phi \sqrt{1 - \phi_z^2} + \phi_z \cos \theta,\end{aligned}\quad (2.4)$$

where Φ is a randomly sampled azimuthal angle defined as $\Phi = 2\pi\xi$. The deflection angle θ is dependent on the anisotropy of the tissue layer, and is defined as

$$\cos \theta = \begin{cases} \frac{1}{2g} \left\{ 1 + g^2 - \left[\frac{1-g^2}{1-g+2g\xi} \right]^2 \right\} & \text{if } g \neq 0 \\ 2\xi - 1 & \text{if } g = 0 \end{cases}, \quad (2.5)$$

where g is the anisotropy factor of the current tissue layer.

During step s , the photon may encounter a tissue boundary. The distance to the closest tissue boundary in the photon propagation direction is defined

$$d_b = \begin{cases} (z_0 - z)/\Phi_z & \text{if } \Phi_z < 0 \\ \infty & \text{if } \Phi_z = 0 \\ (z_1 - z)/\Phi_z & \text{if } \Phi_z > 0 \end{cases} \quad (2.6)$$

where z_0 and z_1 are the z coordinates of the tissue boundaries above and below the current photon position. If the size of the evaluated step s is greater than the distance to the closest boundary, i.e., $d_b \mu_t \leq s$, the current step size is reduced to $s = s - d_b \mu_t$ and interaction with the tissue boundary is simulated.

Depending on the angle of incidence, $\alpha_i = \cos^{-1}(|\mu_z|)$, the photon has a chance to be either transmitted or internally reflected. If α_i is greater than the critical angle $\sin^{-1}(n_t/n_i)$, where n_i and n_t are the refractive indices of the media that the photon is incident from and transmitted to, the likelihood of internal reflection, $R(\alpha_i)$ is 1. Otherwise $R(\alpha_i)$ is defined

$$R(\alpha_i) = \frac{1}{2} \left[\frac{\sin^2(\alpha_i - \alpha_t)}{\sin^2(\alpha_i + \alpha_t)} + \frac{\tan^2(\alpha_i - \alpha_t)}{\tan^2(\alpha_i + \alpha_t)} \right] \quad (2.7)$$

where α_t is the angle of transmission, defined as

$$\alpha_t = \sin^{-1} \left(\frac{n_i \sin \alpha_i}{n_t} \right). \quad (2.8)$$

Whether the photon is internally reflected or transmitted to new layer is based on a random number ξ . If $\xi \leq R(\alpha_i)$ the photon is reflected, otherwise it is transmitted to a new layer. In the case of internal reflection, the photon propagation direction is mirrored, i.e., $\dot{\Phi}_z = -\Phi_z$. In the case of transmission, the propagation direction is changed according to

$$\begin{aligned}\dot{\Phi}_x &= \Phi_x n_i / n_t \\ \dot{\Phi}_y &= \Phi_y n_i / n_t \\ \dot{\Phi}_z &= \begin{cases} \cos \alpha_t & \text{if } \Phi_z \geq 0 \\ -\cos \alpha_t & \text{if } \Phi_z < 0 \end{cases}\end{aligned}\quad (2.9)$$

At the end of each propagation step, the energy remaining to the photon is compared against a minimum energy threshold E_{th} . If $E < E_{th}$, the photon has a chance (independent of the remaining energy) of being annihilated. If the photon fails the annihilation test, the propagation stops and the propagation of a new photon is started. Otherwise the propagation continues as before. Following these rules, the photon propagation continues until the photon either escapes the media, and its remaining energy is added to either reflection or transmittance (in the fundus model only reflection is considered), or the photon is randomly annihilated after its energy is reduced to zero. Typically a large number of photon propagations are simulated, and the sum of the weights of the photons that escaped the media form the resulting reflectance spectrum. Examples of photon paths are illustrated in Figure 2.5.

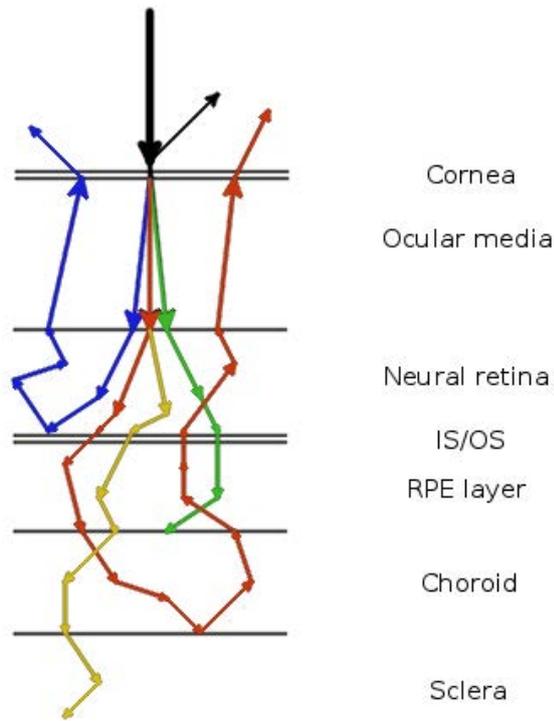


Figure 2.5: MC modelling of spectrum formation.

MODEL PARAMETER SELECTION

As the tissue layers of the model are characterised by thickness, refractive index, absorption and scattering coefficients, and anisotropy factor, the selection of these parameters is crucial for a realistic model of fundus image formation. As no single study of the optical properties of the eye containing estimates for all the required parameters exists to the knowledge of the author, the parameter values were selected based on a variety of studies.

Hammer et al. [83] used the double-integrating-sphere technique to measure the collimated and diffuse transmittance, and diffuse reflectance of the retina, RPE, choroid, and sclera layer of the eye fundus. From the measured reflectance and transmittance spectra, the absorption and scattering coefficients, and the anisotropy of scattering were estimated by inverse MC simulation.

The corneal refractive index used in the model is derived from Fitzke III [68]. The mean of the individual values of the epithelium, stromal anterior and posterior surfaces of the cornea was used to represent the refraction in the cornea.

The transmittance for the ocular media were taken from Boettner and Wolter [28], who measured the transmission in human ocular media *in vitro* from freshly removed eyes. Both the total transmittance and the transmittance of the individual media, cornea, aqueous humor, lens, vitreous humor, were measured. The refractive index reported in [124] was used for vitreous.

The yellowing of the lens is dependent on the age of the patient and has to be considered separately from the model generation. The average lens transmission function for lenses of different ages from [179] was used to correct the simulated spectra to account for the age related lens yellowing. The model spectra were corrected individually based on the age of the patient whose spectral image data was analysed using the model.

Two neural retina layers with identical parameters were used to enable the insertion of a layer simulating the interface between the neural retinal tissue and the photoreceptors within the neural retina layer. The free model parameters retinal haemoglobin and macular pigment density are present in the neural retina layer. Refractive indices for retina reported by Knighton et al. [124] were used.

An estimate of the refractive index and scattering of the ISOS layer is derived based on the physical and biological properties of the ISOS junction [43]. The absorption coefficient was assumed to follow that of the neural retinal layer.

The RPE layer holds retinal melanin, the distribution of which is a free parameter in the model. An estimate for the refractive index of the RPE layer is derived by Hammer et al. [85] from literature and OCT measurements.

No reported values of the refraction index of choroid were found in literature. However, no experimental evidence (i.e., a reflection in OCT scan indicative of an interface between layers with different refractive indices) was found of difference in refractive indices between choroid and sclera. There was assumed to be no (significant) difference in the refractive indices of the choroid and the sclera. The similar (collagen matrix) structure of the layers would also support this assumption. [43]

The sclera is the final layer simulated in the model. Any light transmitted through the sclera is considered to be completely absorbed or scattered, as the amount of light surviving back to the detector after passing to layers below the sclera can safely be assumed to be negligible. The refractive index of the sclera reported by Bashkatov et al. [15] was adopted for the model.

In addition to the characteristic optical properties of the individual layers, the main contributors to the formation of the spectra are the haemoglobins and melanin, both of which are strong absorbers, and the thickness of the layers. The model values were

taken from literature: haemoglobins from Horecker [98], melanin from Anderson and Parrish [10], and layer thickness from Rohen [192]. An 80% oxygen saturation level was assumed for the haemoglobins based on Alm and Bill [9].

The absorption, scatter and refractive indices for the different layers are considered as constant. The model has five free parameters that can vary within histologically plausible limits: concentration of macular pigments in the retina, concentration of haemoglobins in the retina, concentration of melanin in the RPE, concentration of melanin in the choroid and concentration of haemoglobins in the choroid.

The estimations of the optical characteristics of the retinal molecules RPE melanin and macular pigment, and haemoglobin can be expected to be relatively accurate as they can be expected to stay constant between individuals and can be measured in laboratory conditions. To a lesser degree, similar assumption can be made regarding the cornea, the ocular media, and the tissues of the neural retina, RPE, choroid and sclera.

The thickness of the different layers, however, is a subject to greater individual variation. Another potential source of inaccuracy is the level of haemoglobin oxygenation. The level of oxygenated blood is affected by the phase of circulation, the size of blood vessels at (or near) the location, and changes in circulation due to disease. The model also expects majority of the retinal tissue to be free of pathologies and not (significantly) affected by any systemic disease. As it is not possible to determine what the values of these parameters were at the time of the acquisition of a spectral retinal image, it is difficult to measure the representativeness of the values used in this work.

2.4 Spectral fundus image acquisition

A number of spectral fundus imaging systems have been developed (e.g., [18,54,106,166]). This thesis considers the composition and applications of the spectral images from two spectral fundus imaging systems with significant differences in both the image acquisition approach and the desired features for the data.

2.4.1 30-channel spectral fundus camera

Fält et al. [56] modified a Canon CR5-45NM fundus camera system to acquire spectral images of the eye fundus. Leaving the original fundus microscope optics, the camera of the system was replaced by a QImaging Retiga 4000RV digital monochrome CCD camera. A rail for a filter rack and a placement for an optical cable were fitted to the camera casing. The original light source was replaced by broad-band illumination from an external Schott Fostec DCR III lightbox with a 150 W OSRAM halogen lamp using a daylight-simulating filter, guided to the camera system by a fibre optic cable. The system is shown in Figure 2.6.

The setup contains four acrylic glass filter racks with a total of 30 Edmund Optics narrow bandpass filters with central wavelengths in the range 400 nm to 700 nm. The filters are changed manually by sliding the filter rack along the rail, with a mechanical stopper ensuring that each filter is correctly positioned after moving the rack. The broad-band light exiting the cable is filtered by the selected narrow-band filter and guided to the eye

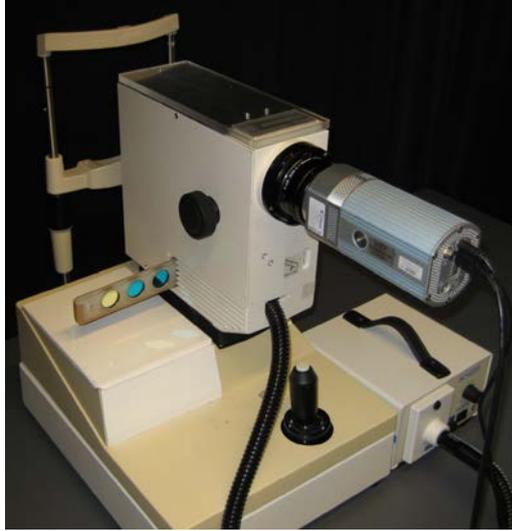


Figure 2.6: Spectral camera system by Fält et al [56].

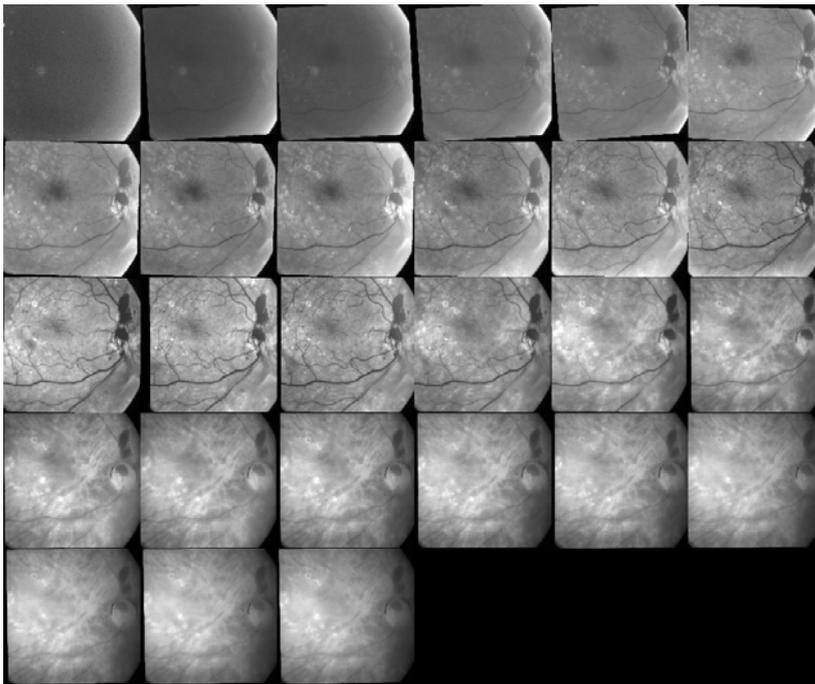


Figure 2.7: Montage of channel images acquired with the system by Fält et al [56]. Images normalised for visualisation.

fundus through the camera optics. The reflected light captured by the camera system represents the fundus reflectance for that wavelength. An example is shown in Figure 2.7.

A suitable exposure time was estimated individually for each filter from the area in the retina with the highest reflectivity (typically the optic disk). For each filter, five successive channel images were acquired to avoid motion blur or significant difference in the imaging angle due to eye movement. After a qualitative evaluation, the highest quality image at each wavelength was selected and the images were automatically aligned using the algorithm by Stewart et al [214]. Manual registration was performed for the image pairs for which the automatic alignment failed. The registered spectral channel images were composed into a spectral image with each channel normalised to unit exposure time (i.e., 1 s).

2.4.2 Six-channel spectral fundus camera

Styles et al. [216] modified a Zeiss RCM250 fundus microscope. The original camera body was replaced by a QImaging Retiga EXi 12-bit monochrome digital camera and a Cambridge Research Instruments VariSpec LCD programmable filter was added in front of the camera, with an additional lens to fit the image to the 1/3 inch CCD sensor array of the Retiga EXi, which is significantly smaller than the original 35 mm film. A halogen lamp was used to illuminate the fundus through the camera optics instead of the original xenon flash. The xenon flash was considered unsuitable due to sharp emission peaks in its illumination spectrum, and the transient (instead of steady-state) nature of the provided illumination. The setup is shown in Figure 2.8.

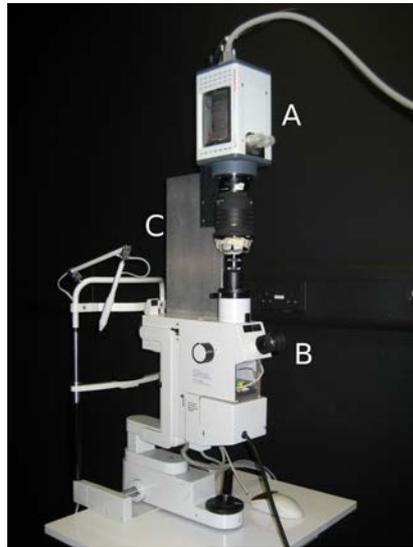


Figure 2.8: The spectral camera system by Styles et al [216].

The VariSpec LCD programmable filter is a configurable interference filter capable of implementing Gaussian narrow-band filters with central wavelengths in the range 400 nm to 700 nm. The spectral image is composed of six sequentially acquired channel images, filter central wavelengths 507, 525, 552, 585, 596 and 611 nm (the selection of the wavelength is related to the application and is discussed in detail in Chapter 4.3), with each channel image normalised to 1 s exposure time. An example is shown in Figure 2.9.

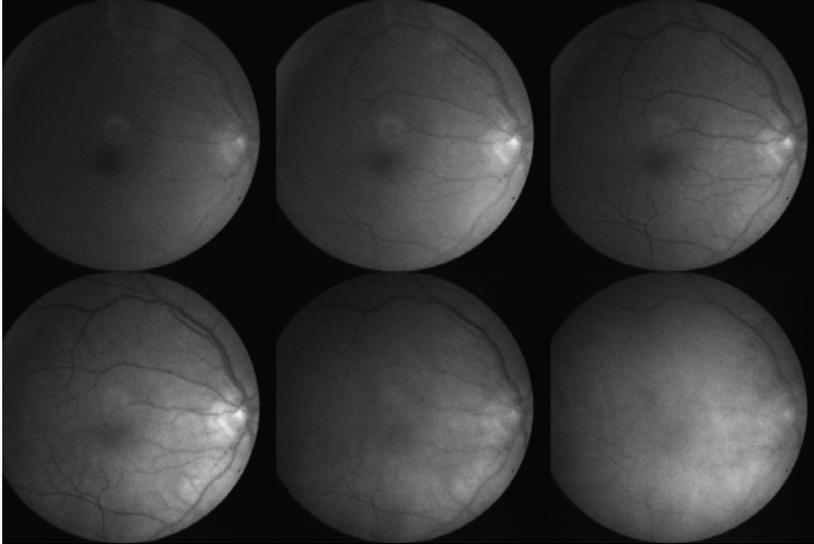


Figure 2.9: Montage of channel images acquired with the spectral camera system by Styles et al [216]. Images were normalised for visualisation.

Further development of the spectral fundus camera system is described in [54]. The halogen white light source was replaced with a light source composed of 12 programmable light emitting diodes (LEDs). LEDs of different emission spectra can be individually addressed, allowing the precise control of intensity, illumination time, and the sequence of illumination.

The total acquisition time for a set of channel images was 0.5 s. To minimise eye movement between the acquisition of the channel images, three image sets were acquired consequently for a high probability of capturing at least one set that contains no movement. The absence of inter-channel movement was confirmed by registering the images using the method by Stewart et al. [214] and examining the resulting transformation. If the transformation required to align the images was below 2.3 pixels, any eye movement present in the images was deemed to fall within the system error and the spectral image composed of the set of channel images was accepted. The system error was derived from the maximum registration error over a set of images, acquired using the system, where no observable eye movement was present.

The system providing the data used in this thesis is a modification of the spectral fundus camera presented in [54]. While attaining a short acquisition time, the LED illuminant of the system caused refraction patterns to appear in the channel images. The LED light source was replaced with a white light source and a VariSpec LCD filter.

2.5 Spectral camera calibration

Fundus cameras offer a non-invasive view to the ocular fundus and are an important tool for diagnosing a number of eye and systematic diseases, e.g., AMD and DR [1]. A

fundus camera system has several independent components and their characteristics that contribute to the features and quality of the acquired image. These include the sensor, the light source and the optics, both optics for guiding the light from the light source to the eye and to guide the reflected light to the camera, with attributes that are often not (accurately) known. Due to the small size and the proximity of the target (i.e., the eye), and the magnification of the eye lens, special optics are used to acquire images with a reasonable field of view, making radial distortions [136] and vignette (i.e., the decrease of image intensity values towards image edges) [96] prevalent in fundus images.

There are a number of fundus camera characteristics that should be taken into account when analysing the acquired images. The wide-angle optics cause increasing deformation to perceived objects as their distance to the principal point of the image increases [217], which is likely to cause error in measurements, complicate image registration and result in an accumulation of error when compiling longitudinal data or data from multiple sources. For any dimensional measurement of fundus features (absolute measurements are important for certain clinical purposes such as the classification of AMD [207]), the spatial resolution of the image has to be known or estimated. An uneven illumination field may hinder the diagnosis and statistical classification and segmentation based on pixel intensities, and cause problems with longitudinal data. Dirt, dust and stains on or inside the optical system of a camera cause artifacts to the images acquired by the system. The artifacts can cause false positive detections in an automatic analysis algorithm or even be misclassified as lesions by a human analyst. When combining data from different imaging systems, the accumulation of the artifacts may have unforeseen consequences if not taken into account.

In the case of spectral imaging, the error to the spectra due to uneven illumination distribution can be significant. Furthermore, as light passes through the multiple lenses of the optical system in a fundus camera, wavelength-dependent differences in the refraction indices of the lens materials and coatings may cause aberrations at different wavelengths of the captured light. While not a significant issue in grayscale or RGB-imaging, the spectral aberrations may cause significant errors in the captured spectra.

Quantifying the effects of and the distortions caused by the imaging system on the image data becomes especially important in the case of longitudinal studies. When studying the retinal changes or the progression of a pathology over years or even decades, the imaging parameters, protocols and even equipment are likely to change between the examinations. If the imaging systems are not properly characterised and calibrated, it may be difficult or even impossible to differentiate between the changes in the data due to changes in the clinical condition and changes due to differences in the data acquisition.

This section presents a protocol for calibrating a fundus camera, with special consideration to spectral fundus cameras. The calibration steps include geometric and spectral calibration, determining the spatial resolution, consideration for correcting uneven illumination and vignetting, and accounting for dirt and scratches in the optics. Practical examples of calibrating the interference filter based spectral camera system by Fält et al. [56] will also be discussed.

2.5.1 Related work

Xu and Chutatape [247] compare the errors of two calibration methods for a fundus camera, one method based on a 3D target and the other on a planar calibration target. The method using the planar calibration target was found to produce more stable and accurate results for fundus camera calibration.

In [141], Lujan et al. use spectral domain optical coherence tomography (SD-OCT) to calibrate fundus cameras by determining the distance between the optic nerve and the centre of the fovea from both the SD-OCT scans and the fundus image, giving the same measurement in millimetres and pixels.

Deguchi et al. [46] calibrate a fundus camera by utilising a transparent acrylic plate with a regular grid painted on both sides with different colours. Using the imaged grid points, the lines passing through the calibration planes are identified and used to account for the optical distortions of the camera when constructing a 3D reconstruction of the fundus from stereo images.

Martinello et al. [145] discuss the calibration of a stereo fundus camera and models required for estimating the distortions caused by the lens system, in the context of 3D-reconstruction of the eye fundus.

Spectral calibration of a fundus camera is discussed by, e.g., Ramella-Roman et al. [182], who use Spectralon reflectance standards to determine the effect of their camera and filter system on the acquired spectra.

The majority of the work including fundus camera calibration seems to have a focus outside calibration, and deal with calibration only to the degree it is relevant to the specific goal of the work. This section presents a general protocol for fundus camera calibration, encompassing imaging system characteristics necessary to be determined when analysing longitudinal data, or data from multiple sources or imaging systems.

2.5.2 Methods

CORRECTION OF GEOMETRIC DISTORTIONS

Imaging a calibration target with a regular pattern with known dimensions, the camera parameters and lens distortions can be approximated. While a planar calibration pattern cannot represent all the distortion present in retinal images, as the outer parts of the eye and individual retinal curvature contribute to the distortion, the distortion caused by the camera system can be characterised and corrected. This is important when dealing with data acquired by different camera systems with different distortion characteristics.

If significant vignette is present, the illumination field of the images may need to be corrected (see Section 2.7 for details) to properly extract the reference points, such as corner points or grid centroids, from the calibration target. A corner detector or thresholding can then be applied to extract the reference points.

Knowing the grid centroid locations in the image space and the dimensions of the physical target, the intrinsic camera parameters including the principal point, focal length, and radial and tangential distortion can be estimated using the calibration approach

of Zhang [253], with the intrinsic camera model from Heikkilä and Silvén [91] for more accurate modelling of the distortion. The model by Heikkilä and Silvén is defined as

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} d_u s_u (\tilde{u} + \delta u_i^{(r)} + \delta u_i^{(t)}) \\ d_v (\tilde{v} + \delta v_i^{(r)} + \delta v_i^{(t)}) \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \quad (2.10)$$

where (u_0, v_0) are the image centre coordinates, (\tilde{u}, \tilde{v}) are the points projected onto the image plane by a pinhole camera model, d_u and d_v are the pixel width and height and s_u is the camera (intrinsic) scale factor. The radial distortion terms, $\delta u^{(r)}$ and $\delta v^{(r)}$, and the tangential distortion terms, $\delta u^{(t)}$ and $\delta v^{(t)}$, are defined as

$$\begin{bmatrix} \delta u^{(r)} \\ \delta v^{(r)} \end{bmatrix} = \begin{bmatrix} \tilde{u}_i (k_1 r_i^2 + k_2 r_i^4 + \dots) \\ \tilde{v}_i (k_1 r_i^2 + k_2 r_i^4 + \dots) \end{bmatrix} \quad (2.11)$$

and

$$\begin{bmatrix} \delta u^{(t)} \\ \delta v^{(t)} \end{bmatrix} = \begin{bmatrix} 2\rho_1 \tilde{u}_i \tilde{v}_i + \rho_2 (r_i^2 + 2\tilde{u}_i^2) \\ \rho_1 (r_i^2 + 2\tilde{v}_i^2) + 2\rho_2 \tilde{u}_i \tilde{v}_i \end{bmatrix}, \quad (2.12)$$

where k and ρ are the radial and tangential distortion coefficients, and $r = \sqrt{\tilde{u}_i^2 + \tilde{v}_i^2}$. The geometric calibration steps are given in algorithmic form in Algorithm 2.1.

Algorithm 2.1 Geometric calibration.

Input: Calibration image I , intensity threshold th , eccentricity threshold th_e
Output: Distortion parameters p_d
Correct uneven illumination of I using Algorithm 2.2
Apply threshold th to I to extract binary image I_{BW}
Perform connected component analysis on I_{BW}
Remove connected components with eccentricity $> th_e$
Extract centroids X of remaining components
Fit distortion model to X and solve p_d

When analysing data that were not acquired recently, are from a third party or from multiple sources, the imaging system used is often not available for acquiring images of a calibration target. Several methods have been developed for determining the camera target parameters without the use of a calibration target, e.g., [37, 58, 67, 180]. However, many of the methods rely on the movement of the camera or the imaged scene between the images. In the case of fundus images, the movement between frames is generally minuscule, and a data set is not guaranteed to have multiple images of the same eye.

DETERMINATION OF SPATIAL RESOLUTION

When the physical dimensions of the imaged target can be measured, determining the spatial resolution of an image is trivial. This, however, is not the case with retinal images. While the retinal features can be measured, it requires separate measurements with specialised equipment. Determining the spatial resolution for specific imaging parameters will not be sufficient as the size of the eye (and consequently the distance between the retina and the imaging plane) varies from person to person. However, the spatial resolution of the image can be approximated by comparing the size of the retinal features to

known reference measurements. Hemminki et al. [92] include a meta-analysis of the size of the optic disc from 2774 eyes from a total of 13 publications. The mean optic disc area was found to be 2.65 mm, with standard deviation of 0.17 mm. It should be noted that, in order to estimate the spatial resolution based on the cited reference measurements, the optic disc has to be fully visible in the image and not obscured or distorted by e.g., haemorrhages, scar tissue, glaucoma or inter-cranial pressure.

The area of the optic disc can be automatically approximated from images acquired by the camera system under calibration. A number of methods for the automatic segmentation of the optic disc from retinal images have been developed, e.g., [11, 156, 242]. It should be noted that the geometric distortions caused by the lens systems may produce significant error to the measurements. The estimation and removal of these distortions is discussed in Section 2.5.2.

CORRECTION OF UNEVEN ILLUMINATION

Numerous methods for the correction of uneven illumination in different applications can be found in the literature, and the selection between well-working methods for specific data can be largely arbitrary. The method by Narasimha-Iyer et al. [154] was chosen as an example due to good previous performance with the data from DiaRetDB1 diabetic retinopathy database (DiaRetDB1) [116]).

Narasimha-Iyer et al. examine the correction of uneven illumination in the context of detection and classification of changes in longitudinal retinal images due to diabetic retinopathy. Selecting the green channel of an RGB image, retinal features including the optic disc, macula and blood vessels are masked out, and an illumination model is fit to the intensities of the remaining background pixels. As retinal pathologies may cause error in the estimation, the estimation is iterated removing the highest and the lowest 10th percentile of the values of the green channel from the estimation. The process is repeated until the difference in the image locations used in the estimation between iterations becomes small or the maximum number of iterations is reached. The steps are detailed in Algorithm 2.2. A fourth-order polynomial model was found to effectively model the light field attenuated by the ocular media [154].

Algorithm 2.2 Illumination correction.

Input: Image I

Output: Corrected image I_c , maximum number of iterations i_{\max} , acceptable illumination field difference ϵ

Remove blood vessels, macula and optic disc from the mask of accepted values $mask$

while Difference to previous model $> \epsilon$ && iterations $< i_{\max}$ **do**

Fit a polynomial illumination model f_{ill} to $I(mask)$

$I_c = I_c / f_{\text{ill}}$

Remove highest and lowest 10th percentile intensities from $mask$

end while

The parameters of the least-squares fitted illumination field estimate f_{ill} relate to the

image values as

$$I(x, y) = \begin{bmatrix} x_1^4 & x_1^3 y_1 & \dots & y_1 & 1 \\ & & \vdots & & \\ x_i^4 & x_i^3 y_i & \dots & y_i & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{15} \end{bmatrix} = \mathbf{A}\tilde{\mathbf{p}}, \quad (2.13)$$

where (x, y) are image coordinate values belonging to *mask*, \mathbf{A} is the design matrix determining the illumination model and β is the vector of model parameters. The parameters of the illumination field estimate f_{ill} are gained by

$$\tilde{\mathbf{p}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \begin{bmatrix} I(x_1, y_1) \\ I(x_2, y_1) \\ \vdots \\ I(x_i, y_i) \end{bmatrix}. \quad (2.14)$$

An example result of the presented illumination correction approach is shown in Figure 2.10.

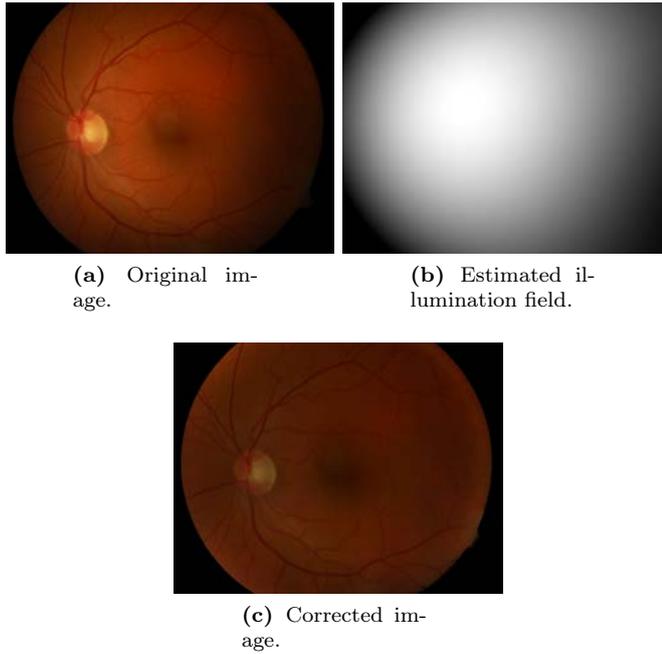


Figure 2.10: An example of uneven illumination correction (images from [116]).

CORRECTION OF SPECTRAL ABERRATIONS

The spectral aberrations (i.e., the effect of different wavelengths of light refracting at different rates in the camera optics) can be determined for each wavelength by imaging a reference white reflectance target (see Figure 2.11). For example, Labsphere Spectralon

diffuse reflectance targets reflect $> 98\%$ of the light for wavelengths between 300 nm to 1600 nm. Before estimating the spectral aberrations based on a stack of channel images captured with narrow-band interference filters, the illumination fields of the channel images need to be corrected. Presuming that the camera stays fixed in relation to the reference target during the acquisition, the shape of the illumination field can be assumed to be constant across the channels, and the illumination field can be estimated from the image resulting from taking the mean over the channels. Illumination correction is discussed in Section 2.7.

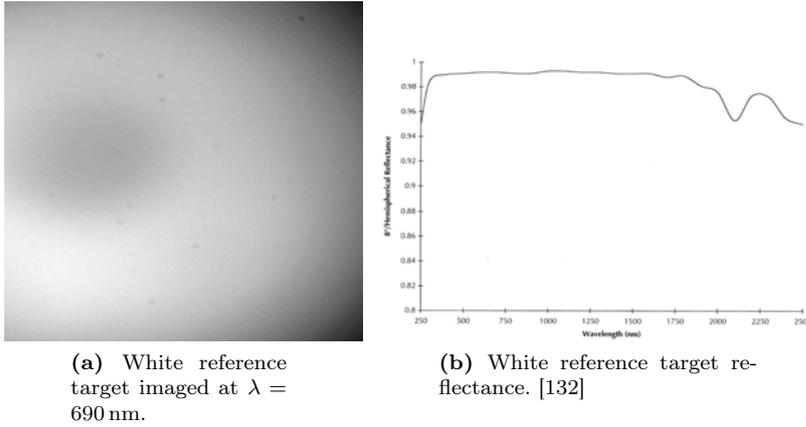


Figure 2.11: White reference target.

The spectrum for each pixel acquired by a spectral retinal camera can be defined as

$$\mathbf{v} = \mathbf{W}\mathbf{T}_{\mathbf{v}_0}^2 \mathbf{v}_0 + \eta, \quad (2.15)$$

where $\mathbf{T}_{\mathbf{v}_0}$ is a diagonal matrix of the ocular media transmittance, \mathbf{v}_0 is a vector of reflected intensities, η is the noise term, and \mathbf{W} is a diagonal weight matrix describing the wavelength dependent aberrations. As the white reference target is expected to reflect significant majority of the light at the visible spectrum, the transmittance can be approximated by

$$\mathbf{T}_{\mathbf{v}_0} = \mathbf{I}. \quad (2.16)$$

Assuming minimal noise and uniform reflection for all wavelengths, the corrected spectrum can be calculated as

$$\hat{\mathbf{v}}_0 = \mathbf{W}^{-1}\mathbf{v}, \quad (2.17)$$

where \mathbf{v} is the vector of values captured by the camera system. [56]

DETECTION OF DUST AND DIRT PARTICLES

Dirt, dust and stains on or inside the optical system of a camera cause artifacts to the images acquired by the system. An example of the effect of dirt in the lens system on images can be seen in Figure 2.12.

Averaging a large number of images taken with the same camera system has been used in previous work [204] to determine the artifacts caused by the system. However, the

white reference target used for determining the spectral aberrations can be used to detect the artifacts caused by dirt and scratches in the optics. The white reference images are preprocessed by applying the illumination correction method detailed in Section 2.7 to produce a more even illumination field and to facilitate segmentation.

The extended-minima transform [209] is applied to the preprocessed images to determine the regional minima. Regional minima represent connected pixels with values with low variation and of lower intensity than their neighbourhood. The extended-minima transform runs through a range of thresholds and combines the binary masks with the logical or operation. The segmented regions are thresholded based on the area to remove artifacts.

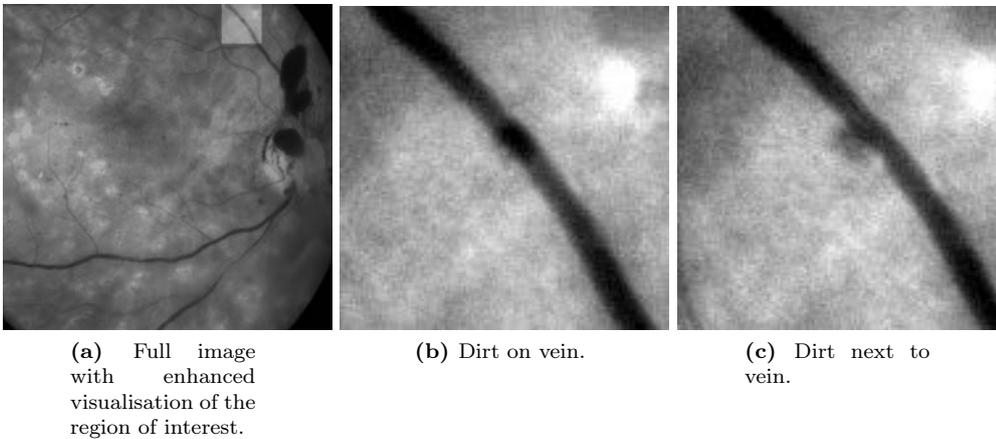


Figure 2.12: Effect of dirt in optics; channel images at $\lambda = 600$ nm of the same patient.

2.5.3 Experiments and results

This section describes the methods for the distortion correction, and the estimation of dirt particle locations and spectral aberrations. As *in-vivo* reference measurements of the human retina for calibration purposes are very challenging, and a phantom encompassing the relevant characteristics of the eye (e.g., curvature, optics, transmission and reflectance) does not exist yet, only qualitative experiments are presented. In the experiments, the system by Fält et al. [56] was used.

GEOMETRIC DISTORTIONS

The distortion generated by the optics of the example system was estimated by imaging a grid distortion target (Edmund Optics #46-250 [170]), with 0.25 mm to 1 mm diameter chrome circles on glass at 0.5 mm to 2 mm intervals, against a white background and determining the mapping from the dot centroids of the imaged grid to the dot locations of the physical target.

29 channel images of the target were acquired with the camera using filters within the range of 410 nm to 694 nm. To test if the distortion varies between the channels of the spectral image, the distortion was estimated individually for each channel. As no systematic inter-channel change in the estimated distortion parameters was observed, suggesting a wavelength-independent distortion, the dot centroid means over all channels were used for the calibration to reduce error from the segmentation.

To determine the accurate centroids and to remove binarisation artifacts, the binary regions resulting from the thresholding were tested for low eccentricity and the regions not meeting these criteria were discarded. While a more sophisticated methods for extracting the dot centroids could be used, the number of successfully extracted centroids using the aforementioned approach was sufficient for estimating the distortion caused by the optics. The calibration parameters were estimated from the centroid locations using the toolbox by Bouguet [30]. The result of correcting the calibration images using the estimated distortion can be seen in Figure 2.13.

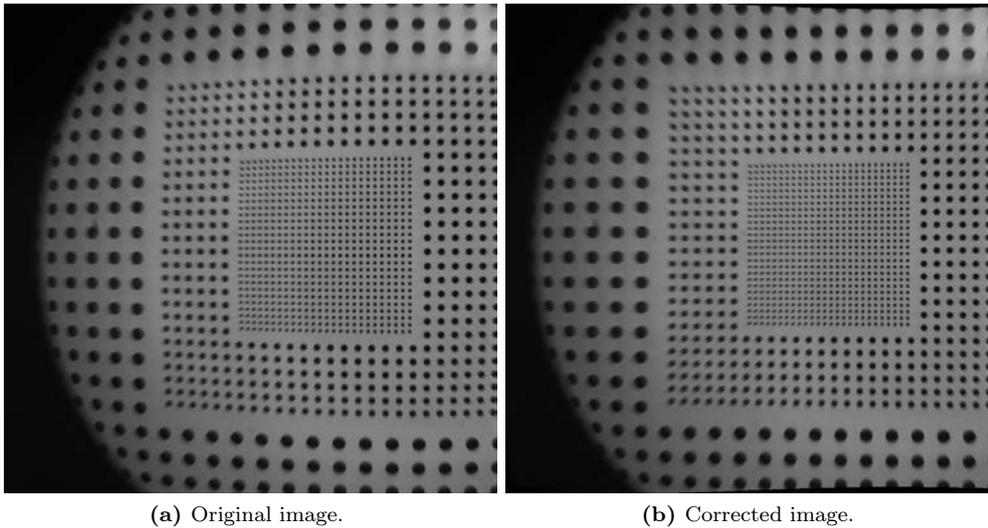


Figure 2.13: An example of distortion correction.

SPECTRAL ABERRATIONS

The system described in [56] uses a set of 30 narrow-band interference filters to acquire spectral images in the range of 400 nm to 700 nm, with ~ 10 nm steps. The intensities of each channel of the spectral images acquired with the system were normalised by the image exposure time. With the near-uniform reflectance of the white reference target and the exposure time accounted for, the remaining differences in the intensity levels between the channels can be accounted to the spatial and spectral characteristics of the illuminant. The spatial distribution of the illumination was normalised by taking the mean over the channels, fitting a polynomial model to the resulting image and dividing each individual channel with the estimated model.

Taking the mean over the normalised intensities of each channel, the resulting mean vector is the (relative) spectrum of the illumination. Dividing the intensities of each channel by the illumination spectrum, the channel images can be normalised to correspond to images acquired with uniform white illumination. The normalisation can be summarised as

$$\hat{I}_\lambda = \frac{I_\lambda}{\bar{I}_\lambda t_\lambda}, \quad (2.18)$$

where I_λ is the channel image acquired at wavelength λ , \bar{I}_λ is the mean and t_λ is the exposure time for the wavelength λ .

After the normalisation, if no spectral aberrations are present, each individual channel should be identical (except for small variation due to noise and numerical error due to normalisation). The only significant source of intensity variation is the illumination field which has been normalised with respect to wavelength.

The spectral aberrations were studied by taking a circle integral over the normalised channel images. The area of the integrating circle was set to 10000 pixels centred at the image focal point and with an increasing radius, with the artifacts due to dirt and scratches masked out (see Section 2.5.3). The results are visualised in Figure 2.14.

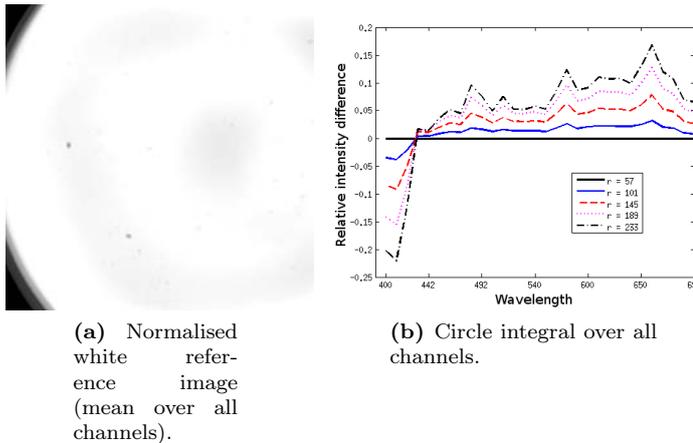


Figure 2.14: Effect of aberrations on measured spectra.

The results show significant spectral aberration especially at the shorter wavelengths of the spectrum. The maximal intensity difference due to aberration, 22%, was measured at $\lambda = 410$ nm. Therefore, any measurement or application relying on accurate spectra from the example system should correct or otherwise take into account the aberrations.

DIRT AND DUST

The dirt and dust in the optics of the example system were determined by acquiring a 30-channel spectral image of a Labsphere Spectralon SRS-99-020 diffuse reflectance target [132] and applying the segmentation approach detailed in Section 2.5.3. The

resulting spatial map of the dirt particles is shown in Figure 2.15, with the pixel intensity indicating the effect of a dirt particle.

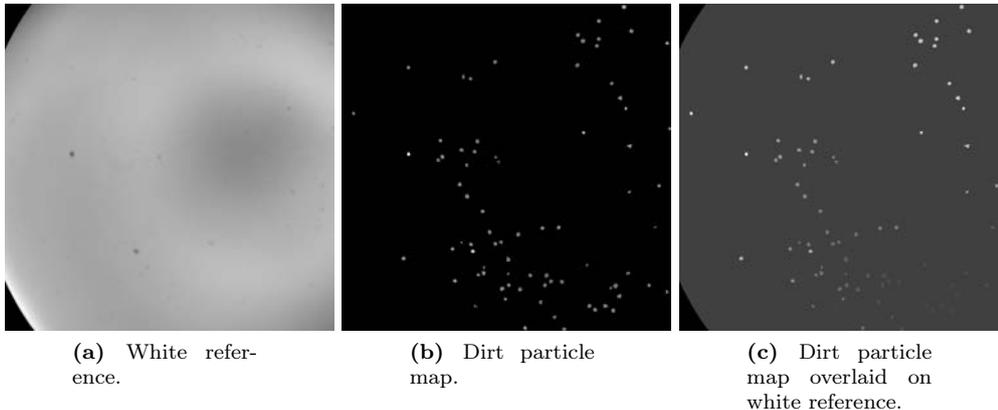


Figure 2.15: Spatial map of dirt particles.

Any image-based measurement or statistical analysis performed on the data should take into account the artifacts caused by the extra particles. While the intensity values of the white reference image could be used to correct any subsequent images taken with the system, the resulting corrected values would be rough estimates of and have higher uncertainty than the true intensities.

2.5.4 Discussion

The distortion caused by the camera system can be characterised by imaging a planar calibration target. As the physiology of the eye varies from person to person, the contribution of the ocular media and the retinal curvature is more difficult to estimate. The magnitude of the contribution of the curvature on the distortion could be estimated by a set of distortion targets with different curvatures within the range of physiologically plausible retinal curvatures.

No specific reason for the maximal aberration to emerge at $\lambda = 410$ nm (see Figure 2.14) was found in the literature. Among the likely contributors to the result are the fundus camera optics. As the amount of short wavelength light re-emitted from the eye is very low, it may be that fundus camera manufacturers are less concerned of the optical properties of the camera system at these wavelengths. Another notable factor is the relatively low intensity of the halogen light source at the short wavelengths. As the ratio of both the ingoing and re-emitted light is low for the short wavelength channels, noise becomes more prevalent. This is likely to affect the aberration measurements.

2.5.5 Summary

A protocol for calibrating a fundus camera was presented. The protocol includes geometric and spectral calibration, determining the spatial resolution, consideration for

correcting uneven illumination and vignetting, and accounting for dirt and dust in the optics. Based on the presented qualitative evaluation results, significant improvements can be achieved related to the quality of spectral retinal images, longitudinal data and data from multiple sources.

Due to the specialised optics and illumination of fundus cameras, imaging a planar white reference target may not accurately capture the spectral aberrations experienced when imaging a curved retina. In general, an appropriate reference target or phantom with similar characteristics as the human eye would improve the calibration process. The future work includes quantitative evaluation of the different sub-tasks of the calibration by using specialised targets, or *in vivo* eye measurements when possible.

2.6 Spectral image composition

Depending on the imaging technology, composing spectral images based on a set of channel images may require image registration at some stage. The purpose of the registration is to find the geometric transformation needed to align the floating image with the base image. Manual registration by selecting corresponding points in image pairs becomes difficult and even infeasible when the number of individual channels increases, or when there are large differences between the images due to differences in viewpoint or modality. When the wavelength difference between the channel image bands increases, the salient image features become very different, which makes manual registration slow, difficult and error-prone.

There is a significant body of work in the field of image registration to solve the problem of image alignment. However, the majority of the approaches are designed for images from the same imaging modality. While registration approaches specific for multimodal data exist (e.g. [236,255]), the modalities are expected to have similar information content. Neither of these prerequisites are necessarily true for the channel images when the difference in acquisition wavelength is large. Despite the fact that the channel images are captured with a single modality, the image appearance varies significantly.

To evaluate the performance of different registration methods on spectral fundus image data, a comparison of image registration approaches for the composition of spectral retinal images is presented. The registration methods are quantitatively evaluated on a set of channel images of an eye phantom and a retinal image set, with each image deformed by a known transformation.

2.6.1 Related work

Medical image registration (and image registration in general) is a widely studied problem and, consequently, a vast body of literature on proposed solutions exists. Maintz and Viergever [143] present an extensive review of medical image registration approaches published between the year 1993 and 1997. A significant number of the methods reviewed deal with the registration of radiological images, with methods dealing with monomodal, multimodal and modality to model registration.

A significant part of medical image registration literature concentrates on MRI, computed tomography (CT) and other radiological modalities. In their paper, published 2001, Hill

et al. [93] review the main approaches (of the time) for registering radiological images. The review presents an overview of rigid feature-based methods and intra- and intermodal (voxel) similarity-based methods.

Feature-based registration methods have been popular in general (i.e., without emphasis on medical modalities) image registration. Zitová and Flusser [256] present a review of general feature- and area-based registration methods. The paper discusses different approaches to feature detection, feature matching, mapping function design, and image transformation and resampling. The evaluation of registration performance of feature and area based registration methods is also discussed.

Deformable transformation models have been widely used in medical image registration. In their 2004 paper, Crum et al. [45] present an overview of deformable medical image registration, with the presented methods dealing mostly with radiological modalities. A number of well-known similarity measures, such as sum of squared differences (SSD), correlation coefficient (CC) and mutual information (MI) are included. Non-rigid transformation models including splines and demons are also discussed. Bhatia et al. [26] present a qualitative evaluation of similarity metrics for group-wise non-rigid registration, including a novel metric. The methods are evaluated on MRI data. Sotiras et al. [212] present a comprehensive study of recent approaches to deformable image registration. A large number of deformable registration methods, classified by the deformation models, matching criteria and optimisation approach used, are described. While not limited to the application area, the study puts an emphasis on methods dealing with the registration of medical images.

A more recent review, published in 2012, of registration methods for medical data has been conducted by Markelj et al. [144]. Three classes of registration base and strategy are identified: feature-, intensity- and gradient-based methods, and projection, back-projection and reconstruction strategies. However, while the paper cites a number of other modalities, the scope of the review is limited to 3D-to-2D registration. More modern approaches to image registration are presented in the study by Wyawahare et al. [245]. In addition to the registration approaches discussed in earlier reviews, e.g., [143], registration methods using wavelets, neural networks, genetic algorithms, fuzzy sets and rough sets are discussed. In [169], Oliveira and Tavares describe the geometric transformations, similarity measures and optimisation methods in common (medical) registration approaches. In addition, available registration software and methods for performance evaluation are reviewed.

While a large part of medical image registration literature is concentrated on radiological modalities, methods focused on the registration of retinal images have been studied. As a part of their review on methods applicable to the automatic screening of diabetic retinopathy, published 2002, Teng et al. [220] present an overview of feature-based registration methods (and two methods that utilise the whole image) used in relation to retinal images. The features used include matched filter responses, vessel branching points and manually marked anatomy markers. The reviewed methods were constrained to rigid transformation models. Laliberté et al. [133] quantitatively evaluate registration methods on retinal colour and fluorescein angiography images. A novel method based on vessel network structure is also presented.

Methods for the registration of spectral images have been proposed by e.g., [89,178,215].

However, the papers on the registration of spectral images seem to generally deal with remote sensing data, and are not (directly) applicable to spectral retinal data. While a significant body of work related to image registration exists (including applications in medical imaging and spectral images), little attention has been given to the inter-channel registration of spectral retinal images.

2.6.2 Methods

A number of methods from the registration literature were selected for a quantitative comparison of registration accuracy on retinal channel images. To offer a broader selection of different approaches, a number of methods based on different frameworks (similarity-, demons- and feature-based methods) found in medical image registration literature were chosen for the comparison. Publicly available implementations were used for all the methods.

DEFORMABLE REGISTRATION BY SIMILARITY

In local similarity based registration, the floating image is deformed in such a manner that a set of local similarities are maximised. In the framework used in this work, the nodes of an $n \times n$ grid are iteratively moved based on local similarity of the base and floating image.

The final transformation is obtained by b-spline interpolation using the grid nodes as control points. Several similarity measures were quantitatively evaluated in this framework: CC, similarity measure by Myronenko et al. (CD2) [153], similarity measure by Cohen and Dinstein (MS) [44], MI [236], minimisation of residual complexity (RC) [152], sum of absolute differences (SAD) and SSD.

Mutual Information (MI) [236] is a measure of similarity (or dependence) between two data sets, measured by the distance between their joint probability distribution and the independent probability distribution. Methods based on maximisation of mutual information have seen frequent use in registration of multimodal medical images.

Myronenko and Song [152] present a similarity measure, RC, that accounts for spatial intensity distortions and based on the minimisation of the complexity of residual image. The method outperforms state-of-the-art similarity measures in several medical registration problems (including retinal image registration), but is limited to monomodal data.

The maximum likelihood (ML) based motion estimation scheme by Cohen and Dinstein [44], MS, differs from other ML schemes by assuming both images to be corrupted by multiplicative noise that follows the Rayleigh distribution in the likelihood function formulation. Another ML based scheme by Myronenko et al. [153], CD2, assumes Rayleigh and blurred speckle noise, and uses spherical coordinates.

REGISTRATION BY DEMONS

Thirion [221] has presented a registration approach (Demons) where the floating image is considered as a deformable grid model diffusing through semi-permeable membranes defined by, e.g., the gradients of the base image. The floating image is transformed by

a grid that is deformed by internal forces (relations between grid points) and demons that locally push a diffusing model (grid node) based on the underlying base image. Vercauteren et al. [233, 234] have extended the demons framework to the space of diffeomorphic transformations by Lie groups, and further extend the diffeomorphic demons framework into log-domain (Log-demons) to guarantee the existence of and access to the inverse transformation.

FEATURE-BASED REGISTRATION

The generalized dual-bootstrap iterative closest point (GDB-ICP) [250] algorithm finds a transformation aligning two images by starting from a small area of overlap (bootstrap region) between the images and a locally stable similarity transformation. An initial transformation derived from a scale invariant feature transform (SIFT) descriptor match is refined and validated by feeding edge and corner points inside a growing bootstrap region to a robust iterative closest point (ICP) algorithm.

To reduce the number of incorrect feature matches, the edge and corner points are divided into driving and matchable features. The driving features, having stricter validation thresholds, are matched to a larger pool of matchable features. To increase the bootstrap region stability, GDB-ICP determines both the forward and inverse transformations and uses bi-directional mapping of the feature points to increase the number of constraints.

The edge-driven dual-bootstrap iterative closest point (ED-DB-ICP) [225] algorithm is a modification of GDB-ICP designed for the registration of multimodal fluorescein angiogram (FA) sequences. The main differences of ED-DB-ICP with respect to GDB-ICP are the use of gradient magnitude images instead of intensity images and extending the SIFT descriptors with a shape context descriptor presented in detail in [151].

2.6.3 Registration strategy

In the case of spectral image composition, all channels need to be transformed into the same space to form a complete spectral image. This section discusses different strategies for composing the set of pair-wise registrations into a spectral image where all channels are aligned with respect to each other.

REGISTRATION TO A SINGLE BASE IMAGE AND SEQUENTIAL REGISTRATION

A simple approach for channel image set registration is to register each image to a previously selected base image. In the case of spectral images, however, this can mean considerable difference in image features as a result of large differences in wavelength between the base image and the image to be registered. The difference in the wavelength means differences with respect to their structure and intensity due to the varying reflectance of different features of the retina.

Large differences in image structure and intensity due to wavelength can be avoided by registering each channel image to its immediate neighbour (in the direction of the base image). For images that are not immediate neighbours of the base image, the transformations steps can be sequentially combined to align the images with the chosen base image.

JOINT REGISTRATION

Instead of registering the channel image set as a number of independent pair-wise registrations, the overall registration result is likely to be improved if the registration strategy considers the whole image set. Joint registration has been extensively studied, especially in MRI and tomography [13, 25, 108, 226, 244, 249]. However, the majority of joint registration approaches in the literature are applicable (at least without significant changes) only to registration methods based on a similarity measure, and are often an integrated part of the method.

Considering the promising performance of feature-based methods in retinal image registration, two joint registration strategies applicable to both feature and similarity measure based registration are presented.

Registration using intermediate templates Instead of directly registering to a neighbouring or to the base image, the floating image can be registered to an intermediate template to avoid both accumulation of error from combined transformations and large differences in features and intensities. Each floating image is registered to a template that is a combination of the previously registered intermediate (i.e., channels between the floating and the base) images. Similar approaches for group-wise registration have been presented in e.g., [82, 105].

The template Y_i used as the registration target for floating image I_i can be defined as

$$Y_i = \frac{I_{i-1} + Y_{i-1}}{2}, \quad (2.19)$$

where Y_{i-1} is the template of the previous step and I_{i-1} is the previously registered floating image. In the first step, the base image is used as the template. Each intermediate image can be given an equal weight in the generation of the template, but this was found to yield poor results in initial experiments. As each template has been (not accounting for the registration error) transformed into the same space with the base image, the resulting transformation is of the same complexity for each channel image.

Registration using shortest path The sequential registration strategy can be improved by registering only the images along the shortest path from a floating image to the base image instead of all intermediate images. Here the cost of each registration step is measured in image similarity. A cost matrix is defined by calculating the squared sum of intensity error for each combination of pairs for the image set. For each image, Dijkstra's algorithm is applied to the cost matrix to determine the shortest path to the base image. The final transformation is gained by combining the pair-wise transformations along the path.

2.6.4 Experiments

SYNTHETIC AND SEMISYNTHETIC DATA SETS

To quantitatively evaluate the performance of each method with the data acquired with the system described in [55], five sets of channel images were used to directly estimate

the image registration error. A fully aligned spectral image consisting of 30 channels with spatial resolution of 1024x1024 of a phantom (artificial eye) was used as the basis.

The Carl-Zeiss Meditec eye phantom used as the basis for the synthetic images is a closed container with a small entry pupil fitted with an optical system simulating the lens of the human eye. The back of the container is concave (to represent the curvature of the eye fundus), with painted retinal structures (e.g., vasculature and fovea).

A known transformation of increasing degree was applied to each set to simulate eye movement between the acquisitions of individual channels of spectral retinal images. The initial and the extended performance evaluation use different approaches to generating the known transformation. The approaches are described in Sections and , respectively. The produced sets are referred to as the (synthetic) test set A1-A5.

To evaluate the registration performance on real medical data, a semisynthetic test set was generated by warping a set of five spectral retinal images taken with the system in [36]. The system captures a set of six channel images in approximately 0.5 s to avoid inter-channel eye movement [36]. The retina was illuminated at six selected wavelengths 507, 525, 552, 585, 596, 611 nm [216] using a halogen white-light source filtered through a liquid crystal tunable filter. During the experiments, the maximum inter-channel displacement was found to be 2.3 pixels (referred to as the system error). The same approaches for generating the transformation simulating the movement of the eye were used as with the synthetic test sets A1-A5. The produced sets are referred to as the (semisynthetic) test set B1-B5.

REGISTRATION ERROR AND THE QUALITY OF IMAGE SPECTRA

The channel-wise registration errors can cause significant error in the image spectra. To estimate the deterioration of the quality of the spectra as the registration error increases, artificial systematic misalignment was applied to a spectral image of an eye phantom and metrics measuring the quality of the spectra were calculated.

30 channel images of the eye phantom were translated n pixels in a direction unique for each channel (by a monotonically increasing angle) to simulate a mean registration error of n pixels. The decrease in spectral quality was calculated using a set of quality metrics from the resulting misaligned spectral image y , using the original aligned image as the reference x . An example is shown in Figure 2.16.

Vaiopoulos [228] has implemented a number of technical and reference metrics for the assessment of spectral image quality. The metrics for spectral image y and reference spectral image x adopted for the evaluation were correlation coefficient (CC), quality index (QI) and root-mean-square error (RMSE), defined as

$$CC = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \quad (2.20)$$

$$QI = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)(\bar{x}^2 + \bar{y}^2)} \text{ and} \quad (2.21)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}, \quad (2.22)$$

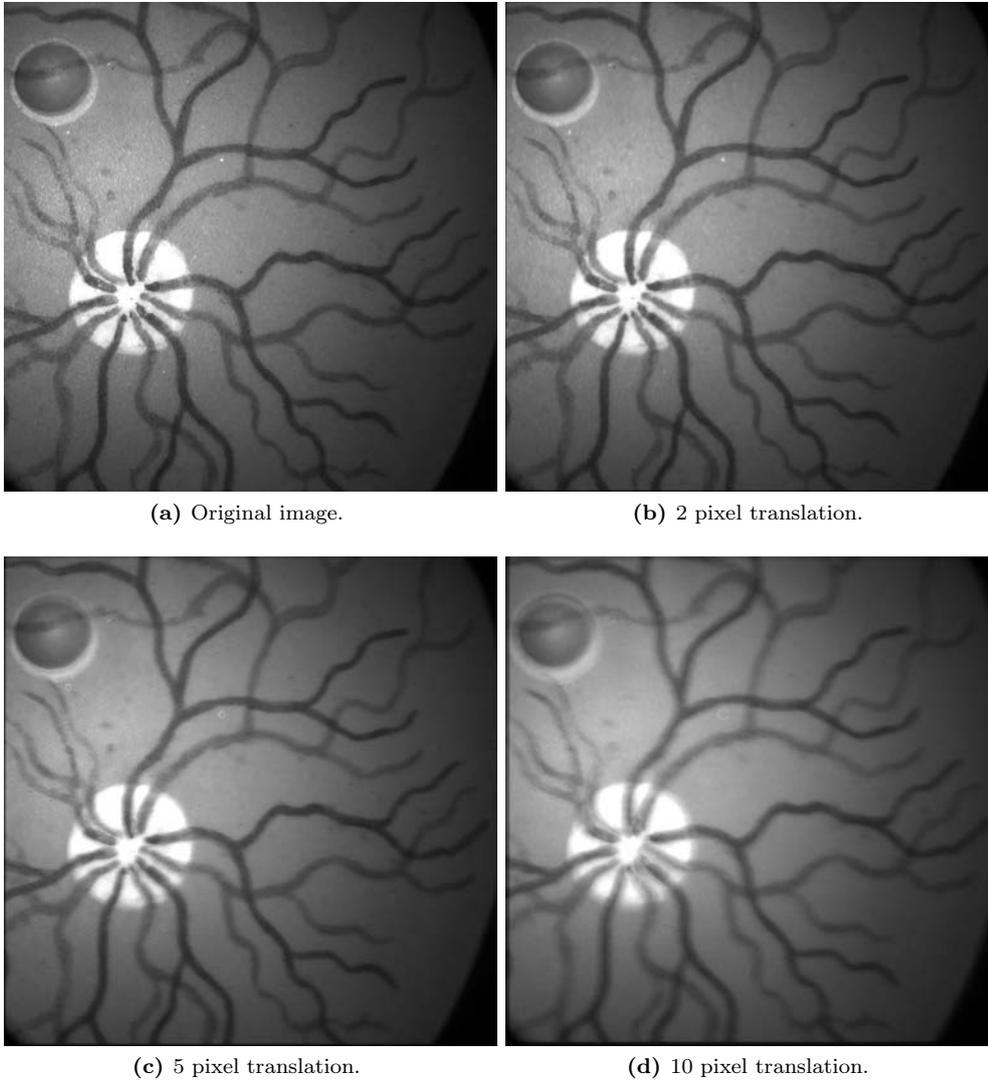


Figure 2.16: Examples of the reference set for the effect of registration error on image quality. The images are mean images over the wavelength channels with each channel translated n pixels into a different direction (except the original). [128]

where n is the number of channel image pixels, \bar{x} is mean over x and σ_x is the standard deviation of x .

Van der Meer [230] presents a number of similarity measures for spectral data. The similarity measures, spectral angle measure (SAM), spectral correlation measure (SCM) and spectral information divergence (SID), can be used as a measure of how well the

image spectra are preserved after the alignment. The similarity measures are defined as

$$\text{SAM} = \cos^{-1} \left(\frac{\sum_{i=1}^d x_i y_i}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}} \right) \quad (2.23)$$

$$\text{SCM} = \frac{\sum_{i=1}^d x_i y_i - \sum_{i=1}^d x_i \sum_{i=1}^d y_i}{\sqrt{\left[d \sum_{i=1}^d x_i^2 - \sum_{i=1}^d (x_i)^2 \right] \left[d \sum_{i=1}^d y_i^2 - \sum_{i=1}^d (y_i)^2 \right]}} \quad (2.24)$$

$$\text{SID} = \sum_{i=1}^d p_i (\log p_i - \log q_i) + \sum_{i=1}^d q_i (\log q_i - \log p_i), \quad (2.25)$$

where

$$p_i = \frac{x_i}{\sum_{j=1}^d x_j}, \quad q_i = \frac{y_i}{\sum_{j=1}^d y_j}, \quad (2.26)$$

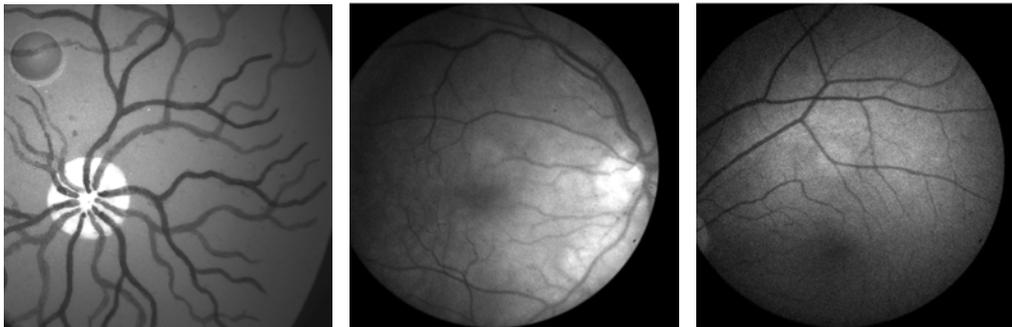
and d is the number of channels in the spectra.

INITIAL PERFORMANCE EVALUATION

An initial performance evaluation was performed with all the presented methods to determine a subset of the methods for more in-depth evaluation [129]. The registration strategies were not considered in the initial evaluation, and a more straightforward approach to the artificial deformations to produce the synthetic and semisynthetic data sets was taken. The methods with the lowest registration errors were selected for further evaluation.

For both the synthetic and semisynthetic test sets, five image sets were generated from the original images by transforming each channel by a known transformation with transformation parameters sampled from a parameter distribution. Based on the experiments in [214], a quadratic transformation was deemed appropriate to represent the deformations caused by eye movement during the acquisition of retinal images. The parameter distributions for the transformation parameters is determined from a previously registered true retinal channel image set (with 1 442 images and successful registration confirmed by an expert) by using kernel density estimate (KDE). The parameters used for deforming the images of the test sets were gained by randomly sampling the distributions. The parameter distributions are shown in the Appendix I.2.

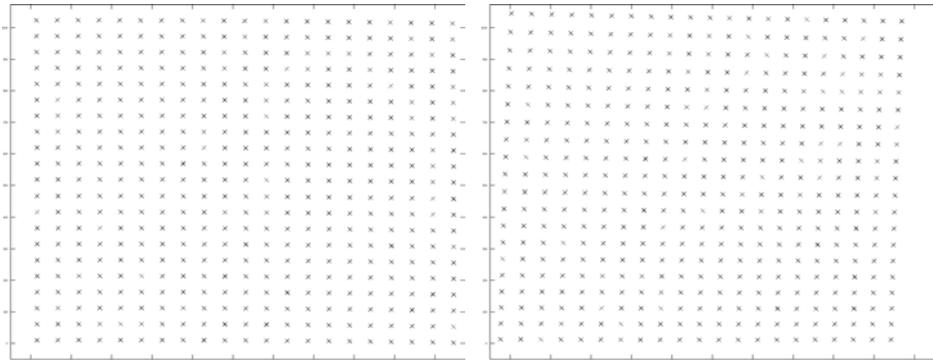
The quadratic term parameters were multiplied by an increasing weighting constant k for each image set to simulate increasingly distorted images. For each set, a channel image near the middle of the wavelength range ($\lambda = 540$ nm for the synthetic and $\lambda = 552$ nm for the semisynthetic, respectively) was chosen as the base image for each registration.



(a) An original image of the synthetic Set A ($\lambda = 600$ nm).

(b) An original image of the semisynthetic Set B1 ($\lambda = 585$ nm).

(c) An original image of the semisynthetic Set B5 ($\lambda = 585$ nm).



(d) An example transformation grid for synthetic Set A1.

(e) An example transformation grid for synthetic Set A5.

Figure 2.17: Examples of synthetic and semisynthetic test sets (enhanced for visualisation). [129]

Examples of the test set images and the corresponding transformations are shown in Figure 2.17.

The parameters for each registration method were systematically selected by registering a subset of the test images whilst varying the parameter values. The parameter combination that produced the smallest error on the images was chosen. The parameters selected for the evaluation are presented in the Appendix in Tables I.1 and I.2.

The implementation of GDB-ICP released by the authors of the method was used for the evaluation. Publicly available third party implementations were used for methods for which the original implementations were not available.

In the semisynthetic test set, it is possible that the registration method, in addition to estimating the synthetic transform, corrects some of the system error. This would manifest itself as increased error despite the more accurate registration. Therefore, error values below the system error of 2.3 pixels were considered as zero for the semisynthetic test set. All registration errors were measured as the Euclidean distance between the grid points of the registered image and their expected locations based on the known transformation.

EXTENDED PERFORMANCE EVALUATION

The three methods with the lowest registration errors in the initial study were selected for the extended performance evaluation [128]. A more realistic transformation model was used for the generation of the test sets, and all the presented image set registration strategies were considered for the three methods.

In the extended performance evaluation, five image sets were generated from the original images of both the synthetic and semisynthetic test sets by transforming each channel by a designed transformation. To simulate the changes in retinal images as the eye moves with respect to the camera, the test set images were first projected onto a semi-sphere with the pixel depth values defined as

$$Z = \sqrt{r^2 - (X - x_0)^2 - (Y - y_0)^2}, \quad (2.27)$$

where X and Y are the x- and y-coordinates of the image pixels (with origin at the image centre), and r is the radius of the semi-sphere, defined as

$$r = \left(\frac{d_U}{2} + A^2\right)/(2 * A), \quad (2.28)$$

where d_U is the image width and

$$A = \frac{d_U}{2}/(\tan \alpha_{fov}/2), \quad (2.29)$$

where α_{fov} is the (horizontal) field-of-view (FOV) angle of the camera.

The semi-sphere was translated and rotated to simulate an offset in the optical axis and movement of the eye with respect to the camera. The rotated and translated 3D-coordinates are gained by

$$V_i = R \begin{pmatrix} X_i + t_x \\ Y_i + t_y \\ Z_i \\ 1 \end{pmatrix}, \quad (2.30)$$

where t_x and t_y are the translation parameters, and R is a homogeneous rotation matrix defined by the rotation angles φ_x and φ_y around the x- and y-axis (with no rotation around z-axis).

The transformed image coordinates were gained from the rotated and translated 3D-coordinates as

$$\tilde{X}_i = \frac{V_{x_i}}{V_{w_i}} \tilde{Y}_i = \frac{V_{y_i}}{V_{w_i}}, \quad (2.31)$$

where V_w is the fourth component of the homogeneous transformed 3D-coordinates. The deformed image was gained by estimating the values at the transformed image coordinates by bicubic interpolation. Examples of the test set images and the corresponding transformations are shown in Figures. 2.18 and 2.19.

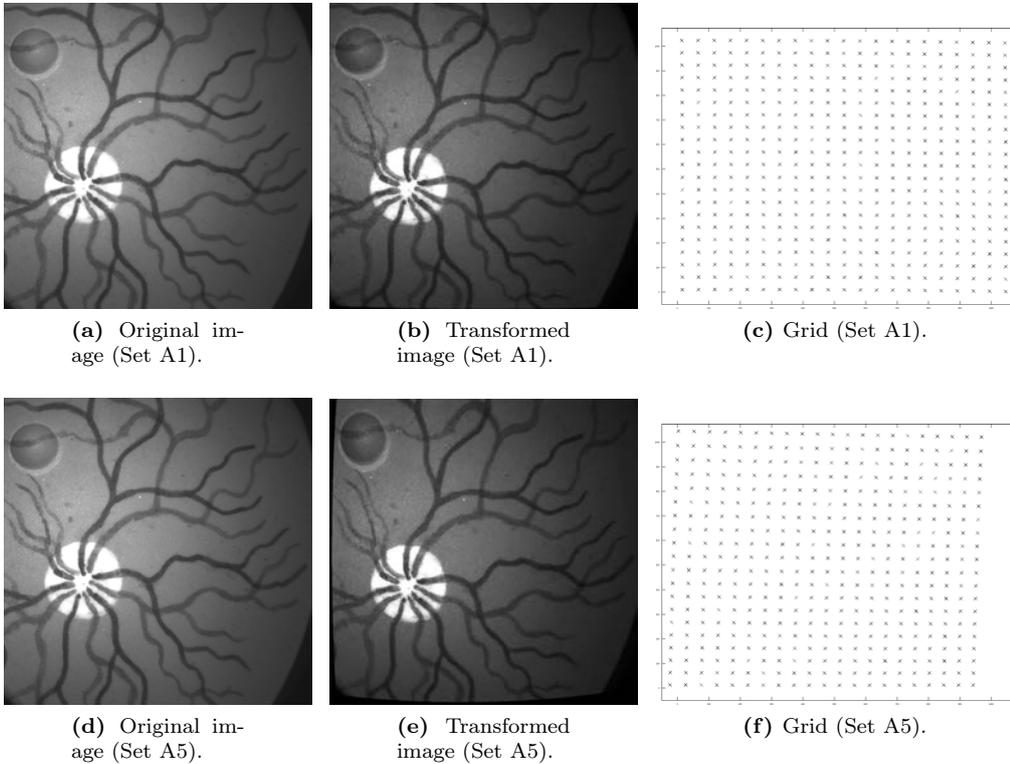


Figure 2.18: Synthetic test set A; examples at $\lambda = 589$ nm of the original and transformed images (enhanced for visualisation) with corresponding transformation grids. [128]

To validate the approach for the generation of the test sets, images of an eye phantom setup, for which the angle between the phantom and camera can be accurately set, were deformed with the approach for test set generation. The set contained three sets of images with the phantom in 6 different rotations in relation to the camera (3° , 1.5° , -1.5° , -3° , -4.5° , -6°). An example is shown in Figure 2.20.

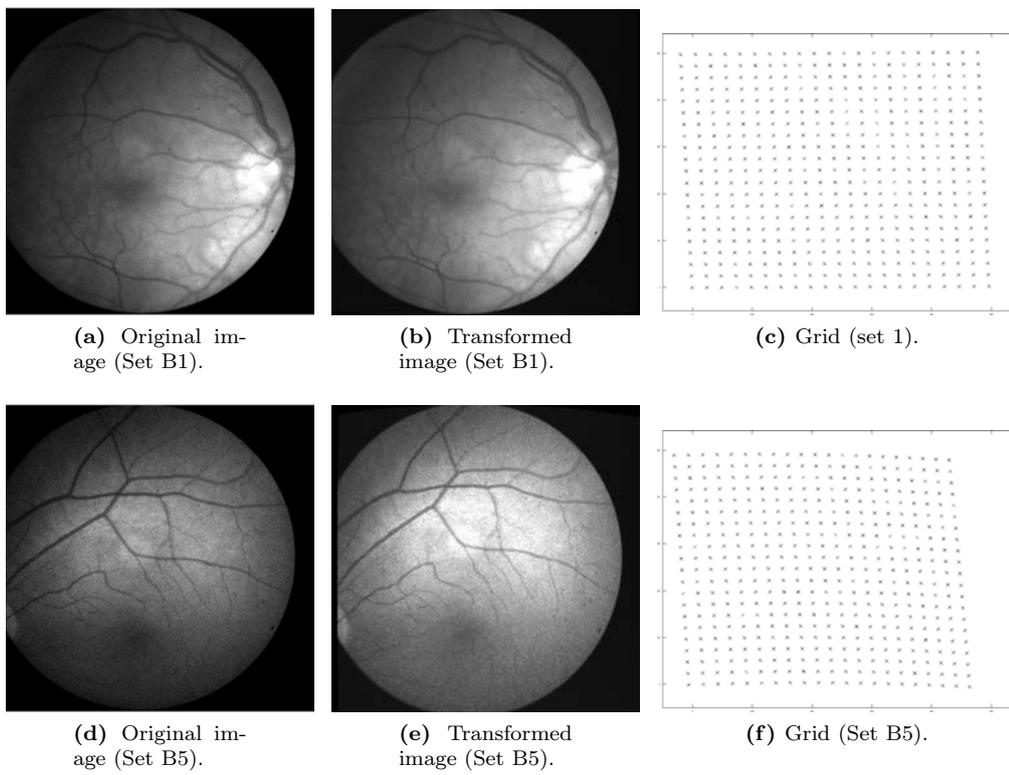


Figure 2.19: Semisynthetic test set B; examples at $\lambda = 585$ nm of the original and transformed images (enhanced for visualisation) with corresponding transformation grids. [128]

An image with the phantom facing the camera directly was deformed with an increasing angle φ_y , and the deformed image was compared to an image with the phantom physically rotated φ_y degrees. The error between the physical and simulated deformation was determined as the displacement of speeded-up robust feature (SURF)-keypoints [16] visible in both images. The numerical results are shown in Table 2.2.

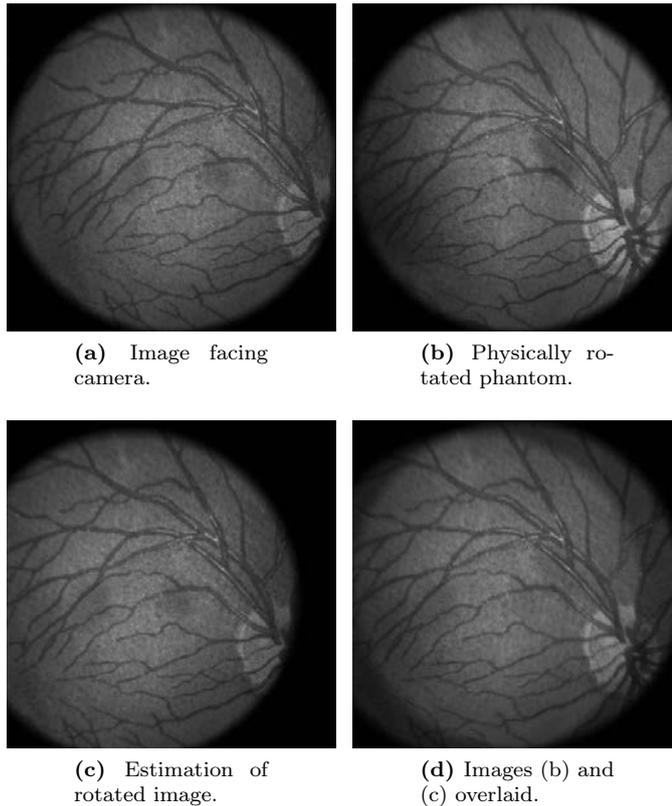


Figure 2.20: Examples of phantom set used in the validation of synthetic data generation. [128]

It should be noted that the artificial retinal features of the phantom used for validating the synthetic data generation approach are painted on a plane (and consequently, the pixel locations of the deformed image are projected onto a plane instead of a semi-sphere). Due to the differences between the true retinal curvature (semi-synthetic data), the curvature of the synthetic test set phantom and an ideal semi-sphere, the deformations in the synthetic ground truth are unlikely to correspond accurately with the eye movements they are simulating. However, the purpose of the deformation approach is not to accurately simulate specific eye movements, but to produce realistic deformations to the synthetic ground truth. Based on the validation results, the approach to synthetic ground truth generation can be expected to produce deformations similar to those caused by the movements of the eye in relation to the camera.

Table 2.2: Validation of the synthetic data generation approach; errors (in pixels) between feature locations in mechanically rotated and artificially deformed images, respectively.

φ_y	-6°	-4.5°	-3°	-1.5°	1.5°	3°
Mean	2.98	1.85	1.38	0.83	1.71	2.74
Std	0.67	0.48	0.45	0.32	1.00	1.07

The same parameters for each registration method were used as in the initial study (Appendix Tables I.1 and I.2). As with the initial study, all registration errors were measured as the Euclidean distance between the grid points of the registered image and their expected locations based on the known transformation, and error values below the system error of 2.3 pixels were considered as zero for the semisynthetic test set.

2.6.5 Results

INITIAL EVALUATION

The results of the image registration performance comparison on the synthetic and semisynthetic test sets are summarised in Tables 2.3 and 2.4, respectively.

Table 2.3: Median (Med.) and standard deviation (STD) of registration error for the synthetic set A. Base stands for the initial error before registration. Corr is the correlation with baseline error. Lowest median errors are in bold. [129]

		Base	CC	CD2	MI	MS	RC	SAD	SSD	Demons	Log-demons	GDB-ICP	ED-DB-ICP
Med.	Set A1	33.7	2.9	64.4	2.3	154.1	2.0	128.8	93.9	110.4	33.5	1.0	0.7
	Set A2	47.3	2.6	101.3	7.2	128.6	2.0	144.3	81.0	323.8	48.4	1.1	0.7
	Set A3	28.2	3.0	105.8	3.6	233.8	2.2	177.4	114.4	143.4	28.9	1.1	0.7
	Set A4	28.9	4.2	29.4	11.9	257.2	4.3	115.3	55.8	18.7	32.8	1.1	0.7
	Set A5	43.0	6.1	105.0	11.8	200.1	5.4	147.4	126.6	139.9	43.3	1.0	0.7
STD.	Set A1	6.6	3.0	46.4	1.9	74.9	4.9	67.8	51.2	50.8	10.3	0.4	0.4
	Set A2	18.6	6.4	40.4	9.3	52.4	4.6	47.4	40.0	210.9	18.7	0.4	0.3
	Set A3	9.8	3.5	22.3	5.4	53.8	6.1	47.1	42.4	63.8	15.2	0.3	0.2
	Set A4	14.6	8.6	28.2	19.4	69.3	18.9	50.7	30.6	18.9	17.4	0.4	0.4
	Set A5	14.5	8.3	20.7	9.0	41.5	8.9	35.2	44.0	75.4	18.4	0.4	0.3
Corr.		1.00	0.16	0.53	0.25	-0.74	0.09	0.01	0.20	0.78	0.98	-0.33	0.05

CC performed relatively well for both the synthetic and semisynthetic sets, achieving a median registration error below five pixels for most images. The variation of error for some images was high, especially for the synthetic sets A4 and A5, however. Excluding some individual images, RC produced low median errors for all sets for both synthetic and semisynthetic images. However, the error variance was high, especially for the fourth synthetic set A5. MI produced relatively good results for the first synthetic set, but showed sensitivity to the increasing level of deformation with the successive sets.

For the semisynthetic set B, MI performed comparably to CC. CD2, MS, SAD and SSD managed to, at least in part, successfully register the images at wavelengths near the base image, but produced high errors elsewhere.

Table 2.4: Median (Med.) and standard deviation (STD) of registration error for the semisynthetic set B. Base stands for the initial error before registration. Corr is the correlation with baseline error. Lowest median errors are in bold. [129]

		Base	CC	CD2	MI	MS	RC	SAD	SSD	Demons	Log-demons	GDB-ICP	ED-DB-ICP
Med.	Set B1	13.8	4.5	36.8	5.8	50.8	4.0	56.5	33.0	39.2	16.1	0.0	3.5
	Set B2	51.4	0.0	18.4	0.0	21.3	9.9	44.7	15.1	15.2	39.0	0.0	0.0
	Set B3	25.7	0.0	32.6	0.0	10.9	0.0	61.4	3.7	15.2	23.3	0.0	-
	Set B4	58.7	4.6	167.7	4.5	6.4	5.1	17.1	4.8	8.7	34.2	0.0	34.7
	Set B5	33.2	3.4	41.3	3.5	11.5	4.2	29.4	10.2	12.4	31.8	0.0	332.2
STD	Set B1	5.1	3.8	14.6	21.6	21.6	11.3	21.4	13.1	14.5	6.0	0.2	2.3
	Set B2	1.3	0.3	7.9	10.9	10.9	14.1	21.5	6.3	6.7	4.4	0.0	0.0
	Set B3	3.5	1.1	2.4	5.6	5.6	5.0	27.6	2.2	5.3	3.5	0.0	-
	Set B4	9.2	2.0	6.7	4.8	4.8	2.4	11.3	2.3	4.8	9.6	0.0	13.0
	Set B5	13.6	3.2	19.9	8.4	8.4	6.4	16.9	8.0	9.6	11.1	0.0	117.6
Corr.		1.00	-0.07	0.59	-0.23	-0.65	0.59	-0.74	-0.57	-0.76	0.91	0.00	-0.15

Demons showed very high standard deviation in error and produced globally unrealistic transformations. The method, however, registered local image areas with high accuracy, resulting in error minimum close to zero for most images and low median errors for the wavelengths close to the base image. Log-Demons performed similarly to Demons, but produced less extreme errors and smaller standard deviations for most images. Furthermore, log-demons showed high correlation between the registration and baseline error, suggesting that the method is sensitive to the initial configuration.

ED-DB-ICP was hindered because of using the gradient information, especially in the semisynthetic test set B. While the successful registrations were highly accurate, the method failed (i.e., the method found no stable transformation) with several images in all sets. GDB-ICP showed very good performance for both the synthetic and semisynthetic test sets. With the exception of failed registrations for the two shortest wavelength images of the synthetic set A, and two partially successful registration in the semisynthetic set B, the method achieved a close to sub-pixel accuracy with minimal standard deviation. The channel-wise performances for Sets A1 and B1 of the test sets are shown in Figure 2.21.

As expected, the errors tend to increase with the wavelength difference of the registered images as the prominent image features change. The channel-wise errors of the best-performing methods are visualised in Fig. 2.22.

EXTENDED EVALUATION

The registration accuracy was measured as pixel error between the resulting grid points and the original ones before the known transformation. MI produced relatively good results for the medium and long wavelengths of the first synthetic set A1, but performed poorly for wavelengths shorter than 520 nm and showed sensitivity to the increasing level of deformation with the successive sets (see Figure 2.23). Furthermore, due to the deformable transformation, regions of high registration error could be found in otherwise well registered images.

Using intermediate templates significantly decreases the registration error in the shorter wavelengths (see Figure 2.24). While the shortest path strategy similarly improved the

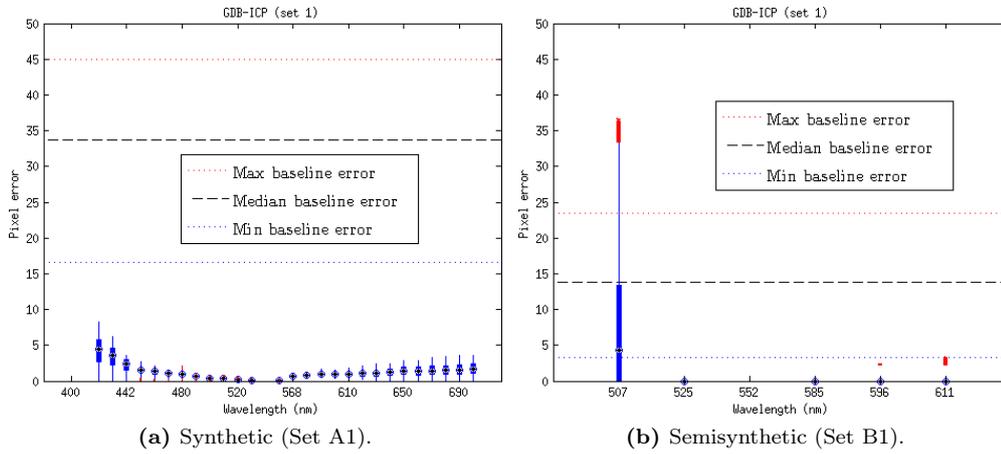


Figure 2.21: The generalized dual-bootstrap iterative closest point (GDB-ICP) errors; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [129]

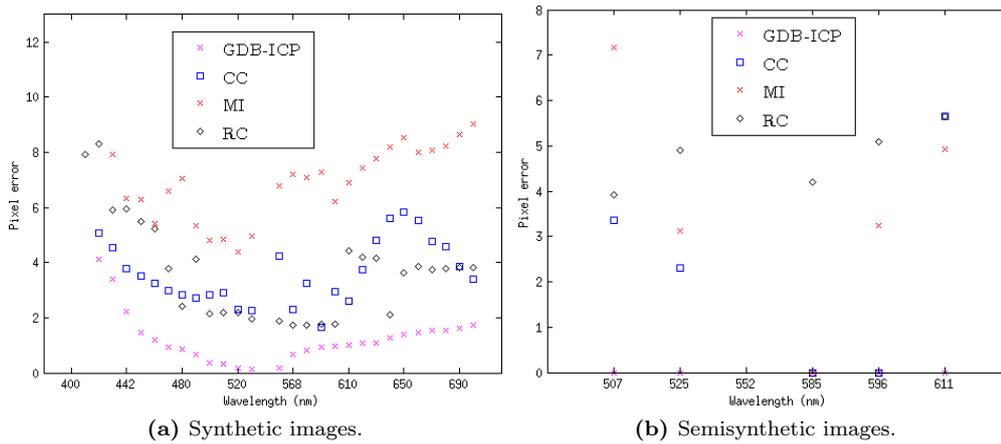


Figure 2.22: Wavelength-wise registration error; Median errors over the sets for each wavelength. [129]

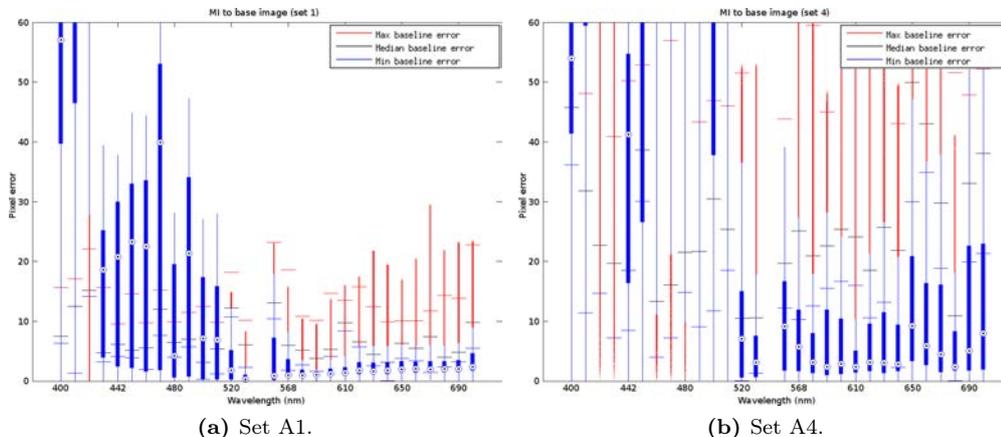


Figure 2.23: The mutual information (MI) errors with the synthetic set A; all images registered to a single base image captured at 540 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [128]

registrations in the shorter wavelengths, there was an adverse effect to the general registration performance (shown in Figure 2.26).

In the semisynthetic set B, the intermediate template strategy improved the performance to a point where the mean registration error was within the system error for majority of the images. The errors are visualised in Figure 2.25. However, most of the registered images in the test set contained regions with higher registration error, i.e., the registrations were only partially successful. For the semisynthetic set B, MI did not show similar loss of performance with shorter wavelengths.

RC performed similarly to MI for the medium and longer wavelengths, but showed significantly better performance in the shorter wavelengths. However, similar regions of larger error were present (although less severe than with MI) in the images, and the method showed similar sensitivity to increasing deformation of the test images. Furthermore, for a few images of the synthetic set A, RC failed completely (i.e., the registration error for all pixels was over 60). Neither of the joint registration strategies provided any significant increase in performance for RC. The errors are visualised in Figure 2.27.

GDB-ICP showed very good performance for both the synthetic and semisynthetic test sets. With the exception of failed registrations for two of the synthetic set, and one failed registration in Set B5 in the semisynthetic set, the method achieved a reasonable registration error with minimal standard deviation. The median registration error remained below 2 pixels for the majority of the images. As the level of deformation increased in the synthetic images, some regions showed increased registration error (denoted as outliers). The results are shown in Figures 2.28 and 2.29

For GDB-ICP, registering to a single base image provided the lowest error for the method in both synthetic and semisynthetic sets. Likely due to restricting the transformation

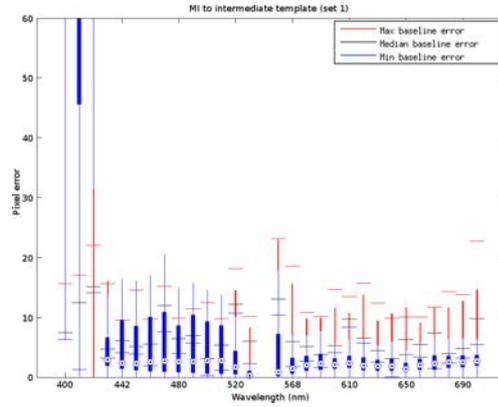
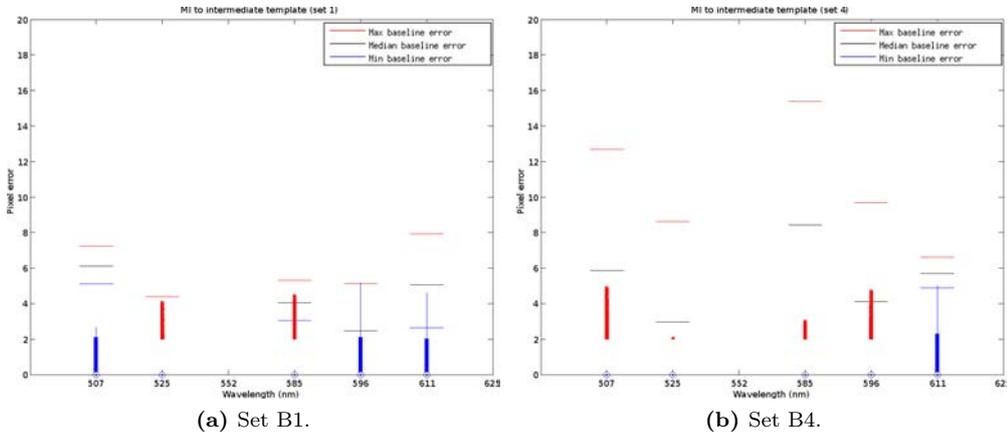


Figure 2.24: The mutual information (MI) errors with the synthetic set A1; intermediate template strategy; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [128]



(a) Set B1.

(b) Set B4.

Figure 2.25: The mutual information (MI) errors with the semisynthetic set B; intermediate template strategy; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [128]

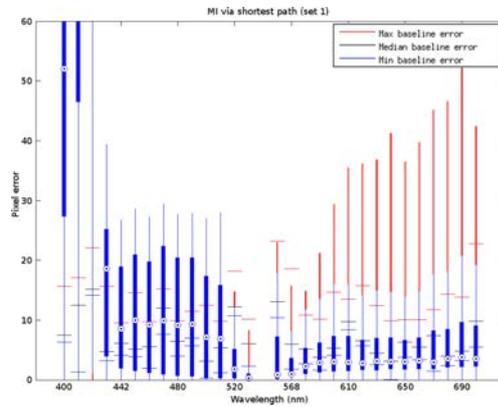
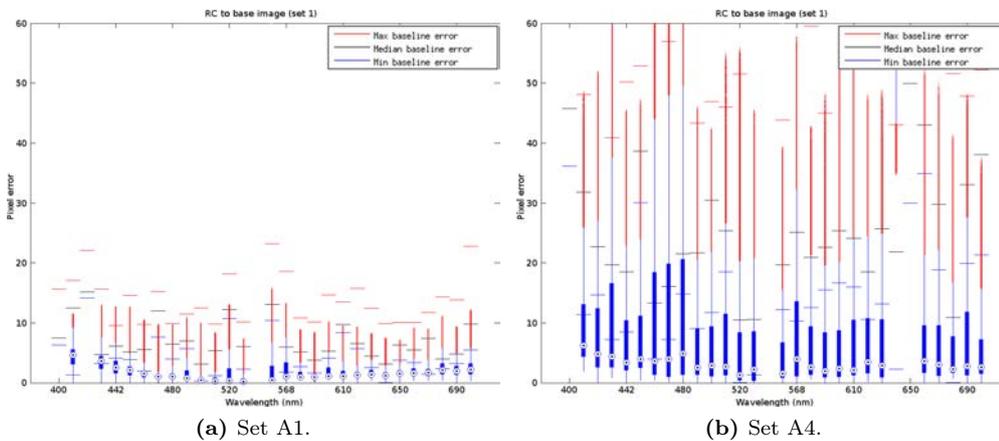


Figure 2.26: The mutual information (MI) errors with the synthetic Set A1; shortest path strategy; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [128]



(a) Set A1.

(b) Set A4.

Figure 2.27: The minimisation of residual complexity (RC) errors with the synthetic set A; all images registered to a single base image captured at 540 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [128]

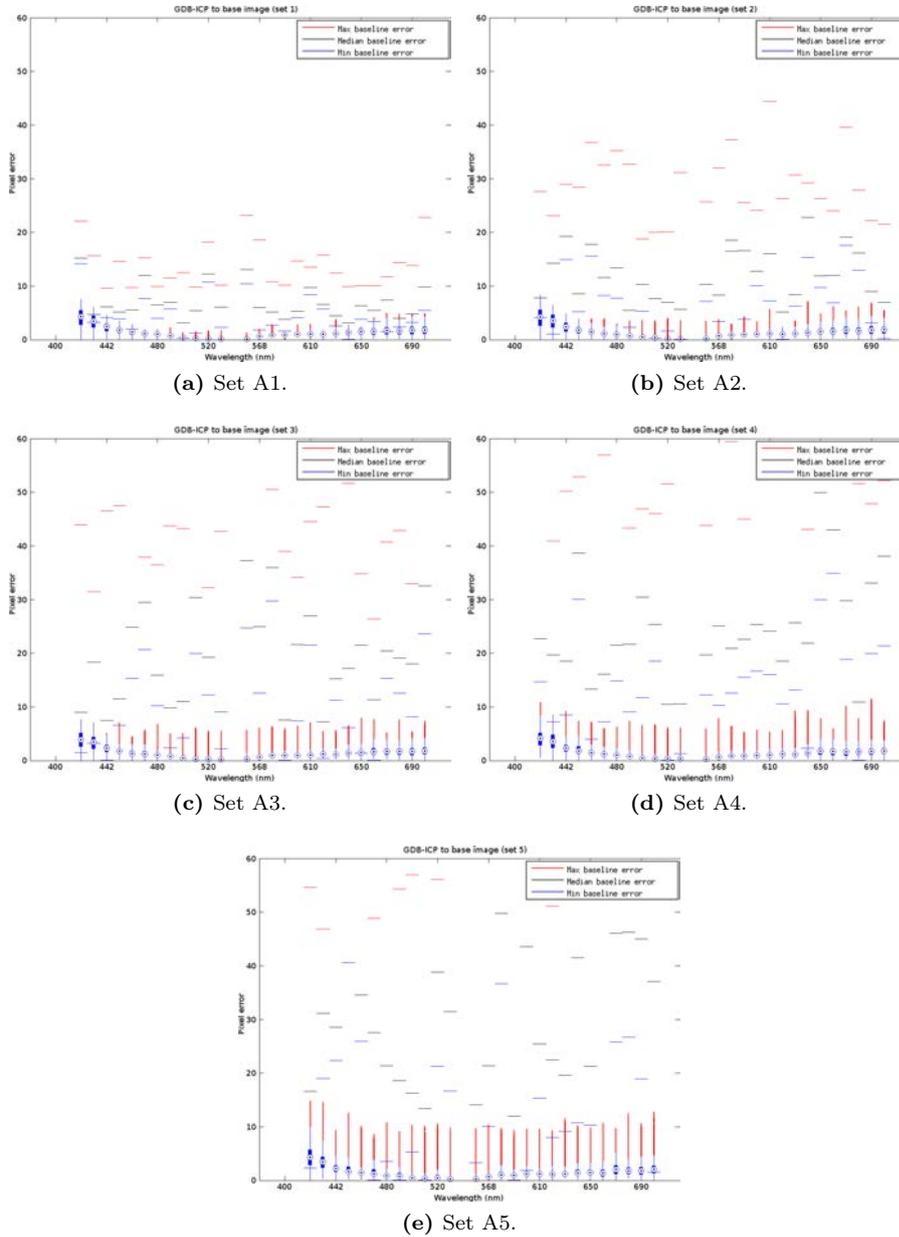


Figure 2.28: The generalized dual-bootstrap iterative closest point (GDB-ICP) errors with the synthetic set A; all images registered to a single base image captured at 540 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [128]

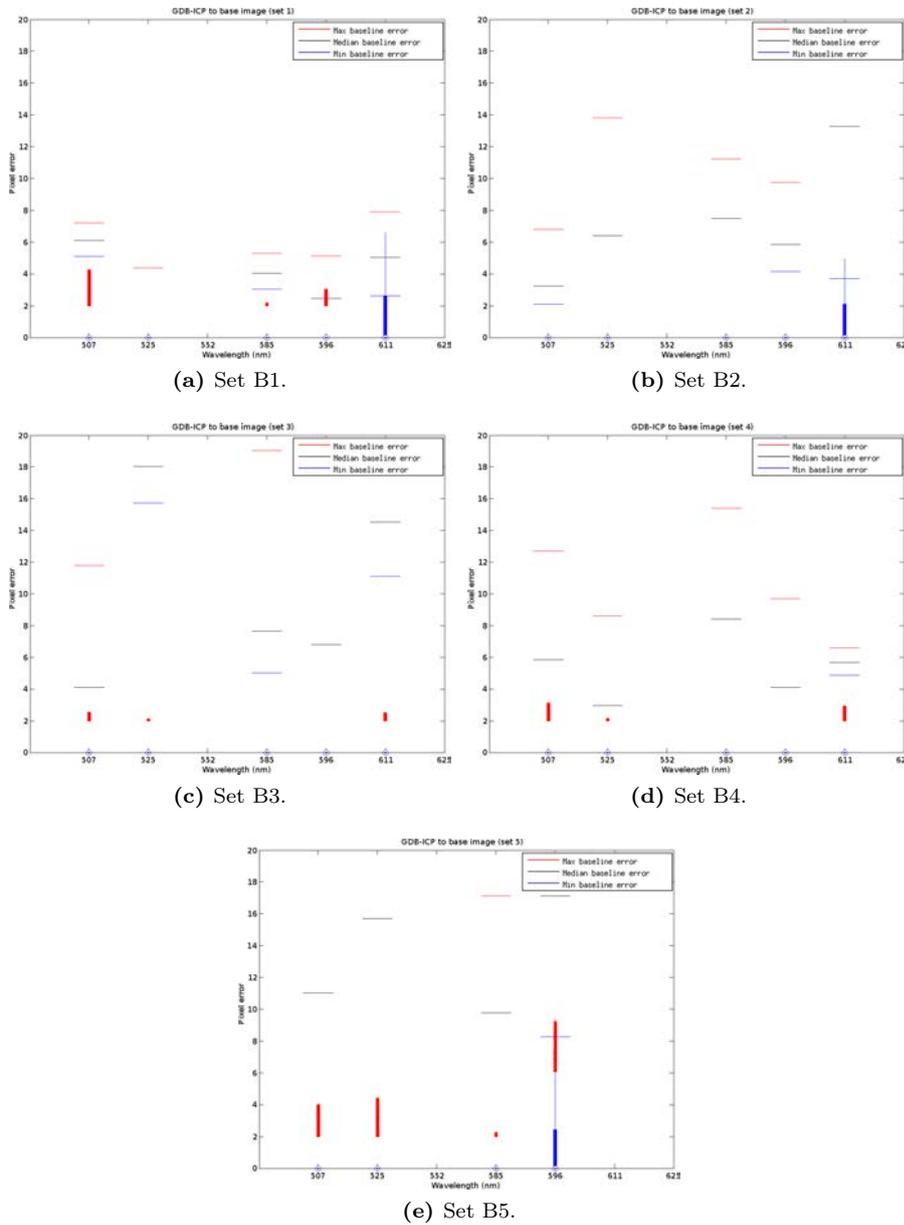


Figure 2.29: The generalized dual-bootstrap iterative closest point (GDB-ICP) errors with the semisynthetic set B; all images registered to a single base image captured at 540 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. [128]

complexity to quadratic, any combining of transformations or using intermediate images resulted in a notable decrease in performance.

The intermediate template strategy worked poorly for GDB-ICP. As any displacement due to inaccurate registration accumulates in the template and the quadratic transformation used by GDB-ICP could not fully compensate for the increasing complexity of the accumulated error, the template became blurry and caused further error in consequent registrations. The strategy of sequential registration did not work well with any of the evaluated methods. The large number of combined transformations lead to a rapid increase in registration error. The results of the image registration performance comparison on the synthetic and semisynthetic data are summarised in Tables 2.5 and 2.6, respectively.

Table 2.5: Median (Med.) and standard deviation (STD) of registration error for the synthetic set A. Init. stands for the initial error before registration, Base for registration to a single base image, NN for sequential registration, T for intermediate template and SP for shortest path. Lowest error for each method/set is displayed in bold.

Method		Init.	GDB-ICP				MI				RC			
Strategy			All	Base	NN	T	SP	Base	NN	T	SP	Base	NN	T
Med.	Set A1	6.1	1.1	1.2	5.5	1.0	2.1	9.1	2.4	3.5	1.3	1.2	1.6	1.4
	Set A2	11.9	1.1	2.0	8.5	1.7	2.3	18.9	6.0	4.4	1.5	3.0	1.7	1.9
	Set A3	18.0	1.2	5.2	11.5	1.4	20.4	186.2	25.0	4.3	2.2	10.5	3.0	2.3
	Set A4	22.7	1.2	4.6	22.3	2.4	8.0	141.1	18.0	7.1	3.0	7.7	3.9	3.9
	Set A5	31.1	1.2	8.2	47.3	2.2	47.5	275.9	42.7	58.3	4.6	35.1	7.1	5.3
STD	Set A1	1.6	0.4	0.7	1.8	0.5	4.1	15.1	2.5	8.9	1.5	2.4	1.3	1.4
	Set A2	3.4	0.5	2.1	3.3	1.0	5.8	33.8	5.1	10.9	3.6	4.7	3.1	3.0
	Set A3	6.2	0.6	3.7	5.1	0.9	19.3	73.2	17.8	15.8	5.8	9.9	6.2	5.8
	Set A4	7.6	0.8	4.0	8.0	1.7	13.3	54.4	8.9	19.3	8.4	11.5	7.4	8.5
	Set A5	10.9	0.9	8.2	6.3	1.8	29.4	83.2	29.7	30.8	12.6	13.4	10.8	10.2

Table 2.6: Median (Med.) and standard deviation (STD) of registration error for the semisynthetic set B. Init. stands for the initial error before registration, Base for registration to a single base image, NN for sequential registration, T for intermediate template and SP for shortest path. Lowest error for each method/set is displayed in bold.

Method		Init.	GDB-ICP				MI				RC			
Strategy			All	Base	NN	T	SP	Base	NN	T	SP	Base	NN	T
Med.	Set B1	4.1	0.0	0.0	0.0	0.0	4.8	4.6	0.0	4.6	2.6	4.0	3.3	2.6
	Set B2	6.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Set B3	7.7	0.0	0.0	0.0	0.0	3.0	4.2	0.0	3.5	0.0	0.0	11.0	0.0
	Set B4	5.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Set B5	15.7	0.0	2.4	7.3	2.3	2.1	3.2	10.2	3.1	2.5	3.6	5.4	2.9
STD	Set B1	1.0	0.9	0.6	0.5	0.7	2.9	2.4	1.0	2.8	3.4	5.2	4.3	3.8
	Set B2	2.7	0.0	0.0	1.4	0.0	1.2	1.5	1.5	1.4	1.3	1.3	3.0	0.9
	Set B3	3.0	0.1	0.8	0.8	0.5	1.9	2.5	2.8	2.1	1.4	7.1	6.7	3.0
	Set B4	2.1	0.1	0.0	1.3	0.0	1.0	0.8	0.9	0.5	1.0	0.8	1.3	0.8
	Set B5	5.0	0.9	1.8	5.0	1.5	2.1	4.7	8.6	4.1	4.4	3.8	7.7	3.8

The deterioration of the spectral image quality of the composed image when registering with GDB-ICP was similar to the image with linear displacement of ≈ 1 pixels, de-

spite the more complex transformation GDB-ICP uses. Similar results were observed on the spectral similarity metrics. The spectral image quality measures are visualised in Figure 2.30. The spectral similarity measures are shown in Figure 2.31.

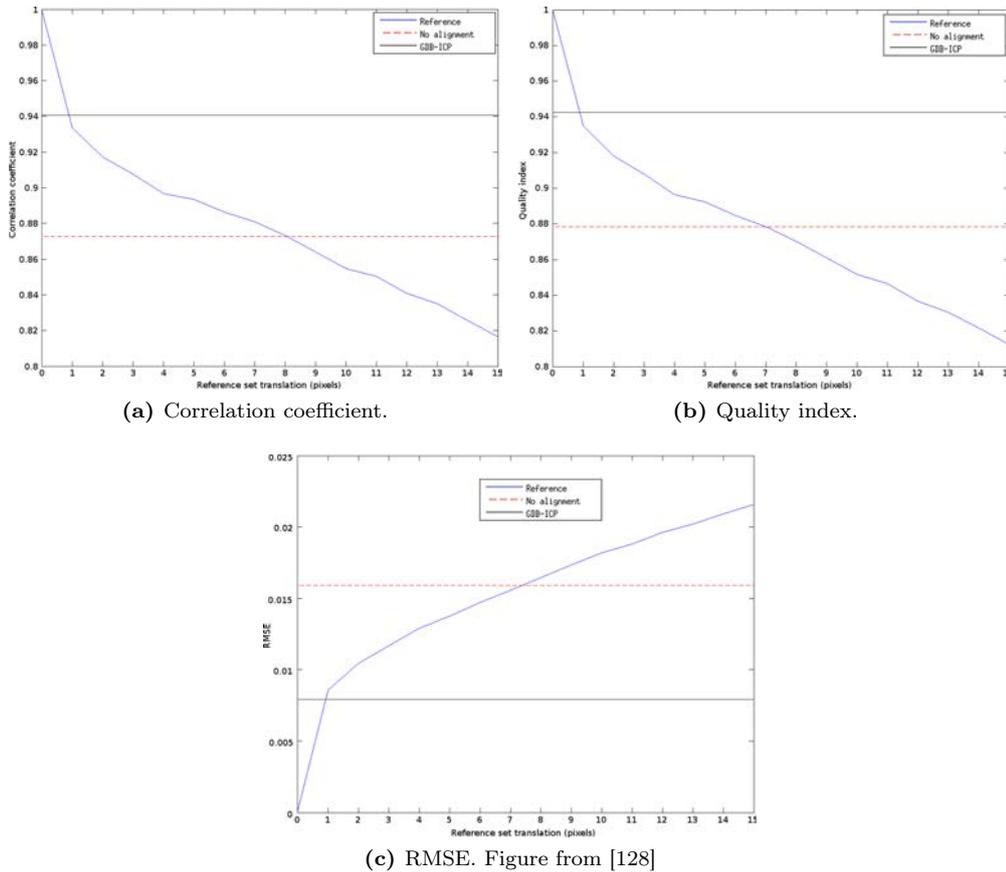
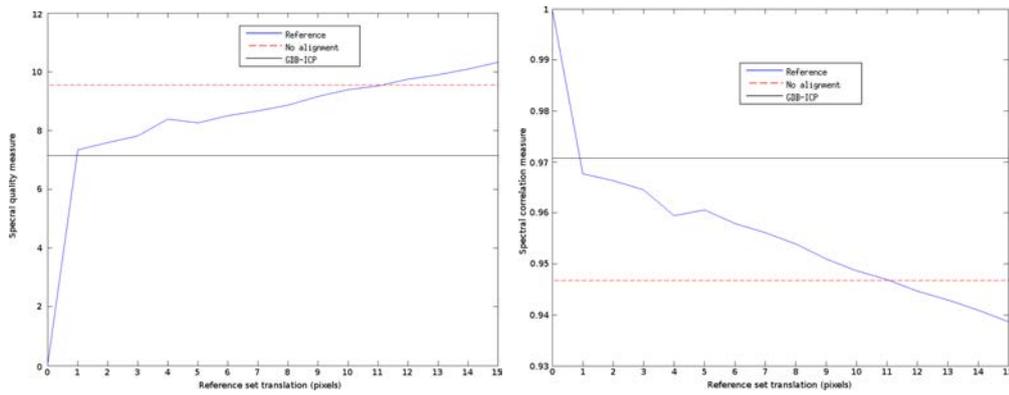


Figure 2.30: The spectral image quality of the image after registration with generalized dual-bootstrap iterative closest point (GDB-ICP) in relation to the reference, with the different metrics. Synthetic set A1, to a single base image.

2.6.6 Discussion

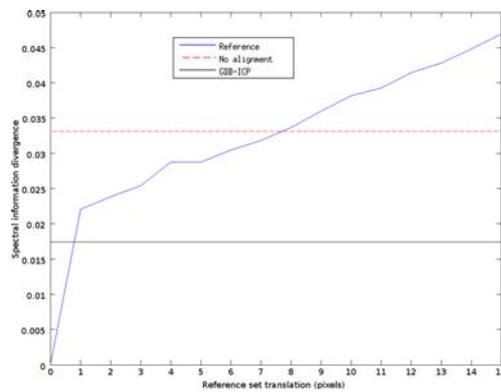
The registration errors tended to be more concentrated towards the image edges opposite to the optic disk (surrogate) where the blood vessels become less prominent. The retinal background was largely unable to provide either features for GDB-ICP or reliable region matches for MI or RC. The lowest registration errors were generally found in regions with high-contrast retinal blood vessels.

The low error rate of GDB-ICP was partially due to the restriction to the transformation complexity. The global transformation ensured that the registration error was reasonably low even in regions where there were few retinal features to guide the registration process.



(a) Spectral angle measure (SAM).

(b) Spectral correlation measure (SCM). Figure from [128]



(c) Spectral information divergence (SID). Figure from [128]

Figure 2.31: The spectral similarity of the image after registration with generalized dual-bootstrap iterative closest point (GDB-ICP) and the reference, with the different metrics. Synthetic set A1, to single base image.

In contrast, MI and RC generated, for some images, transformations where the parts of an image containing well defined features were registered to sub-pixel accuracy and the registration errors of tens of pixels could be found elsewhere.

The sequential registration strategy causes additional problems for non-deformable registration approaches (i.e., approaches that limit the complexity of the transformation used for registration). Even if the transformation error is negligible, as a non-rigid transformation is required to properly represent the deformation due to eye movement, combining the sequential registrations results in different channels being registered with transformations of different complexities. In addition, complex transformations require multiple transformation steps as the combined transformation cannot be expressed as a multiplication of the transformation matrices. Quantitatively, the accumulation of registration error significantly outweighed any benefit gained by having less difference in the image wavelengths.

Using the shortest path can help to avoid large wavelength dependent differences between images and limit the accumulation of error due to combined intermediate transforms. However, unless the length of the shortest path is the same for all images, the images will be registered with transformations of different complexity when using non-deformable approaches. The strategy is also sensitive to the choice of the similarity measure used in constructing the cost matrix (i.e., how the distance between images is measured when determining the shortest path). Based on the results, the shortest path strategy is not an effective strategy for the joint registration of spectral channel images.

In general, with the exception of using intermediate templates with MI, the evaluated registration strategies showed limited benefit. The increase in registration error due to even limited number of combined transformations outweighed the benefit of more similar image content due to smaller difference in wavelength.

Unless a threshold below which the registration error cannot be distinguished from the error inherent in the imaging system, such as in the case of the semisynthetic test set, can be determined, a clear distinction between a successful and a failed registration is application dependent and not trivial. In the case of spectral images, determining an error threshold for a successful registration can be approached through the effect of the registration inaccuracy on the spectra; a registration error of two pixels would require the spatial resolution of the spectral image to be downscaled by a factor of two to provide accurate spectra at each pixel.

Using maximal resolution, the system by Fält et al. [55] is capable of acquiring images at 2048×2048 (at the cost of increased noise). Therefore, a registration error of two or three pixels would still allow high-resolution images with accurate spectrum at each pixel. GDB-ICP is capable of registering most images with the median error within this accuracy. For the first set RC achieves similar median error, but some images contain regions with considerably higher errors. The large errors of MI in the blue wavelength images make the method unreliable in spectral image composition.

2.6.7 Summary

Of the compared methods, GDB-ICP outperformed the others in both number of successful registrations and the accuracy of registration. The experiments on synthetic and

semi-synthetic data showed that the registration error increases with increasing wavelength difference between the floating and the base image. The other well-performing methods were CC, MI and RC, but they could not match the accuracy and success rate of GDB-ICP.

The non-rigid registration approaches performed poorly due to significant differences between the channels. While locally accurate, the resulting transformations were unrealistic and inaccurate globally.

2.7 Illumination correction in spectral images

Due to the fundus camera optics, the curvature of the eye fundus and the limited aperture (pupil), the light used to illuminate the fundus is not evenly distributed onto the fundus. The uneven illumination field adds a low-frequency component to the intensity values of an image, which reduces the visual quality of the image and can have significant effect on automatic segmentation and diagnostics.

The estimation and correction of an uneven illumination field in fundus images has been studied (e.g., [6, 63, 76, 126, 127, 154]). Generally, the illumination field estimation includes (or is limited to) the background of the fundus image. In grayscale or RGB images the fundus background can be expected to form a reasonably uniform texture that does not introduce a significant bias to the estimation of the illumination field.

A number of different parametric models for estimating the illumination field in an image have been proposed. A fourth-order polynomial model for the illumination field of a fundus image is used in, e.g., [154]. A general bivariate polynomial field for location (x, y) is defined as

$$f_{\text{ill}}(x, y) = \sum_{i=0}^n \sum_{j=0}^n \beta_{ij} x^i y^j, \quad (2.32)$$

where n is the degree of the polynomial and $\beta_0 \dots \beta_n$ are the estimated model parameters.

Grisan et al. [76] assume the main contributors of the uneven illumination field in fundus images to be vignette and the possible luminosity glare at the image edges. They propose to model the vignette by elliptic paraboloids, and the glare by a sigmoid function defined as

$$S(d_r; p) = \frac{1}{1 + e^{-(d_r + s_1)/s_2}}, \quad (2.33)$$

where d_r is the radial distance of the point (x, y) to the image centre and $p = (s_1, s_2)$ is the vector of sigmoid function parameters. The elliptic paraboloid modelling the vignette is defined as

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ L_i \end{bmatrix} = \mathbf{R}_y(p_9) \mathbf{R}_x(p_8) \mathbf{R}_z(p_7) \begin{bmatrix} x - p_1 \\ y - p_2 \\ p_6 \left(\frac{(x-p_1)^2}{p_4} + \frac{(y-p_2)^2}{p_5} \right) + p_3 \end{bmatrix}, \quad (2.34)$$

where \mathbf{R}_x , \mathbf{R}_y and \mathbf{R}_z are rotation matrices and $p_1 \dots p_9$ are the estimated ellipsoid parameters.

Kang and Weiss [113] propose an illumination model that takes into account off-axis illumination, vignette and the tilt of the camera. The model is defined as

$$f_{\text{ill}}(u, v) = (1 - \alpha_0 r) \cos \gamma \left(1 + \frac{\tan \gamma}{f} (u \sin \varphi_x - v \cos \varphi_x) \right)^3 \frac{1}{(1 + (r/f)^2)^2}, \quad (2.35)$$

where f is the camera focal length, α_0 is a radial vignetting factor, γ is the camera tilt angle, φ_x is the rotation of the target around x-axis, (u, v) is the image point (x, y) relative to the image principal point, and $r^2 = u^2 + v^2$.

All the presented models have been designed for grayscale or RGB images. Spectral fundus images have a number of qualities that affect illumination correction. A fundus background with a uniform texture is not a reasonable assumption for spectral fundus images. Due to the wavelength-dependent absorption and scattering of different retinal and subretinal features, and the varying penetration depth of different wavelengths, the information content at what constitutes as the background in grayscale or RGB images will vary for different wavelengths.

In imaging systems where the channels of the spectral image are acquired individually (i.e., with a separate exposure for each channel), the resulting illumination fields of the channel images are independent of each other. Any inter-channel movement of the eye or the camera, or difference in exposure time can cause significant nonlinear changes in the spectra of the spectral fundus image.

2.7.1 Illumination field estimation using the image spectra

The simplest approach to spectral image illumination correction is to individually estimate and correct the illumination field for each channel. However, the reflectance and scatter of retinal tissue is wavelength dependent, and longer exposure times are generally required towards the shorter wavelengths (due to the increased absorption of the crystalline lens at those wavelengths). Consequently, the range of illumination field shapes may vary between channels (requiring different models for accurate estimation).

Furthermore, channel-wise individual approximation does not take into account the consistency of the image spectra. The (channel-wise) utilisation of any method that performs contrast enhancement or otherwise modifies the image intensities relative to other channels will likewise result in incorrect spectra.

The consistency of the spectral information can be used to guide the channel-wise illumination field estimations. The reflectance from similar retinal tissues can be assumed to produce similar spectra. As the illumination fields are channel-wise independent, the reflectance spectra from a specific type of tissue contain information on different parts of the illumination fields depending on their spatial location in the image.

REFERENCE SPECTRA FOR GUIDING THE ILLUMINATION FIELD ESTIMATION

If a set of image spectra that sufficiently represent the spectral variation of a specific retinal structure while containing uniform bias (i.e., the illumination field is flat at their spatial locations) for all channels can be found, the spectra can be used as reference for the

other spectra where there is inter-channel difference in the bias. Subtracting the reference spectrum with similar underlying histological parameters (i.e., similar tissue structure), \mathbf{s}_{ref} , from a spectrum \mathbf{s}_i , the resulting residual vector $\mathbf{res}_i = \mathbf{s}_i - \mathbf{s}_{\text{ref}}$ should consist only of the illumination field bias and noise (from various sources). Fitting channel-wise illumination fields to the residuals provides illumination field estimates where the image spectra remain consistent after correction.

Retinal blood vessels have several characteristics that make them a good candidate for the reference spectra. Due to the strong absorption of haemoglobin (excluding the longer wavelengths), the spectra of the retinal blood vessels can be expected to have a lower variance than the background of the fundus, as the light reflected from the background can be the product of reflections from multiple tissue layers. In contrast, with the exception of smaller retinal vessels, the amount of light reaching the fundus camera after passing twice through a retinal vessel and interacting with the tissue below the vessel can be assumed to be negligible.

The main contributors to the reflectance spectra of retinal blood vessels are therefore absorption and scatter by haemoglobin and a low-frequency bias caused by the uneven illumination field. The haemoglobin absorbance depends on the oxygen saturation level of the blood (the absorption spectrum is different for oxygenated and deoxygenated haemoglobin). It should be noted that, unless the acquisition time for the whole spectral image is short, the oxygen saturation level at a specific region in a vessel will vary between the channels due to natural flow of blood. Thus, directly using the haemoglobin absorption spectra to guide the illumination field estimation is not applicable.

Furthermore, the retinal vasculature is relatively easy to segment. Numerous automatic segmentation methods are available, e.g., [64, 134, 158, 189, 208, 213].

EXTRACTING THE REFERENCE SPECTRA

As the set of reference spectra is defined by the uniformity of illumination, an initial estimation of the illumination field for each channel is required for the extraction of the reference spectra. A polynomial model is fitted to the blood vessel intensity values for each channel using least-squares approximation. The resulting n illumination field estimates $g_{1\dots n}$ (where n is the number of wavelength channels) model the low-frequency intensity component within the channels, but provide no consistency over the spectral dimension.

Based on the estimated channel-wise illumination fields, regions of the field with similar estimated illumination fall-off can be extracted. Selecting a set of linearly spaced thresholds t with values between $[\max g_i, \min g_i]$, the set of regions with similar estimated fall-off can be defined as

$$S_{w_i} = t_i \leq g_w \leq t_{i-1}, \quad (2.36)$$

where g_w is the estimated illumination field of wavelength channel w .

As the amount of light projected onto the eye remains the same and the angle of rotation between the eye and the camera should be relatively small, regions of the same set level at different channels have a similar relation to the peak illumination of that channel. Therefore, the effect of the illumination field on the spectra can be expected to be similar for the region where the masks of specific level overlap for all channels.

The regions where the reference spectra can be extracted are defined as

$$S_{\text{ref}} = \bigcup_i (S_i \cap S_{i+1} \cap \dots \cap S_n). \quad (2.37)$$

While the levels of bias due to the intensity fields in S_{ref} vary, the bias can be expected to remain relatively consistent for all the wavelength channels. The reference spectra used to calculate the residual vector are taken from the spatial locations in S_{ref} that correspond to a blood vessel position. With the set of reference blood vessel spectra with mostly coherent wavelength dimensional illumination bias, residual vectors can be calculated for all the image spectra corresponding to a blood vessel location.

However, as the haemoglobin absorption is wavelength dependent and especially low towards the longer wavelengths, the amount of haemoglobin contribution to the image spectra is different between the channels. As a consequence, the blood vessel intensities in the longer wavelength channels are influenced by the reflections from the retinal tissues below the blood vessel. In practice, the blood vessel spectra can be expected to have a higher variance in the longer wavelength component, which is likely to cause error in the selection of a corresponding reference spectrum for image spectrum \mathbf{s} .

To reduce the influence of the longer wavelength channels in the selection of the reference spectrum, the distance between \mathbf{s} and \mathbf{s}_{ref} is weighted by the haemoglobin absorption spectrum. The reference spectrum for the spectrum \mathbf{s} corresponding to the location on a blood vessel (x, y) is selected based on a weighted distance

$$\operatorname{argmin}_i \left(\sum \omega \sqrt{(\mathbf{s} - \mathbf{s}_{\text{ref}_i})^2} \right), \quad (2.38)$$

where $\mathbf{s}_{\text{ref}_i}$ is the i :th reference spectrum and ω is a weighting factor based on the wavelength dependent haemoglobin absorption coefficient.

With a reference spectrum for each blood vessel spectrum in the spectral fundus image, the residuals for estimating the illumination field are defined

$$\mathbf{res}_i = \mathbf{s}_i - \mathbf{s}_{\text{ref}_i}. \quad (2.39)$$

The global illumination field is estimated by fitting n channel-wise independent illumination fields g_1, \dots, g_n to the channel-wise residual values. Dividing the wavelength channels using the calculated illumination field estimates reduces the bias due to uneven illumination while retaining the consistency of the image spectra.

2.7.2 Experiments and results

The performance of the proposed illumination correction method was evaluated using semisynthetic images based on six-channel spectral images acquired with the system by Styles et al. [216]. The system acquires all six channels in a quick succession using a single exposure. Therefore, while the formation of an illumination bias field cannot be avoided, all six channels can be assumed to have an identical (or close to identical) illumination field.

Computing the mean over the spectral dimension of the spectral image, a good estimate of the illumination field can be achieved as channel-wise differences are averaged while the illumination field is retained. Estimating the illumination field of the mean image \bar{I} using a fourth-order bivariate polynomial, the illumination field bias can be corrected by dividing each channel of the spectral image with the estimated illumination field f_{ill} , resulting in reference image I_{ref} with (close to) uniform illumination field on each channel.

A test set of 5 semisynthetic images were gained by corrupting each channel of each of the 5 reference images I_{ref} using a different fourth-order bivariate polynomial (simulating a biased illumination field) with randomly sampled parameters (normally distributed samples with standard deviations based on parameters of estimation of f_{ill}).

The artificially corrupted semisynthetic images were corrected using both the proposed method and (fourth-order polynomial model) channel-wise correction for reference. The difference between the reference spectral image I_{ref} (with no synthetic illumination field bias) and the corrected images was measured using spectral distance measures SAM and SID [230], defined as

$$\text{SAM} = \cos^{-1} \left(\frac{\sum_{i=1}^d x_i y_i}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}} \right) \quad (2.40)$$

$$\text{SID} = \sum_{i=1}^d p_i (\log p_i - \log q_i) + \sum_{i=1}^d q_i (\log q_i - \log p_i), \quad (2.41)$$

where

$$p_i = \frac{x_i}{\sum_{j=1}^d x_j}, \quad q_i = \frac{y_i}{\sum_{j=1}^d y_j}, \quad (2.42)$$

and d is the number of channels in the spectra. The results, visualised in Figure 2.32, show smaller spectral distance to the reference image after correction with the proposed method. The channel-wise correction does not retain the consistency of image spectra, and consequently the spectral distance to the reference image is increased. The results are consistent for the majority of the images in the test set.

For qualitative evaluation of the performance of the presented illumination correction method, a set of 7 spectral fundus images with visually identifiable unevenness in the illumination field were chosen from the DiaRetDB2 data set. The selected spectral retinal images were processed using the presented spectral image illumination correction method, and both the corrected channel images and RGB projections (gained by integrating the image spectra with a set of virtual filters) were visually evaluated.

Visual examination of the results suggests that the method was able to reduce the illumination field bias in the spectral image while respecting the consistency of the spectral dimension. An example are shown in Figure 2.33, where the wavelength channels show reduction in the illumination field bias. Two RGB projections are shown in Figure 2.34.

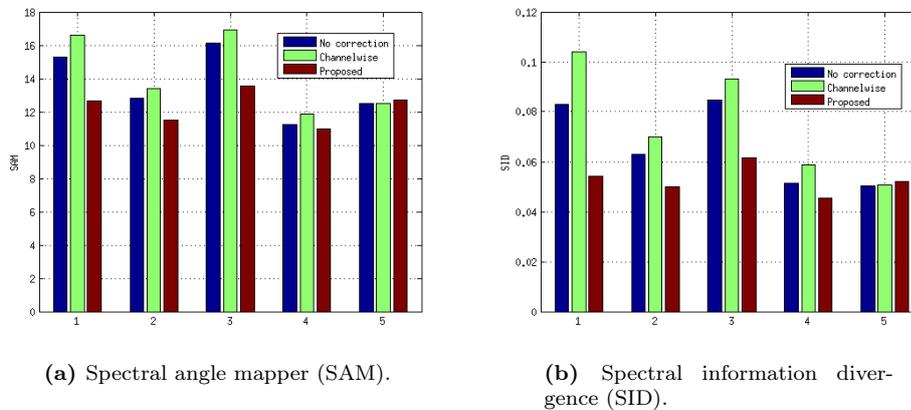


Figure 2.32: Spectral difference to the reference spectral image (without synthetic illumination field bias).

2.7.3 Discussion

As the method relies on retinal blood vessels to extract the reference spectra, the lack of blood vessels in locations critical to the illumination field estimation can limit the performance of the method. An example of this can be seen in Figure 2.33b, where a region of higher intensity (due to illumination field bias) remains after correction.

The consistency of the image spectra is difficult to evaluate without ground truth measurements. Indirect proof of the proposed method retaining the consistency of the image spectra is presented in Section 4.4.2, where the image spectra are used to estimate histological parameters of the retina.

2.7.4 Summary

The fundus camera optics, the curvature of the eye fundus and the limited aperture cause an uneven illumination field in fundus images, adding a low-frequency component to the intensity values of an image and reducing the quality of the image spectra. Channel-wise individual approximation and correction of the illumination field does not take into account the consistency of the image spectra.

A method for correcting the uneven illumination field in spectral images while retaining the spectral consistency, based on extracting reference haemoglobin spectra was presented. A set of 7 spectral retinal images and 5 semisynthetic test images were processed using the presented spectral image illumination correction method. The method was able to reduce the illumination field bias in the spectral image while respecting the consistency of the spectral dimension.



Figure 2.33: Channel images before and after correction with the presented method.

2.8 3D-reconstruction of the retina from spectral images

The images from spectral fundus imaging systems that acquire each spectral channel using separate exposures offer a unique possibility for image processing. As some eye movement between the exposures is unavoidable, some disparity of spatial structures will be present in the data. Using stereoscopic principles, the disparity in the spatial location between two (or more) images can be used to estimate depth information at that spatial location.

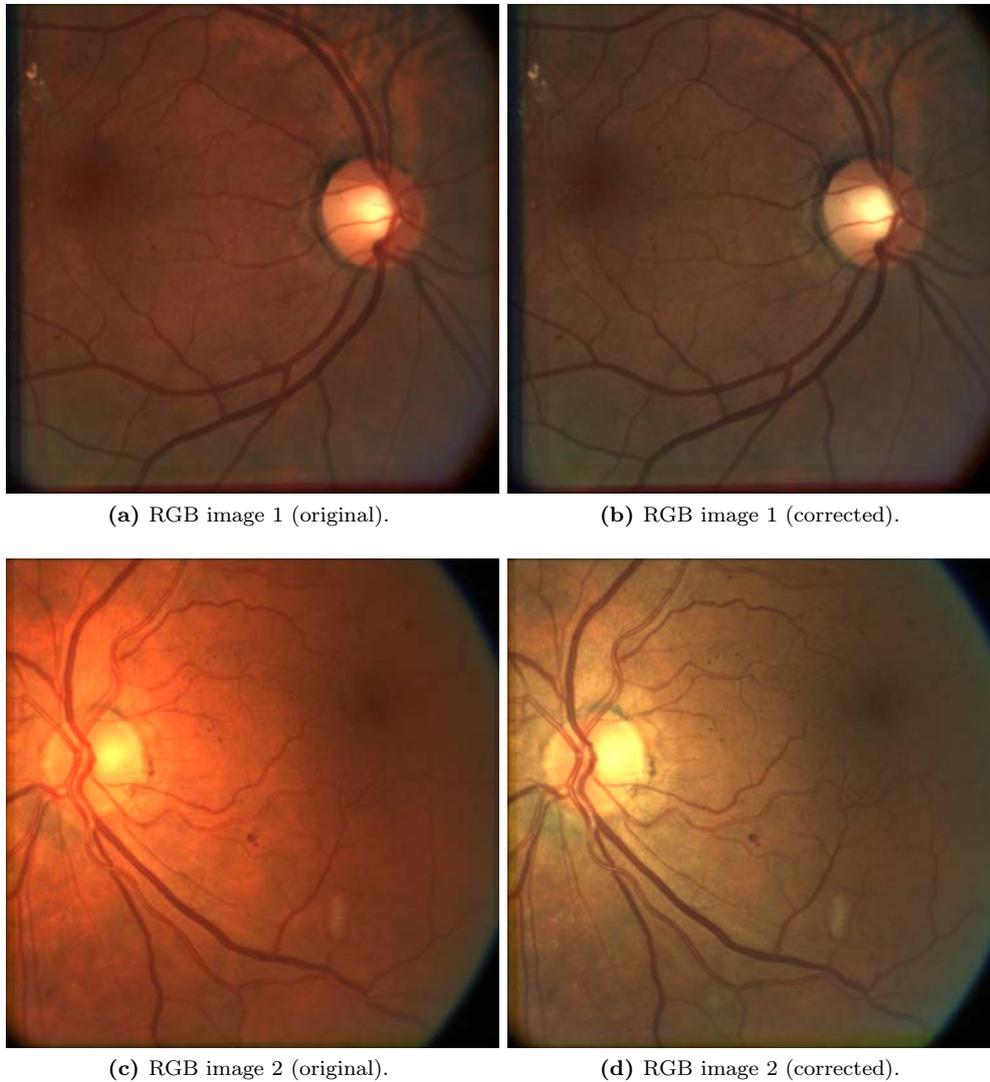


Figure 2.34: RGB projection images before and after correction with the presented method.

Stereoscopic reconstructions of the fundus have clinical value, as stereo fundus photography is used to diagnose e.g., macular edema and glaucoma. Reconstructions of (parts of) the eye fundus from fundus images have been suggested for the diagnosis of glaucoma [218], study of vasculature [140, 146], and other ophthalmic diagnostics [38, 39, 47, 73].

Ideally, a stereoscopic reconstruction from fundus images has several benefits over OCT which can be considered as the *de facto* method for extracting structural information on the retina. Firstly, the FOV of a fundus image is significantly larger than that of a typical

OCT reconstruction. Multiple OCT reconstructions could be taken to acquire data over the same FOV, but a globally accurate reconstruction would not be guaranteed, as the acquisition could be difficult and the registration of the reconstructions would not be trivial.

In practice, however, the acquisition of retinal images with sufficient disparity for accurate stereo reconstruction can be challenging. Martinello et al. [145] studied the light paths in reconstruction with both stereo fundus camera and monocular stereo reconstruction (i.e., from images with the target in different poses). They concluded that despite the changes in the pose of the eye, the light reflected from the retina travels through the centre of the pupil, in practice limiting the disparity between the images and consequently increasing the reconstruction error. Giancardo et al. [73] suggest that the problem of disparity in monocular retinal images can be avoided by using an aperture smaller than the eye pupil.

While the depth information of the individual OCT scans considerably surpasses the accuracy of what a reconstruction using stereoscopy can be expected to achieve, the inexpensiveness and larger FOV of 3D-reconstructions based on (monocular) retinal images can be valuable for diagnostics, image registration, image post-processing or other image analysis applications. This section presents a method for the 3D-reconstruction of the retinal shape from (monocular) retinal images. The method is evaluated quantitatively on synthetic data, and reconstruction examples on real medical data are given.

2.8.1 Methods

The method presented in this work is based on the stereoscopic reconstruction method for fundus images by Lin and Medioni [139]. There are three major modifications to the original algorithm: the replacement of SIFT and MI as the features for the coarse correspondence and dense correspondence search with SURF and binary Robust Invariant Scalable Keypoints (BRISK), the use of multiple images to calculate the dense reconstruction, and the addition of a step to remove (dense) correspondences with multiple good matches. The selection of the combination of SURF and BRISK features is justified in Section 2.8.2.

IMAGE PREPROCESSING

As the features used for determining the coarse correspondence rely on image edges, pre-processing is performed to increase the contrast of edge responses and to facilitate feature extraction. Before calculating the edge responses, histogram normalisation is performed on the images (using the same μ and σ for all images) to extend the dynamic range in the images. Extensive stretching of the dynamic range tends to generate significant amounts of noise. A non-local mean filter [34] is applied to the images to reduce the increased noise and to avoid false edge responses.

The original method by Lin and Medioni [139] uses edge response images calculated from the contrast enhanced images using a Sobel filter. For the modified method the use of the contrast enhanced images directly was found to produce better results.

2D-2D CORRESPONDENCE

The initial coarse 2D-2D correspondence is based on SURF features extracted from the preprocessed images. Instead of directly matching the SURF features (based on the distance of feature vectors) and estimating the homography from the matches, an iterative matching scheme similar to the ICP algorithm is employed to produce more accurate and robust homography estimations. The process is visualised in Figure 2.35.

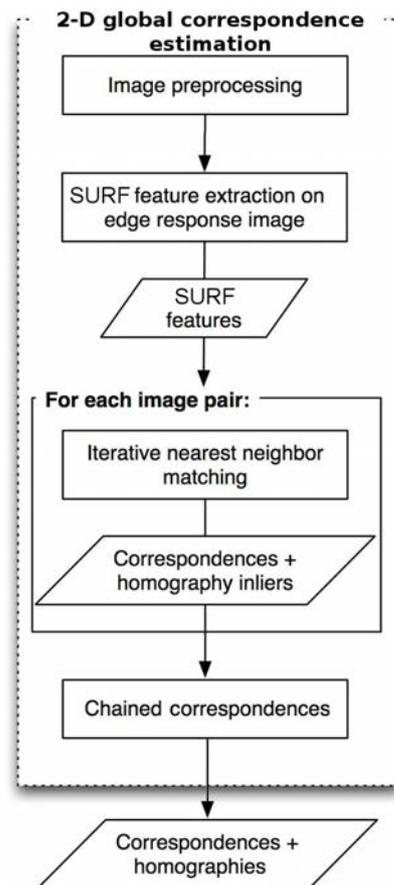


Figure 2.35: Flowchart for determining the global 2D-2D correspondence (modified from [139]).

The robust homography estimation starts with the estimation of the homography as identity (i.e., a transformation that does not change feature point positions). The SURF feature points (of image I_i) are matched by pairing them with the feature point (of image I_j) with the most similar feature vector (measured as Euclidean distance) within a (initially large) radius r . Using random sample consensus (RANSAC) [62], a robust estimation of the homography is calculated based on the matches, and the feature points are transformed using the estimated homography. For the next iteration, the search radius r is halved. The steps are repeated until the search radius falls into a predetermined

minimum (r_{min}).

In classic RANSAC, the best homography is determined by the size of the consensus set (i.e., set of points that would produce the transformation within some error). If a large number of the feature points receive poor matches from the previous step, a large consensus set might not guarantee a correct homography. To reduce the effect of poor matches, each match is assigned a weight based on the orientation difference (good matches are expected to have similar orientations) of the feature points. Instead of the consensus set size, the sum of weights of the feature points in the consensus set, ω , is used to rank the estimated homographies.

Correspondence is estimated for all combinations of image pairs. The sum of weights ω associated with the estimated homographies are used as a measure of correspondence accuracy for refinement of the pairwise correspondences.

Another round of pairwise correspondence estimations is performed on the combinations of image pairs to refine the correspondences (see Figure 2.36). When refining the homography H_{ij} , instead of starting with an identity homography, the estimation is initialised as a combination of all transformations in the shortest path between the image pair, i.e., $\prod H_S$ (chained correspondence estimation). As the initial estimation is assumed to be reasonably accurate, the search radius r is initialised to r_{min} (resulting in the estimation converging after a single iteration). ω_{ij} is updated, and all homographies with their shortest path including the step H_{ij} are re-evaluated. The process continues until new estimates with improved scores ω can no longer be found.

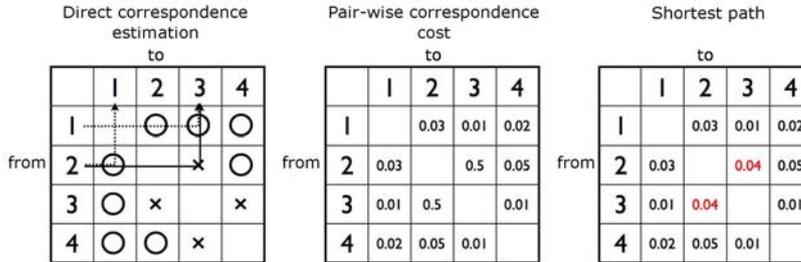


Figure 2.36: Refinement of 2D-2D correspondence (modified from [139]).

3D CORRESPONDENCE AND RECONSTRUCTION

Once the pairwise coarse correspondences are estimated and refined, the image with the minimum distance to all the other images (along the shortest paths) is selected as the reference image, I_{ref} , for the camera pose estimation. As with [139], the reconstruction is defined for points on an opaque retinal surface imaged with pinhole cameras. While these assumptions are a gross simplification of the optical system and the light paths in retinal imaging, they facilitate the estimation of the depth information.

The plane+parallax method is used to estimate a fundamental matrix F_i (with respect to I_{ref}) for each image I_i . An initial camera pose is estimated by projecting the feature point locations of I_i (transformed to the reference frame) and of I_{ref} , x_i and x_{ref} , into 3D using the (naïve) assumption that the points lie on a plane in front of and parallel to

the reference camera. Using the resulting set of 3D-points, X_i , the pose for the camera associated with the image I_i (with respect to the reference camera) is estimated by minimising the reprojection error:

$$\epsilon(P_i) = \sum \|x_i - P_i X_i\|^2 \quad (2.43)$$

where P_i is the projection matrix associated with image I_i . The procedure is repeated for all the cameras (i.e., camera poses corresponding the individual images).

After the initial pose estimation for all cameras, X_i is re-estimated using triangulation based on the camera position estimates. After the triangulation step, estimates for both the camera pose and the 3D-structure exist.

Once the 3D-structure and camera parameters have been estimated, the images are rectified and dense BRISK features are extracted from the images. Based on the feature vectors, feature point correspondences are identified across the images (e.g., the pair of matched points a_{ij} in I_i and I_j corresponds to the point b_{kl} in I_k and I_l). The dense BRISK features are matched for each image pair. To reduce false matches, matches where the relative distance to the second closest match candidate is below threshold t are discarded.

The dense matches are used to refine the estimates by bundle adjustment, i.e., both the camera parameters and the 3D-point positions are adjusted simultaneously. The reprojection error is minimised again, this time allowing both the camera parameters and the 3D-locations to vary,

$$\epsilon(P_i, X_i) = \sum \|x_i - P_i X_i\|^2. \quad (2.44)$$

2.8.2 Experiments and results

A number of features were considered for the dense point matching for the proposed method with SIFT and SURF for the 2D-2D correspondence. The features SURF, BRISK, MI and fast retina keypoint (FREAK) were evaluated in the framework of the method using three synthetic data sets.

The method for generating synthetic images described in Section 2.6.4 was used to generate a data set with simulated changes in eye pose with respect to the camera. The synthetic data consisted of retinal images projected onto a segment of a semi-sphere. A set of images with artificial depth disparities was attained by rotating the semi-sphere and reprojecting the image points back onto a plane. An example of the resulting synthetic images are shown in Figures 2.37.

The reconstruction was performed using different features for the dense matching. The implementations of both the original method by Medioni et al. and the proposed extensions used in the evaluation were implemented by the author and co-workers Jarmo Ilonen and Antti Hannuksela.

As the underlying topology of the synthetic test images is known, the reconstruction error can be calculated as the distance from the reconstructed points to the semi-sphere segment. The reconstruction error for reconstructed 3D-coordinates V is defined

$$\hat{e} = \sum_i \sqrt{(V_{zi} - \hat{f}(V_{xi}, V_{yi}))^2}, \quad (2.45)$$

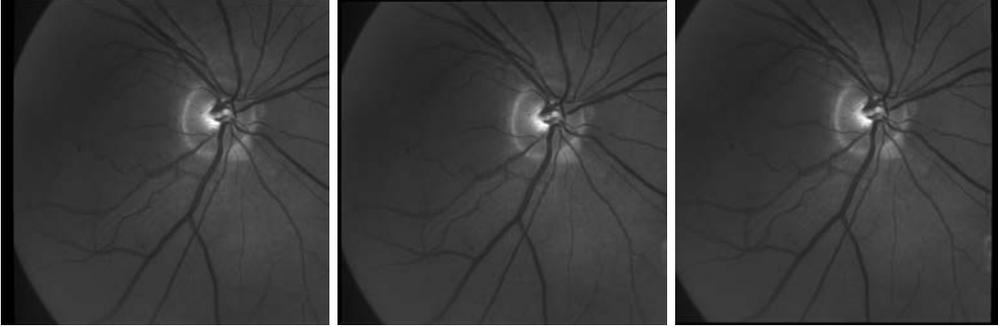


Figure 2.37: Synthetic evaluation data set 1.

where \hat{f} is the bi-quadratic function defining the topology used for distorting the synthetic test images.

BRISK features were found to produce the smallest reconstruction error for both synthetic sets. The reconstruction errors for the different features are summarised in Table 2.7. The reconstruction results for the synthetic data sets are shown in Figure 2.38.

Table 2.7: Mean reconstruction errors using different feature combinations. The values in the parentheses are the reconstruction errors **without** requiring multiple images to support point pairs.

	Set 1	Set 2	Set 3
Lin & Medioni	32.88	26.07	21.21
SIFT + BRISK	13.42 (18.45)	11.88 (15.46)	9.55 (10.16)
SIFT + FREAK	14.80 (22.13)	15.20 (17.83)	11.96 (13.95)
SIFT + SURF	39.57 (42.04)	33.32 (39.42)	29.94 (37.91)
SURF + BRISK	11.54 (16.79)	12.46 (12.67)	9.34 (10.42)
SURF + FREAK	14.07 (21.03)	16.26 (18.21)	13.01 (14.04)
SURF + MI	34.82	30.14	20.96
SURF + SURF	40.47 (42.79)	33.76 (39.23)	30.21 (37.55)

To evaluate the reconstructions with real medical image data, four sets of retinal images were acquired with different parts of the retina centred. Before the acquisition of each retinal image, the subject was asked to focus their eye in different direction to produce disparity within the medical test sets. An example of a medical test set is shown in Figure 2.39. The reconstructions from the medical test sets are shown in Figures 2.40 and 2.41. Before reconstruction, the images were corrected for geometric distortion using the approach detailed in Section 2.5.2. While the full set of reconstructed points are visualised in the plots, median filtering was applied to the point set used for the linear interpolation surface for better visualisation.

2.8.3 Discussion

Judging by the synthetic data, the proposed method is able reconstruct the 3D-structure of the eye fundus relatively well. Majority of the error in the reconstructions from the

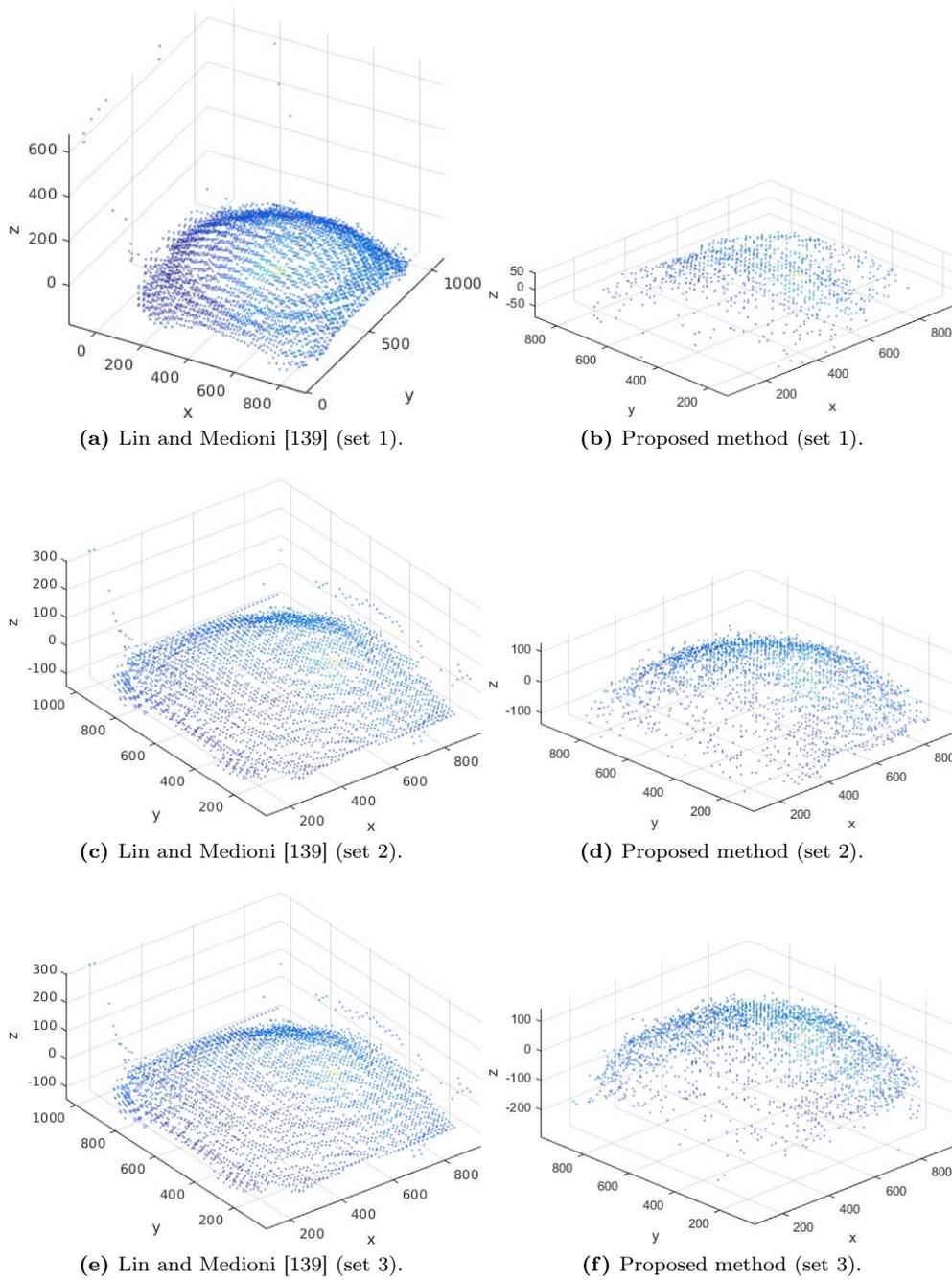


Figure 2.38: Reconstruction results on synthetic data (using BRISK) for dense matching.

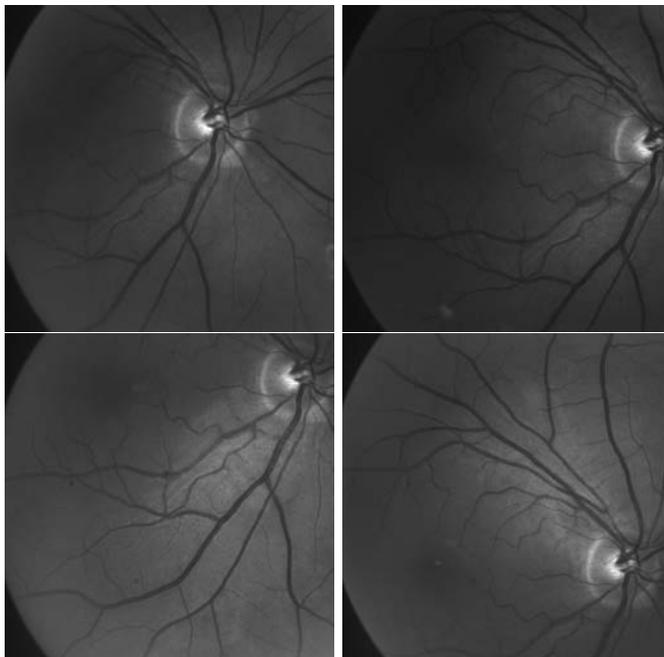


Figure 2.39: An example medical evaluation data set.

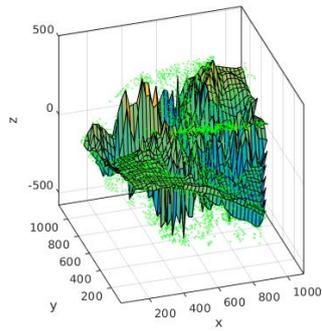
synthetic data was due to outliers, which can usually be removed in post-processing.

The curvature of the underlying structures in the synthetic data sets is exaggerated in comparison to the actual curvature of the fundus region captured by a fundus camera. While the reconstructions from the medical data seem close to planar compared to the reconstructions from the synthetic data, the result is not unexpected as the fundus camera used for acquisition of the images has a high magnification and a relatively low FOV.

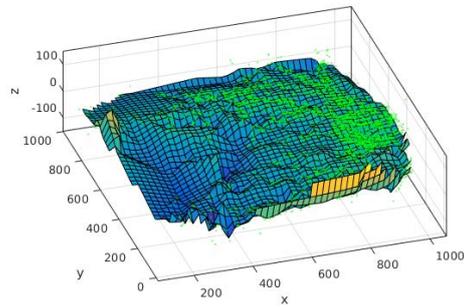
The original method by Lin and Medioni [139] failed with the medical data sets due to a misalignment in the 2D-2D-correspondence phase. As the homography related to one of the images was incorrect, the points from the image were estimated on a plane separate from the rest of the reconstruction.

2.8.4 Summary

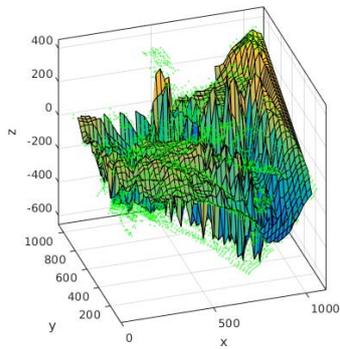
A method for the 3D-reconstruction of the retinal shape from retinal images based on the method by Lin and Medioni [139] was presented. The method was evaluated on both synthetic and medical data. For the synthetic data, the proposed method achieved lower mean reconstruction error than the original method by Lin and Medioni. On the medical data, the proposed method was able to produce a reasonable reconstruction where the original method failed.



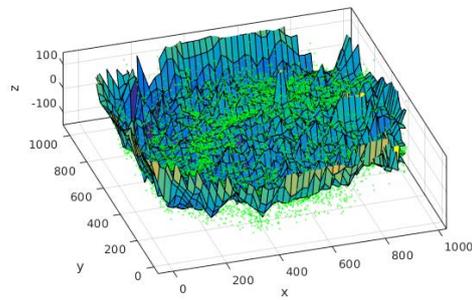
(a) Lin and Medioni [139].



(b) Proposed method.

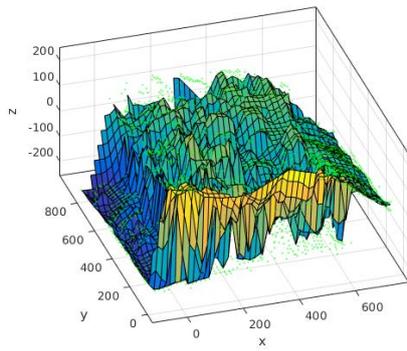


(c) Lin and Medioni [139].

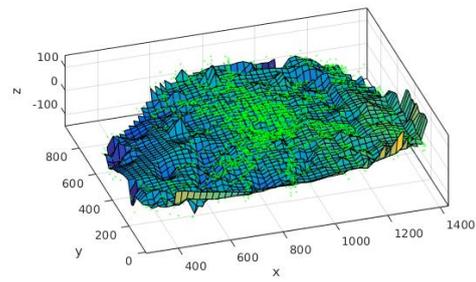


(d) Proposed method.

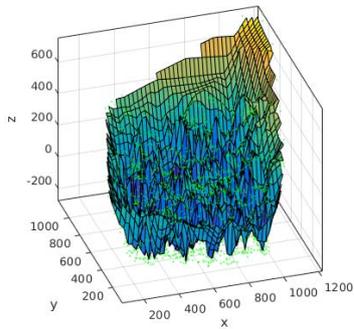
Figure 2.40: Reconstructions from medical data. Reconstructed point clouds with linear interpolation surface. The point sets used for the surface were median filtered for better visualisation.



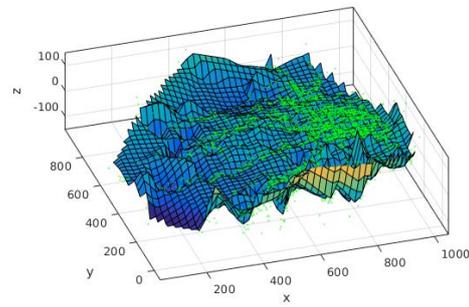
(a) Lin and Medioni [139].



(b) Proposed method.



(c) Lin and Medioni [139].



(d) Proposed method.

Figure 2.41: Reconstructions from medical data (continued). Reconstructed point clouds with linear interpolation surface. The point sets used for the surface were median filtered for better visualisation.

Spectral image database of diabetic retinopathy patients

3.1 Introduction

Publicly available benchmark databases with clearly defined evaluation protocols benefit the development of, e.g., computer vision methods. The evaluation of the method performance is facilitated as the performance reported with descriptions of competing and the state-of-the-art methods are directly comparable. Furthermore, publicly available databases can usually be expected to contain more data for training and testing than an individual research group would otherwise (easily) have access to.

Compared to typical computer vision benchmark databases, medical image databases have a number of important differences. Medical image databases typically concentrate on one organ imaged in a specific modality (e.g., MRI, ultrasound or positron emission tomography (PET)). The number of images is often limited due to the number of available subjects, the cost of image acquisition, and the ethical considerations related to imaging human subjects. Laws related to the privacy of medical information may also limit the publication of medical data.

Medical image databases often contain some relevant information on the images related to the purpose of the database. Typically at least the general diagnostic information is published with the images. A number of databases provide gold standards for structural information such as the location of the optic disk and macula, and the blood vessel profiles. Significantly fewer databases seem to offer gold standards for lesions or pathological regions.

This chapter discusses the acquisition of images and gold standards of lesion locations for the spectral fundus image database DiaRetDB2. The spectral information coupled with gold standards for multiple lesion types provide a good basis for method development in automatic detection and classification of retinal lesions.

3.2 Public fundus image databases

A number of retinal image databases for various purposes exist, with representations of different pathologies. Fumero et al. [66] present a fundus image database with accurate gold standards of the optic nerve head for the development and evaluation of automatic methods for the detection of glaucoma.

Niemeijer et al. [159] provide a database of 100 fundus images. The database includes annotations of microaneurysms from four expert ophthalmologists. The data set is a part of the Retinopathy online challenge [187], a platform for standardised evaluation for automatic microaneurysm detection methods.

Niemeijer et al. [158] compare a number of vessel segmentation methods on a database of 40 images with manually segmented vessel trees. In addition to the publication of the comparison, the database has been made publicly available.

Zhang et al. [254] present an online depository of fundus images with 650 images containing expert annotations focused on biomarkers related to glaucoma diagnosis. A quantitative benchmarking method focused on optic disk and cup segmentation and cup-to-disk ratio is provided.

Budai et al. [35] have composed a database of 15 high-resolution images of healthy subjects with gold standards of the locations of vessels (including differentiation between veins and arteries), macula, and the optic disk. The database was later extended with 15 images of patients with DR and another 15 of patients with glaucoma [167].

Al-Diri et al. [7] present a reference data set of 16 images for automatic segmentation and measurement of retinal vessels. A semi-automatic segmentation method is presented and the resulting 193 vessel segments have been published along with the reference images.

Hoover et al. [97] have published a data set of 20 images (ten of healthy and ten of pathological retinas) with manually drawn reference for blood vessels. The ten pathological retinas were selected based on abnormalities or obstructions in the vessels.

Klein et al. [122] present a database for the evaluation of automatic methods for analysing the severity of DR. The database consists of 1200 images with severity gradings from three ophthalmologic departments.

Owen et al. [176] validate a number of automatic methods for vessel tortuosity measurement on a data set of 28 fundus images from 14 young children. The vessel profiles from two human observers are included in the data set.

Staal et al. [213] present a database of 40 images, 7 with pathology, for the evaluation of automatic vessel segmentation. Manual vessel segmentations by three non-experts trained by an ophthalmologist are included.

Significantly fewer examples of databases of images including spectral information can be found. Styles et al. offer the data set of 35 spectral fundus images described in [216] on request, but to the knowledge of the authors, no previous publicly available database of spectral retinal images with a gold standard for multiple region of interests (ROIs) exist.

3.3 DiaRetDB2 spectral retinal image database with gold standard

3.3.1 Human subjects and ethical considerations

The DiaRetDB2 database consists of the spectral fundus images of 72 subjects ranging from 21 to 81 years of age. 55 of the subjects were diagnosed with various stages of DR, while 17 control subjects had healthy retinas. Of the diabetic subjects 30 were female and 25 male. The control group consisted of 4 females and 13 males. Due to lengthy acquisition time (total of ~ 30 min for a single spectral image) only one eye of each patient was imaged.

The Declaration of Helsinki [12] was followed for all parts of the study involving human subjects. The imaging of the patients was conducted by the Department of Ophthalmology in Kuopio University Hospital (Kuopio, Finland) with the permission of the research ethics committee of University Hospital District of Northern Savo. Written consent was received from all patients prior to the study.

3.3.2 Imaging procedure

The process of acquiring images for the DiaRetDB2 database is described in [55]. A spectral fundus camera capable of acquiring up to 30 wavelength channels, with central wavelengths from 400 nm to 700 nm, was used to capture the spectral fundus images of the voluntary subjects. The details of the acquisition system are discussed in Section 2.4.1.

Oftan Tropicamid (Santen Oy, Finland) eye drops were used to dilate the pupils of the subjects before image acquisition. The head of the subject was supported by a forehead and a chin rest to provide stability and to make the imaging process more comfortable for the subject. A fixation target was used to reduce involuntary eye movements and to align the eye of the subject so that optic nerve head was in the image periphery.

The initial camera focus was determined using a bandpass filter with central wavelength at 550 nm, where the retinal blood vessels are prominent. If necessary, manual adjustments to the initial focus were made during the acquisition with an individual filter. When the subject deemed it necessary, a moment was taken between the acquisition of channels for the subject to rest their eyes before continuing with the imaging.

Images were acquired using all 30 bandpass filters where possible. In practice, however, the low transmittance for the shorter wavelengths in cornea and the lens (see e.g., [28]) resulted in very little information to be captured at the shortest wavelength channels. Channels where no significant information was observed were discarded.

3.3.3 Gold standard annotations

Two ophthalmologists were asked to annotate all abnormalities found in the images, and certain retinal structures such as the optic disk and macula. The ophthalmologists were asked not to consult others during the annotation work to guarantee independent set of annotations. The annotators were given predetermined classes for a number of more common lesion types (e.g., haemorrhage and exudate), and general classes with

mandatory description fields for other pathological and non-pathological abnormalities. Each annotation includes a confidence level on the scale uncertain - somewhat certain - certain.

Depending on the type of the ROI, the ROI was given either as a single point or a closed region drawn freehand. For the freehand regions, the ophthalmologists were instructed to draw a boundary enclosing all the ROI of the same type in that area without intersecting the boundary of the ROI. To provide representative spectra of the ROI, the annotators were asked to delineate a representative region where all the region pixels are inside the ROI in question.

The motivation for requesting the annotators to include the representative regions comes from the requirements of automatic segmentation. Approximate delineation of a lesion or a group of lesions is usually sufficient for a human observer. For an automatic segmentation or classification method, however, coarse annotations may result in the training examples of pathologies to contain a significant portion of background pixels. The loss of representativeness in the pathological training examples can significantly reduce the performance of the automatic method trained or developed with that data. Pixel-wise accurate manual annotations are generally not feasible for more than a couple of images due to the amount of work required.

The coarse manual annotations can be post-processed to produce annotations that follow the lesion boundaries more accurately. While the resulting post-processed annotations can no longer be considered as expert annotations or gold standard, there is a significant potential for the improvement of the quality of training data for automatic segmentation or classification. The refinement of manual annotations is discussed in Section 3.5.

All annotations by an ophthalmologist for a specific image are stored in a text file in an XML-style format:

```
<annotationBegin>
<id>77<id>
<title>New Hard exudate annotation<title>
<visible>true<visible>
<notes><notes>
<certainty>100<certainty>
<color>125,125,0<color>
<isChild>false<isChild>
<children>78,79,80<children>
<type>Hard exudate<type>
<boundaryType>general<boundaryType>
<coordinates>728.3,672.3,...<coordinates>
<annotationEnd>,
```

where *id* is a unique identifier, *visible* determines whether the annotation is visible in the view of the annotation tool (facilitates the annotation of closely grouped ROI of different type), *certainty* marks the confidence level, *color* determines the boundary color (in RGB), *isChild* determines whether the annotation is an enclosing boundary (parent) or a representative region (child), *children* contains the identifiers of the representative

regions associated with the annotation, *boundaryType* (single point or freehand region) instructs the annotation tool how the region is to be drawn, and *coordinates* contain the xy-coordinate pairs that define the region boundary.

In addition to the spatial annotations, the ophthalmologists were asked to give a general diagnosis for each image (e.g., healthy, mild non-proliferative diabetic retinopathy, mild glaucoma, etc.). The diagnostic information is stored in the beginning of each annotation text file in a similar format:

```
<diagnosis>
<isHealthy>false<isHealthy>
<certainty>100<certainty>
<diabetesClassification>3<diabetesClassification>
<AMDClassification>0<AMDClassification>
<glaucomaClassification>0<glaucomaClassification>
<maculopathyClassification>1<maculopathyClassification>
<notes><notes>
<diagnosis>,
```

where *certainty* is the confidence level of the diagnosis, *diabetesClassification*, *AMDClassification*, *glaucomaClassification* and *maculopathyClassification* are the diagnosed state of diabetic retinopathy, AMD, glaucoma and diabetic maculopathy, respectively.

3.3.4 Summary

Publicly available benchmark databases benefit the development of computer vision methods. Medical image databases typically concentrate on one organ imaged in a specific modality with the number of images often limited due to the availability of subjects, the cost of acquisition and the ethical limitations.

The steps in the acquisition of images and gold standards of lesion locations for composing the spectral fundus image database DiaRetDB2 were presented. The spectral information coupled with gold standards for multiple lesion types provide a good basis for method development in automatic detection and classification of retinal lesions.

3.4 Effect of ground truth inaccuracy on lesion classification

A number of different approaches for the automatic and semi-automatic image segmentation and detection of lesions in the retinal images have been proposed, e.g., [59, 177, 224]. A large number of the automatic methods are supervised, i.e., detect abnormalities using a pretrained model to classify regions of interest based on image features such as colour [195], texture [4], or a combination of several features [74, 163].

One of the main problems with supervised methods is the need for representative data of the regions of interest, preferably spatially accurate ground truth, to train the classifier. A common solution for composing the spatial ground truth of the abnormalities in retinal images is to ask medical experts to perform manual or semi-automatic segmentation. However, manual annotation of a set of images is time-consuming and tedious, and the

time of medical professionals qualified to make the annotations is often limited. Consequently, while some retinal image databases contain approximate manual segmentations or delineations of the lesions, spatially accurate ground truth is often unavailable.

When producing manual annotations, approximate delineations result in an inaccurate ground truth. Inaccuracies in the ground truth may have a significant impact on the performance of a supervised method as the classifier is trained with examples that are not fully representative of the lesion class. Especially in the case where multiple small lesions of the same type, such as exudates, are spread over a region, the delineated area may consist predominantly of the background (an example is shown in Figure 3.1).

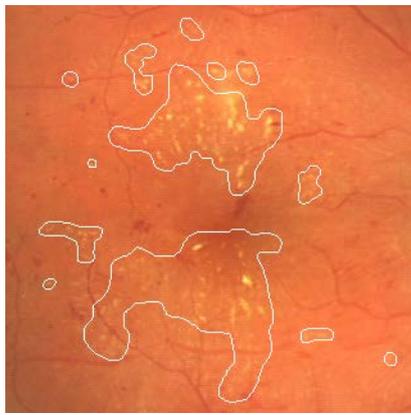


Figure 3.1: An example of spatially inaccurate ground truth of exudate locations. [130]

Different image features are affected differently by the increase of non-representative points in the training data, a factor that should influence the generation of training data and the selection of features for a supervised segmentation method. This section presents a quantitative evaluation of the sensitivity of different image features (including colour, texture and edge features and higher level features) to inaccuracy in the ground truth [130]. The feature sensitivity is evaluated as the classification error of pixel-wise segmentation (i.e., if exudate or background pixels are correctly classified as such) of exudates in retinal images. Separate instances of a classifier are trained using each feature individually to assess the sensitivity of the feature to the inaccuracy of training data.

3.4.1 Related work

Sánchez et al. [196] propose the use of active learning to reduce the workload of the expert annotating medical images. Starting from a small initial training set, a query function selects the sample from an unlabelled set for manual labelling so that it most likely improves the classifier performance.

In [210], Sommer et al. present an interactive toolkit for medical image segmentation. The toolkit provides semi-automatic segmentation by a Random Forest classifier that is fine-tuned by interactive refinement of the segmentation results.

Noronha and Nayak [165] review approaches for automatic detection of diabetic retinopathy. A quantitative evaluation of method performance is presented, with an emphasis on methods for automatic segmentation of exudates.

To the knowledge of the author, no quantitative evaluation of the effect of ground truth inaccuracy on classifier performance exists in relation to fundus images. Furthermore, the literature on supervised classification in the context of retinal images does not, generally, discuss the quality of the ground truth. However, there is awareness of the difficulty of obtaining representative ground truth in medical imaging context and its impact on algorithm validation. The creation of common databases and performance measures for the purpose of standardised performance evaluation has been proposed [103, 224].

3.4.2 Methods

LOCAL STATISTICAL FEATURES

Statistical measures can be used as simple descriptors of local neighbourhoods. While simple statistical features are unlikely to provide enough distinction between exudates and healthy tissue individually, they can be used in conjunction with other features to improve the overall performance of a classifier. The features selected for the evaluation are intensity range, entropy and intensity standard deviation (σ). The features are calculated using the following formulae:

$$\text{Range} = I_{max} - I_{min} \quad (3.1)$$

$$\text{Entropy} = - \sum_{i=1}^L p(z_i) \log_2 p(z_i) \quad (3.2)$$

$$\text{Standard deviation} = \sqrt{\frac{1}{N} \sum_{j=1}^N (x_j - \mu)^2}, \quad (3.3)$$

where I_{min} and I_{max} are the maximum and minimum values, L is the number of different graylevels, $p(z_i)$ is the probability of i :th graylevel value, N is the pixel count, x_j graylevel value of the j :th pixel, and μ is the mean of the neighbourhood.

COLOUR FEATURES

In retinal RGB images, colour is a simple feature with good discriminative properties for a number of retinal structures. The RGB triplets of pixel values can be directly used as feature vectors for the classifier.

The CIE Lab (Lab) colour space is intended to simulate the colour differentiation of the human vision system (i.e., points close together in Lab correspond to visually similar colours). As the transformation between RGB and Lab colour spaces is nonlinear, Lab features may produce different classification results than RGB features.

The transformation from RGB to Lab using CIE D65 illuminant is defined as [101]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.4)$$

$$\hat{X} = \frac{X}{0.950456}, \quad \hat{Z} = \frac{Z}{1.088754} \quad (3.5)$$

$$L^* = \begin{cases} 116(Y/Y_n)^{\frac{1}{3}} - 16, & \frac{Y}{Y_n} \geq 0.008856 \\ 903(Y/Y_n), & \frac{Y}{Y_n} < 0.008856 \end{cases} \quad (3.6)$$

$$a^* = 500 \left[(\hat{X}/X_n)^{\frac{1}{3}} - (Y/Y_n)^{\frac{1}{3}} \right]$$

$$b^* = 200 \left[(Y/Y_n)^{\frac{1}{3}} - (\hat{Z}/Z_n)^{\frac{1}{3}} \right],$$

where X_n , Y_n and Z_n are the tristimulus values of the reference illuminant.

Due to the colour variation of the retina between individuals, the colour features have to be normalised before using them as training data. The approach to colour normalisation was to calculate a colour histogram H_i for each image in the training set and perform histogram equalisation for the set with the mean histogram \bar{H} as the target distribution.

TEXTURE FEATURES

Graylevel co-occurrence matrix (GLCM) describes local texture by calculating the frequency of the pixel intensity value in a specific spatial relationship (e.g., horizontal neighbour) with a pixel of a specific intensity value. GLCM can be seen as a histogram of intensities occurring in that relationship [86]. A number of features can be calculated from the GLCM, e.g., contrast, correlation, energy and homogeneity. As the contrast measure produced the best results for GLCM in the initial evaluation, it was selected as the GLCM feature. The feature is defined as

$$\text{Contrast} = \sum_{i,j} (i-j)^2 p(i,j) \quad (3.7)$$

where i and j are vertical and horizontal coordinates of the GLCM and $p(i,j)$ is the (i,j) :th entry of the normalised GLCM. Another approach to local texture representation are local binary patterns (LBPs) where the neighbourhood of each pixel is thresholded by the pixel intensity value, and the thresholded neighbourhood is summed after weighting the binary values in order with increasing powers of two. The original framework has been extended by, e.g., Ahonen et al. [5], who propose a rotation invariant feature calculated from Fourier transform of LBP histograms, local binary pattern histogram Fourier feature (LBPHF).

HIGHER-LEVEL FEATURES

The SURF [16] features are constructed by computing Haar wavelet responses from square-shape sub-regions around the keypoint. Rotation invariance is achieved by rotating the sampling region by the dominant orientation of the region gained from Gaussian weighted Haar wavelet responses within a sliding window inside a circular region around the keypoint.

Leutenegger et al. [137] present BRISK, a keypoint descriptor based on brightness comparisons from locations uniformly spaced on concentric circles around the keypoint. Prior

to sampling, the descriptor orientation is normalised by rotating the sampling pattern by the characteristic direction of the neighbourhood, to achieve rotation invariance.

The FREAK [8] describes keypoints by calculating a cascade of binary strings from comparisons of image intensities over a sampling pattern imitating the distribution of retinal ganglion cells. An orientation normalisation similar to BRISK is applied. In addition, a keypoint matching approach inspired by human vision system is presented.

CLASSIFIERS

In random forest (RF), a number of tree classifiers are generated from independent random samples from the training data. The class of input data is determined by the most popular class assigned by the independent tree classifiers [32].

RF classifiers have been used in segmentation and classification of retinal images (e.g., [3, 114, 190]). Due to its successful application in [210] and good performance in initial experiments, RF was selected for the formal experimental evaluation of the feature sensitivity.

3.4.3 Experiments and results

Due to the amount of work involved in the manual annotation of retinal images, the annotations of lesion locations available for DiaRetDB2 are not pixel-wise accurate. However, pixel-wise annotations are required for reasonable evaluation of the effect of the inaccuracy of ground truth on the feature performance.

Instead of using DiaRetDB2, the feature sensitivity was evaluated on a non-public eye fundus image database, Bristol retinal image data set (BristolDB), collected by the Bristol Eye Hospital, for which pixel-wise accurate ground truth of exudate locations is available. 147 images, of which 107 were used in the evaluation in this study (17 healthy patients and 90 with exudates), were captured using a 45 degree field-of-view Canon CR6-45 Non-Mydriatic fundus camera with a modified digital back-unit (Sony PowerHAD 3CCD colour video camera and Canon CR-TA). Pixel-wise accurate ground truth of exudate locations was manually annotated in the images by a consultant ophthalmologist. [173]

To evaluate the feature performance at different levels of ground truth inaccuracy, another set of manual annotations was gathered. Using the pixel accurate ground truth as a reference for exudate locations, approximate delineations enclosing the accurate ground truth were manually drawn to simulate expert annotations that are not spatially accurate. Synthetic ground truth with intermediate accuracy was generated by iteratively thresholding the distance transform image (i.e., image where each pixel value is the distance of that pixel to the original binary mask) of the accurate ground truth with increasing thresholds until expanded to the coarse annotations (pixels outside the coarse annotations were discarded at every iteration). An example is shown in Figure 3.2.

For each level of the “ground truth”, each feature was extracted from the pixels designated by the simulated ground truth mask, with the same number of features extracted from randomly selected background pixels within the usable image area.

Initial experiments were run on support vector machine (SVM), Gaussian mixture model (GMM) and random forest (RF) classifiers. However, of the studied classifiers only RF

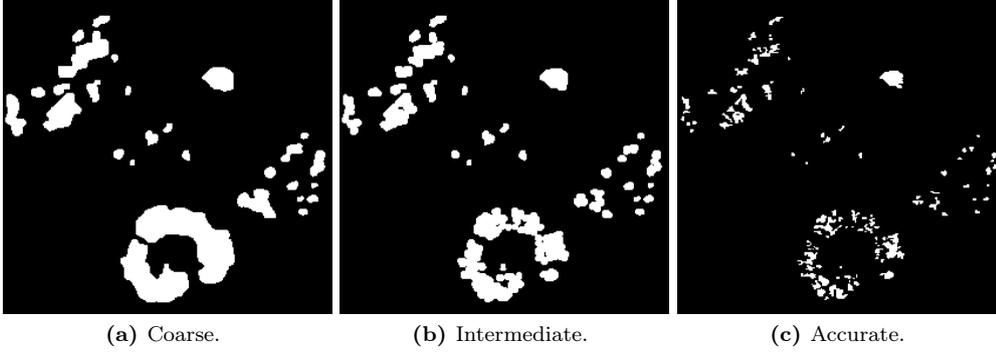


Figure 3.2: Example manual (a, c) and synthetic (b) annotations. [130]

produced reasonable results (as opposed to all pixels being classified in the same or seemingly random class) for all of the features.

A separate RF classifier was trained for each feature and for each level of simulated ground truth accuracy. The classifier performance was cross-validated by using randomly selected two thirds of the images in the database, with the remaining third reserved for testing the classifier performance. Ten cross-validation steps were performed for each classifier.

The classifiers were used to classify each pixel of each image in the remaining one third of the image set into exudate or background class. The classification results were evaluated against the (simulated) ground truth of the same level as used in training the classifier (the goal of a classifier can be seen to be the reproduction of classification defined by the training data on new input, even though more accurate ground truth was available).

To determine the classifier parameters (number of trees and number of samples for training), classifiers with different parameters were trained using each of the features under evaluation. The classifiers were run on the same data sets used for training to determine the learning rate with different parameters. The values selected were 3000 samples (for foreground and background, 6000 samples in total), approx. 5% of the training set, and 100 trees. As increasing the number of samples beyond 3000, $\sim 5\%$ of the training set, and the number of trees beyond 100 (see Figure 3.3) does not significantly improve the learning rate while notably increasing the time required for training, these values were selected for the experiments.

The two commonly utilised metrics that were used to quantify the classification performance, sensitivity (SN), specificity (SP), are defined as

$$\text{SN} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \text{SP} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3.8)$$

where TP is the number of true positives (exudate classified as exudate), TN true negatives (background as background), FP is false positives (background as exudate) and FN is false negatives (exudate as background).

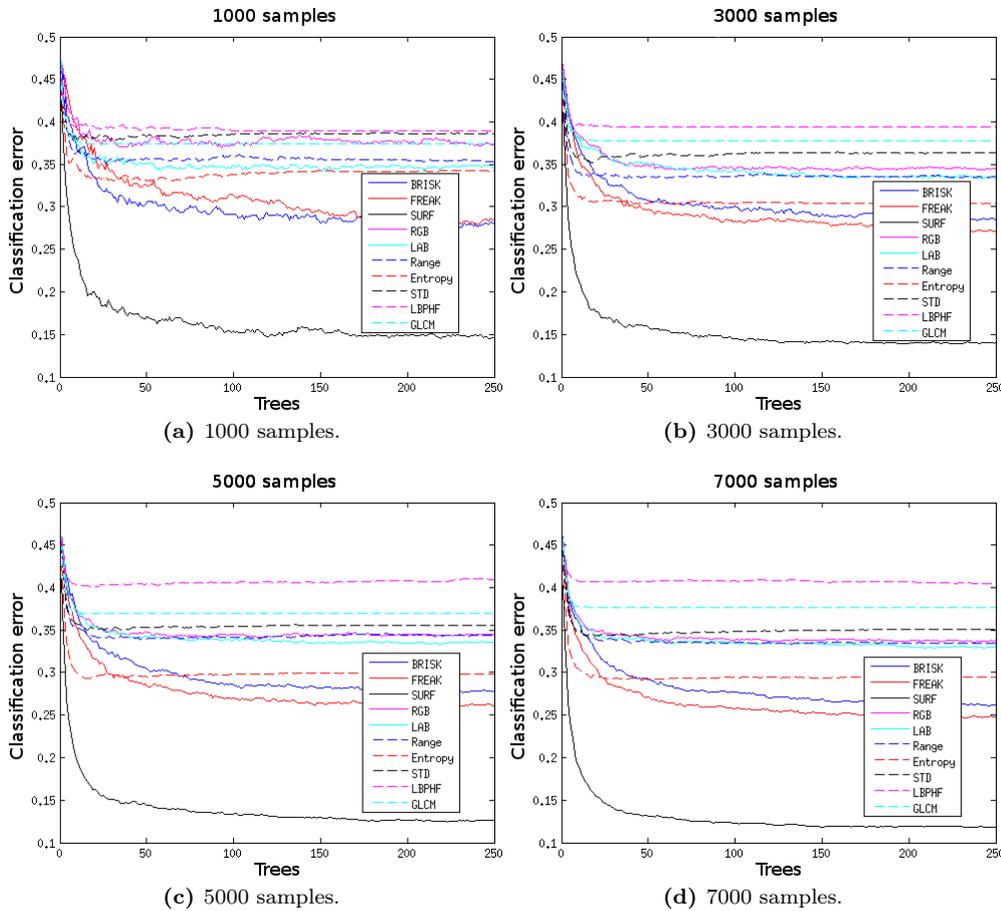


Figure 3.3: Classifier learning on coarse ground truth with different number of samples; binary Robust Invariant Scalable Keypoints (BRISK), fast retina keypoint (FREAK), speeded-up robust feature (SURF), RGB colour (RGB), Lab colour (Lab), intensity range, entropy, standard deviation (STD), local binary pattern histogram Fourier feature (LBP4F) and graylevel co-occurrence matrix (GLCM). [130]

The training data inaccuracy had a significant effect on the classification performance, resulting in a mean decrease of sensitivity and specificity by approx. 20% and 13%, respectively, between the coarsest and the accurate training data. SURF outperformed the other features both in classification performance and in robustness to ground truth inaccuracy by a clear margin.

On the accurate ground truth, RGB colour performed similar to FREAK and BRISK, but its performance degraded faster as the ground truth was expanded. Local statistic features (excluding entropy) showed similar performance to RGB colour, although less robust to higher levels of ground truth inaccuracy. Texture features showed high sensitivity to the ground truth inaccuracy, along with poor general performance. The results

are summarised in Figure 3.4.

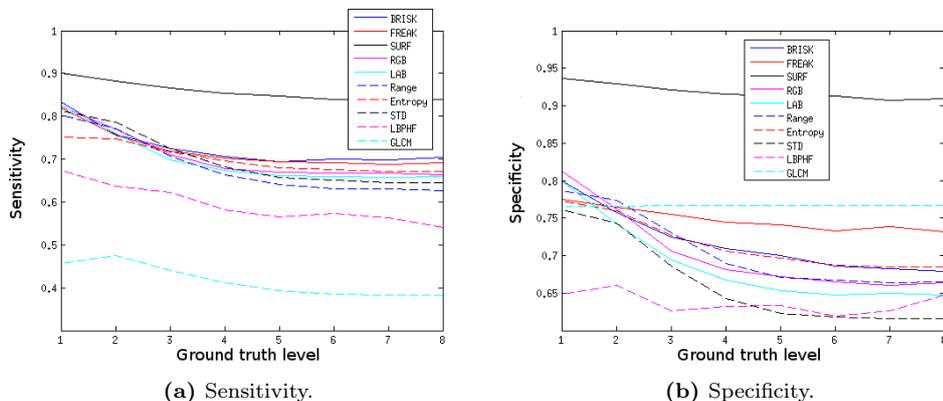


Figure 3.4: Feature sensitivity to ground truth inaccuracy evaluated against the simulated ground truth of the same level. The levels range from the highest (1) to the lowest (8) spatial accuracy. [130]

3.4.4 Discussion

In part, the result can be explained by the longer feature vector of SURF compared to the other features. However, the one-dimensional local entropy feature showed classification performance close to FREAK and BRISK while being less affected by the decreasing ground truth accuracy.

Only SURF achieved sufficient classification performance to be considered as the sole feature of a detector. While most of the features managed to detect most of the exudates, a large number of false positives resulted in worse SP. According to the results, in addition to the higher-level features, colour and local entropy are good candidate features for a supervised exudate detector.

3.4.5 Summary

Supervised methods require representative data of the regions of interest, preferably spatially accurate ground truth, to train the classifier. Manual annotation of a set of images is time-consuming and tedious, and the time of medical professionals qualified to make the annotations is often limited, making spatially accurate ground truth often unavailable.

The sensitivity of features to ground truth inaccuracy was evaluated on a set of 107 images with pixel-wise accurate ground truth of exudate locations manually annotated by a consultant ophthalmologist. Another set of manual annotations was gathered using the pixel accurate ground truth as a reference. Synthetic ground truth with intermediate accuracy was generated by iteratively thresholding the distance transform image of the accurate ground truth with increasing thresholds.

A separate RF classifier was trained for a set of features on each level of simulated ground truth accuracy. The training data inaccuracy had a significant effect on the classification performance, resulting in a mean decrease of sensitivity and specificity by approx. 20% and 13%, respectively, between the coarsest and the accurate training data. SURF outperformed the other features both in classification performance and in robustness to ground truth inaccuracy by a clear margin.

3.5 Annotation refinement

As discussed in Section 3.4, non-representative ground truth of lesion locations can have a significant effect on the performance of automatic segmentation methods. This motivates automatic spatial refinement of manual segmentations to improve the representativeness of the coarse manual segmentation data either at the collection stage (semi-automatic segmentation), or as a post-processing step.

The automatic refinement of manual annotations can be considered as a restricted segmentation problem. The purpose of an automatic refinement method is to provide pixel-wise segmentation while requiring only coarse initial segmentation or delineation of the ROIs. In the case of medical image processing, the regions are typically anatomical structures or lesions related to a medical condition.

In this Section, a method utilising the colour feature and maximally stable extremal regions of lesion likelihoods is presented for the spatial refinement of manual segmentations by experts [130]. The proposed method is compared to a set of methods by using the non-public BristolDB (described in Section 3.4.3) containing spatially accurate segmentations of exudates. Exudates being yellowish lesions typically with a low contrast to the background and without clear borders make the segmentation task a challenging one.

3.5.1 Related work

Various approaches have been proposed for segmenting regions indicative of pathologies and other ROIs in retinal images (e.g., [64, 224]). A short review of different approaches for automatic and semi-automatic segmentation, with the emphasis on retinal image segmentation is presented.

Sinthanayothin et al. [206] use a recursive region-growing segmentation (RRGS) and a neural network (NN) to separate blood vessels from red lesions. RRGS is used jointly with the moat operator increasing the contrast of the edges of the lesions, followed by thresholding.

Hipwel et al. [94] use top-hat transformation on a shade corrected image to remove blood vessels. A Gaussian matched filter is applied to obtain candidate haemorrhage and microaneurysm (HMA) areas used in subsequent classification.

Niemeijer et al. [160] suggest a hybrid method for the detection of red lesions. Candidate regions were obtained by application of a top-hat based method and a pixel-based classification. The true lesion regions were detected with k-nearest neighbour classification.

Ravishankar et al. [185] suppressed blood vessels using morphological operations. The detection of exudate boundaries was obtained from the difference between images dilated at two separate scales. Dynamic thresholding and application of smoothing splines

produced closed contours which, after morphological filling, represent the candidate exudates. Finally, the true exudates were determined using a linear classifier based on the brightness and edge properties of the exudates.

Jafaar et al. [102] segment images with logical intersection of coarse segmentation based on local variation, and the results of adaptive thresholding based on a combination of pure splitting and histogram-based thresholding. The coarse segmentation assumes that hard exudates have clear boundaries, and is based on the automatic thresholding of a standard deviation image using Otsu's method, followed by a classification of features such as major axis length, minor axis length, area and solidity.

Walter et al. [237] proposed a segmentation technique based on morphological operations. During the first step, blood vessels were eliminated by closing. The regions with big local variation were considered as candidate regions after the hole filling from their borders. Afterwards, dilation and optic disc removal were performed. To obtain the final result, the candidate regions were set to zero and reconstructed by successive geodesic dilation. The true exudate areas were produced by thresholding the difference between the original image and the reconstructed one.

Many authors have utilised fuzzy *c*-means clustering (FCM) and its modifications for coarse and final image segmentation. FCM [24] minimises the distances between the samples and iteratively updated cluster centres, with each sample belonging to each cluster with degree of membership w . At each iteration the cluster centres are updated based on sample values weighted by their degree of membership for the given cluster.

Kande et al. [112] apply spatially weighted fuzzy *c*-means clustering (SWFCM) to the graylevel histogram of the image after removing the optic disc. Maximum membership procedure produces the mask corresponding to the true HE regions. Chutatape [252] obtained candidate lesion regions using Improved FCM in Luv colour space after local contrast enhancement. Classification is performed by SVM utilising region edge strength, colour difference between region and its surrounding area, and region size.

Another popular approach of image segmentation is based on the features extracted using different filters. For instance, Osareh et al. [174] utilised a bank of Gabor filters to segment candidate regions obtained by FCM clustering. To identify the best subset of features, a genetic optimisation algorithm was employed. Multilayer perceptron neural network with three layers was used to classify the extracted regions.

Region growing is another clustering technique applied in the detection of hard and soft exudates. Sinthanayothin et al. [206] used a Recursive Region Growing Segmentation in order to find similar pixels in detection of bright lesions. Thresholding was utilised to produce an exudate mask after all pixels have merged into regions.

Reza et al. [188] applied watershed transform for bright lesion detection. They used average filtering that blends the small objects with low intensity variations into the background, and contrast stretching transformation as preprocessing procedures. Afterwards, bright parts of the image were isolated by thresholding. Morphological opening of resulting binary image and the result of extended maxima operator applied to the initial image gave regions with high intensity values, that were used for obtaining markers in subsequent watershed transform.

A significant body of work exists in automatic detection of regions of interest in retinal images based on classifying pixels. In [173], Osareh et al. perform histogram specification and local contrast enhancement followed by evaluation of Gaussian smoothed histograms of each colour band of the image in order to produce initial classification of retinal exudates. FCM clustering is used to assign remaining pixels to the previously determined clusters followed by neural network object classification.

Similar segmentation scheme was realised in [172]. The major difference is observed in the preprocessing step where colour normalisation instead of histogram specification was implemented. Osareh et al. [174] segment candidate regions from FCM using a bank of Gabor filters and identify the best feature subset with genetic optimisation. The extracted regions are classified by a multilayer perceptron NN. Xu and Luo [248] propose a combination of stationary wavelet transform and graylevel co-occurrence matrix as a feature for pixel-level classification by SVM. However, classifiers, such as NN and SVM, cannot be utilised in refinement of manual segmentations as no training data is available as such.

In addition to medical image processing, various segmentation approaches have been applied in numerous other fields. The following short review of segmentation approaches outside the medical field is limited to publications that are relevant to the methods evaluated in Section 3.5.3.

Kakumanu et al. [110] present a review of methods for modelling and detection of skin-colour. The review includes the use of different colour spaces, modelling with Gaussians and Gaussian mixtures, and classification with naïve Bayes classifier.

Donoser and Bischof [223] use colour likelihood maps with modified maximally stable extremal regions (MSER) tracking for robustly tracking a hand through a sequence of images. The likelihood map is based on Kullback-Leibler distance between Gaussian distributions fit to the neighbourhood of a pixel and the background distribution.

Guo et al. [80] propose pixel-based hierarchical features for weak classifiers for hand detection. Adaboost is used to form a strong classifier of the combination of the proposed weak classifiers for skin-colour segmentation.

Wang et al. [240] use graph cuts to segment foreground objects in video frames to obtain more precise colour histogram of the object for mean shift tracking. Graph-cuts (Graph) based segmentation methods [31] view labelled image areas as nodes of a connected graph, and rearrange the labels corresponding to a minimum cut on the graph.

Ning et al. [164] present an interactive image segmentation method. Unlabelled regions are merged with initial regions from mean shift segmentation based on similarity (colour histogram distance) until all regions have been labelled.

GrowCut algorithm (GrowCut) by Vezhnevets and Konouchine [235] is a cellular automaton based semi-automatic segmentation method. The image is labelled by interaction of cellular automata containing different labels, mimicking the growth and competition of bacteria.

Many of the proposed semi-automatic segmentation methods can be considered as suitable for the refinement of manual segmentations. A selection of approaches are evaluated in Section 3.5.3.

3.5.2 Region refinement through stable probability regions

The DiaRetDB1 [116, 117] used as the case study (an example is shown in Figure 3.5) includes manual segmentations for the retinopathy lesions together with representative points inside each segment. To estimate the parameters of statistical models representing sets of pixels or image features, the values of a single pixel closest to that point (representative pixel) are not sufficient, however. Thus, small (area $\sim 5 - 10$ pixels) manually drawn representative regions were used.



Figure 3.5: An example image from DiaRetDB1 [117].

The proposed method for the refinement of manual segmentations is based on stable regions of a naïve Bayesian posterior probability map of lesion presence following the approach in [223]. Empirical probability density functions (PDFs) are fit to the colour values of the background and lesions by kernel smoothing density estimate using a Gaussian kernel. The data samples representing the lesions arise from manually made representative regions.

A probability map of each pixel belonging to the lesion is calculated as $p_{\text{lesion}}(1 - p_{\text{bg}})$ where p is the Bayesian posterior probability of the pixel colour value in the lesion and the background model, respectively. The lesion mask is produced based on the MSER [147] of the resulting probability map.

MSER, can be seen as a systematic search for regions where local binarisation stays stable over a wide range of thresholds. Thus, the selection of an explicit probability threshold can be avoided and small spurious high-probability regions inconsistent with the pixel neighbourhood removed. Algorithm 3.1 presents the details of the proposed method.

3.5.3 Experiments and results

As explained in Section 3.4.3, as only coarse annotations of lesion locations are available for DiaRetDB1, another data set is used for quantitative evaluation of the segmentation performance. The performance of the segmentation methods was evaluated on the

Algorithm 3.1 Stable probability regions based region refinement: kernel density estimate (KDE), maximally stable extremal regions (MSER).

Input: RGB image I , coarse segment set s , representative region set R

Output: Refined segmentation binary mask M_{ref}

$M_{\text{ref}} \leftarrow 0$

for all segment s_i **do**

 Fit KDE G_{bg} to boundary colour values of s_i

 Fit KDE G_{lesion} to colour values of region R_i

for all pixel colour value v_j of s_i **do**

$\text{pmap}_j \leftarrow p(v_j|G_{\text{lesion}})(1 - p(v_j|G_{\text{bg}}))$

end for

$M_{\text{ref}}(\text{MSER}(\text{pmap})) \leftarrow 1$

end for

BristolDB, a non-public data set collected by Bristol Eye Hospital, consisting of 107 RGB-images of subjects with diabetic retinopathy and healthy control subjects. Exudates, when present, were marked with pixel-level accuracy by a consultant ophthalmologist.

The test set with coarse segmentations was produced by collecting coarse manual segmentations of exudate regions and representative regions (foreground for the segmentation methods) for each segment. Possible representative region pixels outside the spatially accurate ground truth were removed.

To implement the method comparison, the following methods were used: binary decision tree (BDT) based classification (with Gini’s diversity index optimisation criterion) [52], FCM clustering [24, 40], segmentation based on Graph-cuts [31], maximal similarity region merging (MSRM) [164], GrowCut [235], and the automatic thresholding by Otsu [175]. Publicly available third party implementations were used for methods for which the original implementation by the author was not available.

The performance of the methods was evaluated by using the Dice similarity criterion between the ground truth and the exudate mask refined by a given method, together with the standard detection performance metrics. Dice coefficient is defined as [14]

$$\text{DSC} = \frac{2|A \cap B|}{|A| + |B|}, \quad (3.9)$$

where A and B are region masks, and $|\cdot|$ is the number of pixels in the region.

In addition, the experiments on the effect of annotation inaccuracy on classifier performance were performed using the refined exudate masks as training data. The experiments related to the classifier performance follow the framework described in Section 3.4.3, and the performance is measured in sensitivity (SN) and positive predictive value (PPV), defined as

$$\text{SN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3.10)$$

and

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (3.11)$$

where TP is the number of true positives (exudate classified as exudate), TN true negatives (background as background), FP is false positives (background as exudate) and FN is false negatives (exudate as background). The method evaluation results are shown in Table 3.1.

Otsu thresholding performed well when there was a clear and uniform intensity difference between the lesion and the background. However, its very large variation over the whole image set made the method unreliable for segmentation refinement. While Otsu achieved the highest SN, it produced a very high number of false positive detections. Poor performance of MSRM was mostly due to the inability of mean shift segmentation to provide reasonable regions in the initial segmentation. BDT segmentation, GrowCut and the proposed method outperformed the other evaluated methods by a large margin.

Table 3.1: Performance metrics of the methods binary decision tree (BDT), fuzzy c-means clustering (FCM), GrowCut algorithm (GrowCut), maximal similarity region merging (MSRM), Bayesian probability regions (NB) and Otsu thresholding (TH5). The metrics are as follows: Dice similarity coefficient, sensitivity (SN), positive predictive value (PPV). [130]

	Dice		SN		PPV	
	Median	Std	Median	STD	Median	Std
NB	0.6634	0.1753	0.6213	0.2318	0.7821	0.1357
BDT	0.6577	0.1444	0.6617	0.1915	0.7188	0.1738
FCM	0.4975	0.1897	0.6674	0.1582	0.4413	0.2823
GrowCut	0.4622	0.1868	0.6101	0.1961	0.4251	0.2184
OTSU	0.3576	0.2027	0.8610	0.1648	0.2301	0.2090
GraphCut	0.3409	0.2425	0.5957	0.2442	0.2287	0.2782
MSRM	0.1242	0.1411	0.1636	0.2533	0.1030	0.2206

A boxplot of the Dice coefficients achieved by the methods is shown in Fig. 3.6. The results show that BDT and NB outperform the other methods with a statistically significant margin. The proposed method has a slightly larger variation in the results and some challenging images (i.e., images containing a number of small low-contrast lesions) cause outliers for both methods. The differences of the best methods are elaborated more in Fig. 3.7 where the resulting Dice coefficients are sorted in increasing order. The differences emphasise that BDT produces fewer bad segmentations than NB, but after the set of weakest segmentations (one third of the image set in size), NB consistently produces higher Dice coefficient values.

The classifiers trained using the exudate masks refined with BDT and NB showed performance close to the classifiers trained with the spatially accurate ground truth. The discriminativeness of NB for the BristolDB data set resulted in some decrease in the sensitivity of the classifier (while still producing results comparable to using the accurate training data), but enabled NB to outperform the other refinement methods when measured in classifier specificity. The mean (over 10 cross-validation test sets) sensitivity and specificity are presented in Tables 3.2 and 3.3, respectively.

In preliminary tests on annotations in DiaRetDB1 [116, 117], the performance of the proposed method was promising (as can be seen in Figure 3.8). The initial results suggest

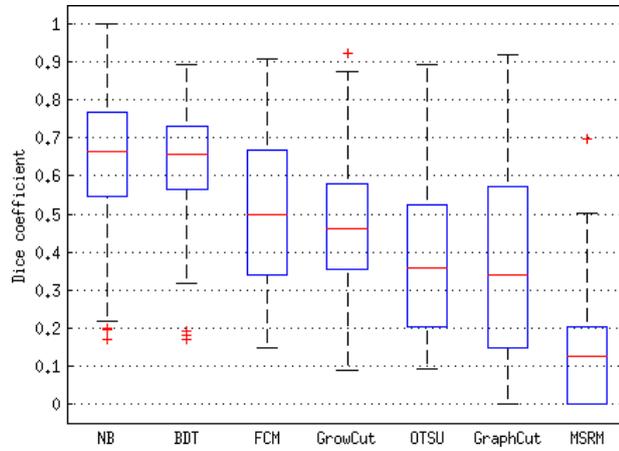


Figure 3.6: Similarity of refined regions and ground truth. The methods are as follows: binary decision tree (BDT), fuzzy c-means clustering (FCM), graphcuts (Graph), GrowCut algorithm (GrowCut), maximal similarity region merging (MSRM), Bayesian probability regions (NB), Otsu thresholding (TH5). The red lines mark the medians, the boxes represent the 25th and 75th percentiles, the whiskers extend to inlier extrema, and the individual markers represent the outliers. [130]

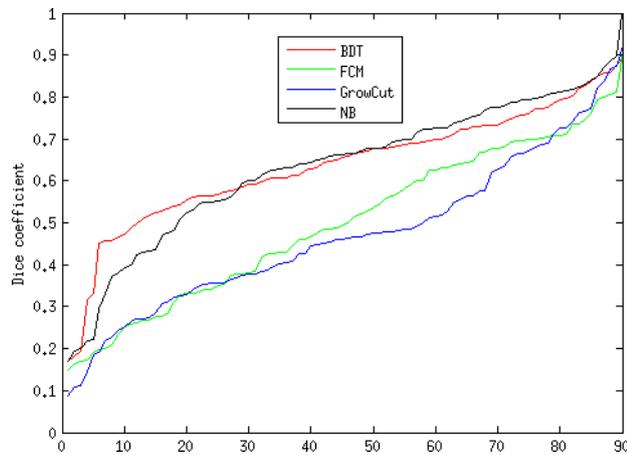


Figure 3.7: Dice coefficients of the best performing methods, binary decision tree (BDT), fuzzy c-means clustering (FCM), GrowCut algorithm (GrowCut) and Bayesian probability regions (NB), sorted in ascending order. [130]

that, for lesions for which colour is a good discriminating feature such as haemorrhages and exudates, the proposed method is able to considerably improve the coarse manual data.

Table 3.2: Classification result sensitivity using refined annotations as training data. binary decision tree (BDT), fuzzy c-means clustering (FCM), graph-cuts (Graph), GrowCut algorithm (GrowCut), maximal similarity region merging (MSRM), Bayesian probability regions (NB), Otsu thresholding (TH5).

	BRISK	FREAK	SURF	RGB	LAB	Range	Entropy	STD	LBPBF	GLCM
BDT	0.8626	0.8294	0.9034	0.8003	0.8097	0.7875	0.7586	0.7883	0.6504	0.4991
FCM	0.8757	0.8469	0.9325	0.7873	0.771	0.7392	0.7203	0.7395	0.7705	0.4988
GraphCut	0.822	0.79	0.9098	0.7448	0.7246	0.7348	0.7293	0.7345	0.6804	0.4991
GrowCut	0.786	0.7359	0.8742	0.8019	0.7858	0.7628	0.7554	0.7715	0.684	0.4991
MSRM	0.7885	0.6897	0.9046	0.7261	0.6881	0.7225	0.7253	0.7287	0.7023	0.4991
NB	0.8098	0.7876	0.87	0.7694	0.7778	0.7768	0.7416	0.7757	0.6605	0.4991
TH5	0.8816	0.8443	0.9319	0.7866	0.764	0.7145	0.7048	0.7293	0.6612	0.4967

Table 3.3: Classification result specificity using refined annotations as training data. binary decision tree (BDT), fuzzy c-means clustering (FCM), graph-cuts (Graph), GrowCut algorithm (GrowCut), maximal similarity region merging (MSRM), Bayesian probability regions (NB), Otsu thresholding (TH5).

	BRISK	FREAK	SURF	RGB	LAB	Range	Entropy	STD	LBPBF	GLCM
BDT	0.7428	0.7587	0.9202	0.815	0.8211	0.7298	0.7202	0.685	0.6629	0.7639
FCM	0.6866	0.6837	0.8395	0.6979	0.6991	0.5894	0.5657	0.5517	0.528	0.7639
GraphCut	0.6007	0.6215	0.8236	0.6375	0.6328	0.5742	0.5671	0.5407	0.6342	0.7639
GrowCut	0.6442	0.6857	0.8815	0.6857	0.6817	0.6084	0.6121	0.5672	0.6272	0.7639
MSRM	0.5797	0.6156	0.8269	0.6062	0.6293	0.5633	0.5892	0.5492	0.5987	0.7639
NB	0.8091	0.7941	0.9496	0.8575	0.8526	0.7946	0.7863	0.7727	0.6463	0.7639
TH5	0.6049	0.6284	0.8228	0.6206	0.6323	0.5643	0.5525	0.5324	0.6413	0.7641

3.5.4 Discussion

The proposed method performed well with most images of the set. The majority of the differences between the ground truth and refined mask were false negatives. Therefore, the refined segments can be considered to provide more robust training data also on images with low contrast between the lesions and the background.

While the median results of BDT and the proposed method were very close, NB repeatedly provided higher Dice coefficients. However, the larger variation in the results decreased the median Dice value for the method. A plausible reason for larger Dice variation with NB arises from the fact that for some larger segments, the data contains representative regions of only a few of many small lesions, resulting in Gaussian mixtures too limited to represent all the exudates within the segment. Nevertheless, the results show that the proposed method has good overall performance in refining the synthetic inaccurate annotations for the BristolDB retinal image database.

3.5.5 Summary

Supervised methods require representative data of the regions of interest to train the classifier. A common solution to composing the spatial ground truth of the abnormalities in retinal images is to ask medical experts to perform manual or semi-automatic segmentation. When producing manual annotations, approximate delineations result in inaccurate ground truth. The training data inaccuracy was found to have a significant

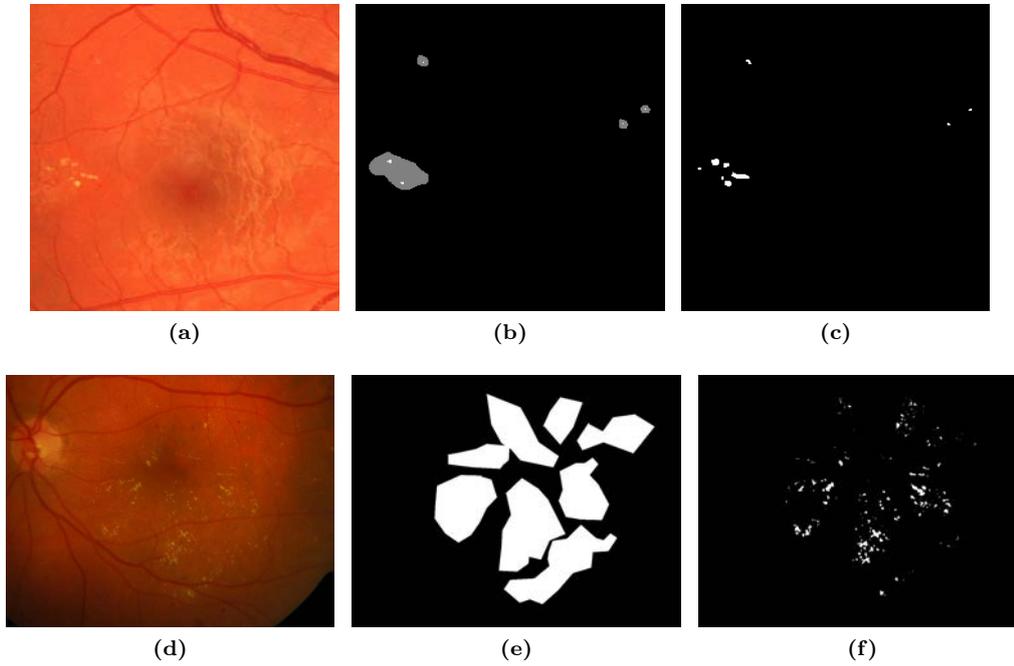


Figure 3.8: Examples of retinal images, manual segmentations and the refinement task. Top row from left to right: RGB image, coarse segmentation of exudates (white pixels mark the representative regions) and spatially accurate ground truth. Bottom row (refinement example by using DiaRetDB1 [117]): RGB image, coarse segmentation of exudates and the refined segmentation result. [130]

effect on the classification performance, resulting in a mean decrease of sensitivity and specificity by approx. 20% and 13%, respectively, between the coarsest and the accurate training data.

A method for refining coarse manual annotations based on stable Gaussian probability regions, NB, was presented. The performance of several segmentation methods in refining coarse annotation was evaluated by applying the methods on a data set where both coarse and accurate ground truth is available. Measured in Dice coefficient, BDT segmentation, GrowCut and the proposed method outperformed the other evaluated methods by a large margin. Classifiers trained using the annotations refined with BDT and NB showed performance close to the classifiers trained with the spatially accurate annotations.

Medical applications of spectral fundus data

4.1 Introduction

Automatic diagnosis of diabetic retinopathy (DR) is typically based on the detection of lesions in the eye fundus caused by the prolonged high blood glucose levels. The earliest signs of DR visible in traditional retinal images are microaneurysms. They are caused by ballooning of capillaries, and appear as small red dots in the retinal images. Larger red lesions with irregular outlines are caused by ruptured small blood vessels in the deeper retinal layers, called blot haemorrhages. The weakening of the capillary walls can result in excess leakage of plasma, manifesting in yellowish-white lesions with well defined borders, called exudates. Furthermore, the increased vascular permeability causes swelling and disruption of the macular structures, which is known as diabetic macular edema. As the delivery of oxygen and nutrients is disrupted due to the failure of and damage to capillaries, the resulting swelling of the retinal nerve fibres causes cotton wool spots, pale areas with poorly defined edges, to appear in the retinal images. [219]

The lack of oxygen and nutrients will stimulate the growth of new blood vessels in an attempt to circumvent the disrupted circulation. The increased growth rate of vessels can cause the dilation of small blood vessels (intra-retinal microvascular abnormalities (IRMA)), and the formation of new vessels. The newly formed vessels are generally very fragile and have a high risk of haemorrhaging, possibly causing a sudden and drastic deterioration of vision. [219] The neovascularisation process is known as proliferative diabetic retinopathy, that is in most cases a result of type 1 diabetes, causing increased health care costs due to the necessity of increased follow-ups, laser treatments and vitreoretinal surgeries, and visual impairment. As IRMA and neovascularisation appear as irregularities of the retinal vasculature, their automatic detection is generally more difficult than that of some of the other retinal lesions. Consequently, they are not as well represented in the (computerised) medical image analysis literature.

Another pathology causing visible changes to the retina is age-related macular degeneration (AMD). As the retina ages, extracellular material may start to accumulate, forming local deposits of the material. The deposits, called drusen, appear as yellowish or grey

spots of various sizes and shapes, and are the earliest signs of AMD. In about 10-15% of eyes suffering from AMD the so called dry AMD progresses to development of neovascularisation and, thus, to an exudative form of AMD commonly causing visual impairment.

As the lesions can be identified from retinal images and are strong evidence of a pathology, the automatic diagnosis of these diseases by the segmentation or classification of retinal images has been studied. When a sufficient number of images with lesions manually annotated by an ophthalmologist is available as training data, supervised classification can be used to classify image pixels to representations of healthy or pathological tissue. Supervised classification relies on training examples of healthy and pathological data to determine rules for good separation of the training examples. The rules are then used to determine classes for new input data.

Pixel-wise supervised classification of retinal images has been used for the segmentation of lesions related to DR (e.g., [161, 162, 211, 238]). Supervised classification has been applied to the detection of drusen [148] but, due to the variability in the appearance of drusen, has been less popular.

Although the reflectance spectra of a fundus image are results of a complicated set of photon interactions with layered tissue, traditional fundus images represent a single spectrum using only up to three values (red, green and blue). As spectral images capture the colour information of an image considerably more accurately than grayscale or RGB images, there is significant potential for the improvement of lesion segmentation. However, as the spectral dimensionality increases, preprocessing and the extraction of meaningful features becomes less obvious.

One great benefit of supervised classification is that the often complicated relationship (especially in the case of spectral data) need not be known. The classifiers simply find an optimal (for the given classifier) separation of the training examples, and uses those parameters to classify new data. Potentially, a good segmentation performance can be achieved without completely understanding the characteristics of the spectra from healthy and pathological tissue. However, all pathological changes in the retina do not show as salient lesions in the fundus images. It may not be possible for even a trained ophthalmologist to diagnose states such as changes in blood circulation or loss of macular pigment from typical fundus images. Still, the information of the pathological change is retained in the reflectance spectrum, and can be extracted if the spectral resolution of the measurement device is sufficient.

Prior knowledge of the formation of the reflectance spectra in spectral fundus images can be applied to the extraction of diagnostically interesting regions from the image spectra. With sufficient understanding of the image formation process and the properties of tissues and histological parameters affecting the reflected spectrum, an estimation of the concentrations of the diagnostically interesting parameters can be derived from the reflectance spectra.

The distributions of certain histological parameters are of diagnostic interest and can indicate a forming pathology. The decrease in retinal circulation has been linked to diabetic retinopathy [125, 241] and glaucoma [183], and the loss of macular pigment has been (indirectly) linked with AMD [20]. Fundus spectra have been used for the measurement of the histological parameters. This includes the estimation of melanin

concentrations [100], the quantitative analysis of retinal blood flow [48], and the detection of changes indicative of retinal pathologies [121,123,201]. Methods for spectral reflectance measurements, especially spectral imaging of the retina, have been developed [22,55,69,107,120,216].

4.2 Lesion detection by supervised classification

The spectral fundus image data presented in Chapter 3 can be directly used for both training and testing of supervised classifiers for the detection of lesions related to retinal pathologies. The ophthalmologists annotating the images were asked to mark a representative region of each annotated lesion (or group of lesions), which contains only pixels that represent the lesion (the gold standard annotations are discussed in detail in Section 3.3.3). The representative regions can be used as positive examples of lesion spectra, while negative examples can be extracted from random regions from the spectral image background (with no lesion annotations).

Several well-known supervised classification schemes, with applications in the segmentation or classification of retinal lesions, do not have restrictions on the dimensionality of the data and are, thus, directly applicable to spectral images. SVM has been used for the classification of retinal lesion (e.g., [2,27,246]). Support vector machine (SVM) is generally used to classify candidate lesion regions (healthy vs. unhealthy or to a specific lesion type) after a segmentation step. Neural networks (NNs) have been used to classify retinal lesions in, e.g., [70,71,155]. Similar to SVM, NNs are usually applied after a separate segmentation step. Random forests (RFs) have been popular especially in the field of medical image analysis. RF classifiers have been used in segmentation and classification of retinal images (e.g., [3,114,190,210]). Kauppi [115] used GMMs as the baseline retinal lesion detection method for the DiaRetDB1.

4.2.1 SVM

SVM is a linear machine classifier with the data preprocessed by mapping them into a higher dimension where the existence of a hyperplane separating the two classes of data can be ensured. Support vectors are the transformed training samples with distance equally close to the separating hyperplane. The optimal separating hyperplane is one that maximises the margin between the support vectors of the different classes. The class of new data is determined by which side of the hyperplane the transformed data resides. [52]

4.2.2 Gaussian Mixture Models

A Gaussian mixture model (GMM) models data by a set of weighted Gaussian distributions parametrised by corresponding mean μ_i and covariance matrix Σ_i . The GMM for multivariate data is defined

$$p(\theta) = \sum_{i=1}^n \omega_i \mathcal{N}(\mu_i, \Sigma_i), \quad (4.1)$$

where $p(\theta)$ is the likelihood of a sample with values θ , n is the number of components, and ω_i , μ_i and Σ_i are the weight, mean and covariance matrix of the i :th component, respectively.

4.2.3 Neural Networks

A (feed-forward) NN consists of a set of computing nodes arranged in layers such that the output of each node in a layer is connected to the input of all nodes of the next layer. The input to a node is the weighted sum of the outputs from the previous layer adjusted by an activation function. The architecture of individual interconnected nodes is reminiscent of the human brain, where a large number of interconnected simple neurons are capable of solving complicated tasks. [75].

4.2.4 Random Forests

In RF, a number of decision tree classifiers are generated from independent random samples from the training data. The tree classifier \mathcal{O}_i is trained using a vector \mathcal{V}_i generated by randomly sampling the training data. The vectors $\mathcal{V}_{1\dots n}$ are independent but follow the same distribution. After the RF model is trained, new (input) samples are classified by the most popular class assigned by the independent tree classifiers [32].

4.2.5 Evaluation

To quantify the performance of each classifier, the receiver operating characteristic (ROC) curves were computed for the pixel-wise classification results for each method. ROC curves have found frequent use in medical imaging in evaluating the performance of computational methods for decision support, diagnosis, and prognosis [135]. The performance of a binary classifier can be characterised by sensitivity and specificity, defined as

$$SN = \frac{TP}{TP + FN}, \quad (4.2)$$

$$SP = \frac{TN}{TN + FP}, \quad (4.3)$$

where TP is true positive classification, TN is true negative, FP is false positive and FN is false negative, respectively. In a medical setting, SN can be seen as a measure for how well a method is able to classify pathological data as pathological. Similarly, SP can be seen as a measure of how likely the data contain pathologies when the method classifies them as such.

For a classifier returning continuous membership values, the calculation of the SN and SP values requires the selection of a threshold to produce a binary classification for the data. As a less discriminatory threshold value results in increased detection rate at the cost of incorrect classifications, and vice versa, the threshold selection becomes a compromise between SN and SP. The resulting SN and SP values are valid only for that specific threshold and consequently do not properly characterise the classifier performance as a whole.

ROC curve is the relationship of SN and SP estimated over linearly increasing thresholds. The shape of the curve, and area under the curve (AUC), gives a more complete understanding of the method performance.

4.3 Histological parameter maps from spectral images

Instead of classifying image spectra into representations of pathologies or healthy tissue, the model of fundus image formation (see Section 2.3.2) can be used, together with the result spectra, to estimate the histological parameters that contribute to the formation of the spectra by the absorption and scatter of light. With a one-to-one mapping (or an approximation of one) between the model and image data, the model of image formation can be used to estimate what molecular concentrations formed the reflectance spectrum captured in the image.

4.3.1 Model generation

The model is characterised by the concentrations of retinal haemoglobin, RPE melanin, choroidal melanin, macular pigment and choroidal haemoglobin. The optical parameters, such as absorption and scatter, of the retinal tissues and the molecules they contain are considered constant between individuals. While the optical properties of the retinal tissues and molecules can be expected to be constant between individuals, the thickness of the tissue layers can have a larger variance. However, the *in vivo* measurement of the layer thicknesses is difficult. The selection for the constant optical parameters of the model is discussed in Section 2.3.2.

The model spectra were computed for a set of parameters, with 9 unique values for choroidal haemoglobin, 7 for choroidal melanin, and 10 unique values for RPE melanin, macular pigment and retinal haemoglobin. As the number of unique parameter values is constrained by the increasing computational cost, the choice of limiting the number of simulated choroidal layer values was influenced by the fact that no prior distributions for the choroidal parameters are known (i.e., the possibility to evaluate the choroidal parameter maps extracted from the model is very limited). The combination of parameters resulted in a total of 63000 spectra. The parameters are shown in Table 4.1.

Table 4.1: Parameters for the model generation (arbitrary units).

Choroidal haemoglobin	0.95	1.90	2.85	3.80	4.75	5.70	6.65	7.60	8.55	-
Choroidal melanin	0.0	0.165	0.33	0.495	0.66	0.83	1.0	-	-	-
RPE melanin	0.80	1.22	1.64	2.375	3.11	3.845	4.58	5.07	5.56	6.05
Macular pigment	0.0	0.025	0.05	0.0713	0.10	0.185	0.30	0.40	0.50	0.60
Retinal haemoglobin	0.0	0.0588	0.12	0.1792	0.25	0.3651	0.50	0.625	0.75	0.875

4.3.2 Alignment of model and image data

Without proper calibration, no mapping between the image spectra and the spectra generated by the model exists. As the concentrations of the molecules in the retina vary between subjects, and *in vivo* measurements are not possible in practice, calibration is

generally not available. Styles et al. [216] propose the use of image quotients to allow the mapping to be determined for uncalibrated data. Instead of using the absolute spectral intensities, the n spectral channels are divided by the longest wavelength channel to produce an $n - 1$ channel vector of relative image quotient values.

The longest wavelength channel is used due to the optical properties of the eye towards the infrared. Haemoglobin, a strong absorber in short and middle wavelengths, is nearly transparent in the longer wavelengths. With the strong scattering of long wavelengths by water, the lack of strong absorption results in a near diffuse reflection with limited information of individual retinal structures.

In addition, as the general reflectance of the eye fundus is highest in longer wavelengths [51], the maximal reflectance values can generally be found in the channel with the longest acquisition wavelength. Using that channel to produce the relative quotients results in quotient values in the range [0 1]. For spectral imaging systems acquiring all channels with a single exposure, the largely diffuse and strong reflection of the long wavelength channel is a relatively close estimate of the illumination field present in all the channel images. Therefore, computing the quotients will remove much of the effects of uneven illumination.

However, this approach assumes that all channels are acquired with the same exposure. In systems where the channels are acquired individually, the variance in exposure times (resulting in different levels of motion blur on different channels) and illumination fields can cause significant error in the computation of the quotients.

Prior knowledge of the histology of the retinal structures can be used to find a mapping between the model and the image data. The values of the foveal region form reasonably distinct distributions in the histological parameter space. Additionally, the foveal region is usually distinct in retinal images and can be segmented with reasonable accuracy. Running the MC simulation with parameters realistic for the foveal region (e.g., high concentrations of macular pigment and no haemoglobin) provides quotient values that can be compared with the image quotients to align the model with the image data. As a number of fovea detection methods have been presented [157, 203, 205], the process of model alignment can be automated.

The transformation aligning the model values with the image data is estimated by aligning the model and the image foveal region data clouds in principal component analysis (PCA) space. A scaling factor s is derived from the ratio of standard deviations of the image foveal region data PCA scores and the model foveal region values Q projected into the image data PCA space

$$s = \frac{S_\sigma}{\tilde{S}_\sigma}, \quad (4.4)$$

where S is a matrix of image foveal region data principal component scores and \tilde{S} is defined as

$$\tilde{S} = CQ, \quad (4.5)$$

where C is a square matrix where each column is a principal component vector of the image foveal region data.

The translation t is derived from the difference in the means of the foveal region values after projecting the model data into the image foveal data PCA space. The translation

is defined

$$t = \bar{Q}_{img} - s(\mathbf{C}\bar{Q})\mathbf{C}^T, \quad (4.6)$$

where Q_{img} is a matrix of image foveal region values.

The aligning transformation is achieved by projecting the model data into the space defined by the image foveal region data principal components. Multiplying the projected data by the scaling factor s , projecting back to the original space, and adding the translation vector t to the projected model data aligns the data clouds. The full alignment is defined as

$$\tilde{M} = s(\mathbf{C}Q)\mathbf{C}^T + t. \quad (4.7)$$

The transformation parameters that align the foveal region data are used to align the full model with the image data. Examples of the resulting alignments on the full model and image data are shown in Figure 4.1. Some portion of the image data falls outside the model. This is in part explained by the illumination field bias in the channel images, and the spectra of the optic disk, which is not included in the model.

4.4 Experiments and results

4.4.1 Lesion detection by supervised classification

The performance of SVM, GMM, NN and RF in the pixel-wise detection of lesions was evaluated using the 55 spectral images of diabetics from DiaRetDB2 presented in Chapter 3. The representative regions of exudates and haemorrhages, and random background regions outside the full annotations where no lesions were present, were used to train the different classifiers. For a reference to the performance of the spectral data, RGB projections of the same images were used to train another set of classifiers, with the performance evaluated using the (RGB projections of the) same test set. Matlab built-in functions were used for the SVM, NN and RF classifiers, and the MVPRMATLAB-toolbox [111] implementation for GMM.

Due to the limited number of examples (not all images contained all lesion types and many images contained only a few lesions), the classifier performance was evaluated by leave-one-out cross-validation. Each of the 55 images were in turn used as the test image, with the lesions in the remaining images providing the training examples.

As the representative regions amount to only a part of the complete set of lesion pixels, testing the classifier output against only the representative regions would cause significant bias to the classification result. Similarly, the full annotations contain a significant number of background pixels, again biasing the classification result as some of the target classes would be opposite of what the classifier was trained with. While pixel accurate annotations would be required for a completely unbiased evaluation, the bias can be reduced by classifying the representative regions as lesions and areas outside the full annotations as healthy. As the true class for the regions inside the annotation but outside the representative region is not known, those areas are not included in the evaluation. While less biased than the other two approaches, some positive bias will be introduced to the results.

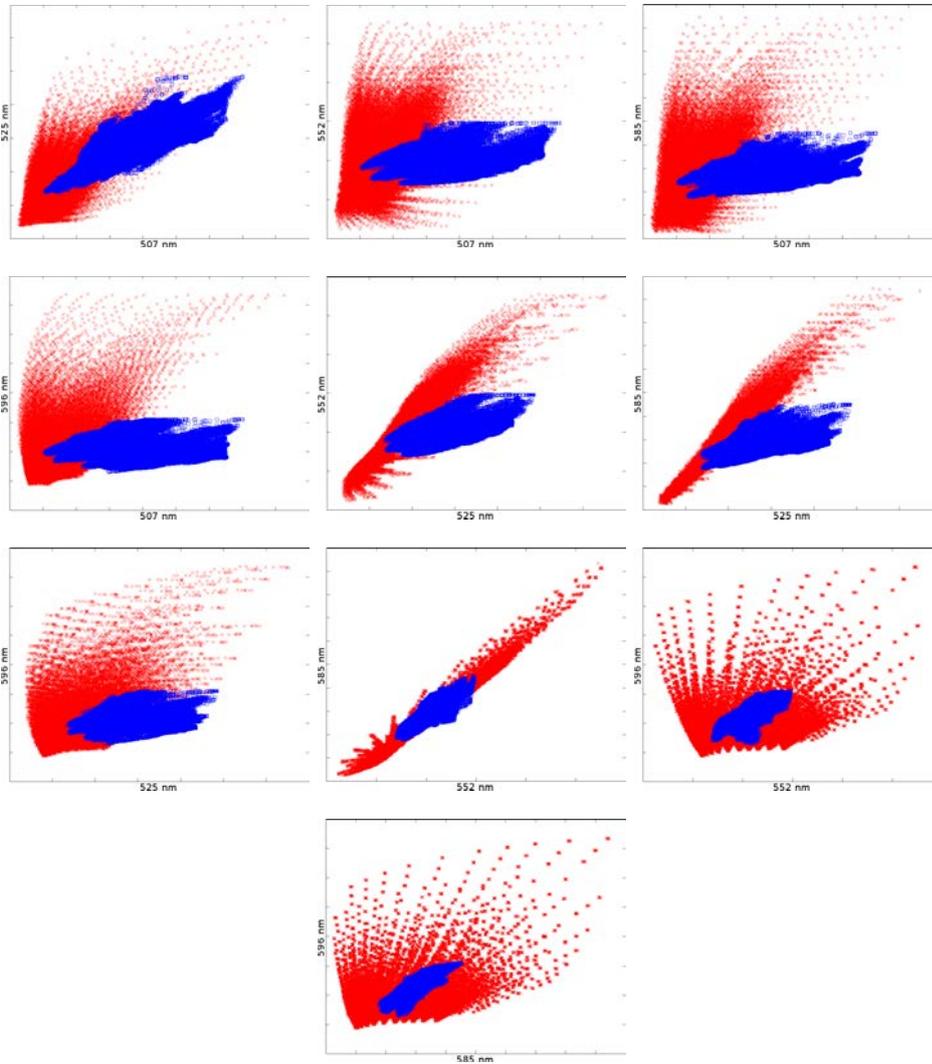


Figure 4.1: Model and image data after alignment (image 1). The dimensions of the five-dimensional data are plotted pairwise. The model values are plotted in red and the image data in blue.

Trained with the unprocessed image spectra, the SVM, RF and NN classifiers generally produced high scores for the true lesion locations. However, especially the classifiers trained with examples of soft exudates give high scores for the background spectra. Examples of the classifier scores are given in Figures 4.2, 4.3 and 4.3.

The inter-expert variation in the annotations was evaluated by considering, in turn, one set of annotations as ground truth and evaluating the performance (measured as sensitivity and specificity) of the remaining set of annotations. The mean results of the

inter-expert evaluation over all annotations are presented in Table 4.2.

Table 4.2: Mean and standard deviation (in parentheses) inter-expert agreement.

	Exudates	Soft exudate	Hard exudate	Haemorrhage
Annotator 1				
Sensitivity	0.4038 (0.3688)	0.5653 (0.4778)	0.3121 (0.3672)	0.4220 (0.3714)
Specificity	1.0000 (0.0000)	1.0000 (0.0001)	1.0000 (0.0000)	1.0000 (0.0001)
Annotator 2				
Sensitivity	0.8757 (0.3232)	0.5691 (0.5324)	0.9711 (0.0744)	0.8637 (0.3378)
Specificity	0.9997 (0.0017)	0.9998 (0.0010)	0.9999 (0.0004)	0.9997 (0.0009)

The ROC curves and the AUC for the different classifiers summarising the results are shown in Figure 4.5, Figure 4.7 and Figure 4.6, and Tables 4.3 and 4.4. It should be noted that due to the spatial coarseness of the ground truth, the pixels inside the lesion annotation but outside the representative region are not considered in the quantitative evaluation. This will cause a positive bias to the classification results.

Likely due to the limited amount of training data compared to the dimensionality of the spectra, the GMM classifiers trained with the raw spectra consistently failed to produce reasonable results (i.e., the construction of the GMM failed or the resulting GMM returned uniform class probabilities for the whole image). Therefore, the GMM results are omitted. The SVM classifier trained with the RGB data exhibited similar behaviour, explaining the linear shape of the ROC curve.

Table 4.3: Classifier performance with annotations from Annotator 1. Mean area under the curve (AUC) for support vector machine (SVM), random forest (RF) and neural network (NN) classifiers.

	Exudates	Soft exudate	Hard exudate	Haemorrhage
Spectral data				
SVM	0.8219	0.7721	0.8343	0.7924
RF	0.8327	0.7543	0.8372	0.7892
NN	0.8159	0.6396	0.8133	0.8096
RGB data				
SVM	0.6286	0.4200	0.7250	0.5272
RF	0.8118	0.7216	0.8263	0.7801
NN	0.7879	0.6832	0.7790	0.7616

The inter-expert agreement between the annotators was relatively low. While the high specificity of the annotations is explained in part by the ratio of background and foreground (i.e., the ratio of true negative (TN) and false positive (FP)), there was a significant difference in the sensitivity between the annotators. Annotator 2 generally marked more area as containing lesions than Annotator 1. The annotators disagreed especially on the classification of soft exudates.

The classifiers trained with the spectral data showed similar performance. However, there was a significant difference between classifiers trained with data extracted from the

Table 4.4: Classifier performance with annotations from Annotator 2. Mean area under the curve (AUC) for support vector machine (SVM), random forest (RF) and neural network (NN) classifiers.

	Exudates	Soft exudate	Hard exudate	Haemorrhage
Spectral data				
SVM	0.7836	0.5230	0.8178	0.7907
RF	0.8103	0.6125	0.8213	0.7982
NN	0.7965	0.7421	0.8143	0.8099
RGB data				
SVM	0.5271	0.4214	0.5794	0.8055
RF	0.8035	0.7167	0.8144	0.7979
NN	0.7936	0.7221	0.8134	0.7996

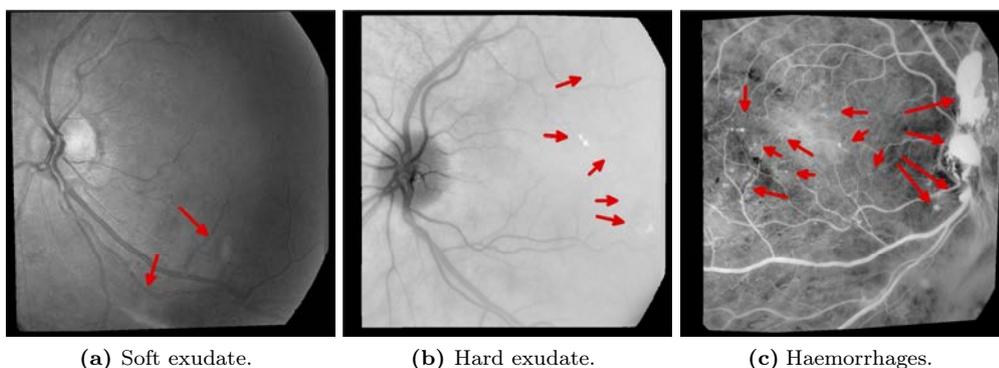


Figure 4.2: Support vector machine (SVM) classifier scores for different lesions (trained with spectral data). Intensity value depicts the estimated likelihood of the pixel belonging to the lesion class.

annotations of Annotator 1 and Annotator 2. The classifiers provided best accuracy for hard exudates, while soft exudates were classified significantly more poorly.

Using the refined annotations did not significantly improve the performance of the NN and SVM classifiers using spectral data, but the performance of the classifiers trained with RGB data improved somewhat when using refined annotations (see Tables 4.5 and 4.6). Contrary to the coarse annotations, the SVM classifiers trained with RGB data using the refined annotations produced relatively good results. The results are visualised in Figures 4.8, 4.9 and 4.10 .

4.4.2 Histological parameter maps from spectral images

Quantitative evaluation of the histological parameter maps derived from image spectra by inverting the model is not possible as neither *in vivo* or *in vitro* measurement of the molecule concentrations in the target eye can be performed. Qualitative assessment of the resulting parameter maps is also problematic to a degree. There is only some

Table 4.5: Classifier performance with annotations from Annotator 1. Mean area under the curve (AUC) for support vector machine (SVM), random forest (RF) and neural network (NN) classifiers using refined annotations.

	Exudates	Soft exudate	Hard exudate	Haemorrhage
Spectral data				
SVM	0.8140	0.7390	0.8358	0.8043
RF	0.8279	0.7064	0.8373	0.7918
NN	0.8213	0.7022	0.8266	0.7831
RGB data				
SVM	0.7839	0.6774	0.8077	0.8133
RF	0.8039	0.7151	0.7876	0.6877
NN	0.7974	0.7161	0.8328	0.7974

Table 4.6: Classifier performance with annotations from Annotator 2. Mean area under the curve (AUC) for support vector machine (SVM), random forest (RF) and neural network (NN) classifiers using refined annotations.

	Exudates	Soft exudate	Hard exudate	Haemorrhage
Spectral data				
SVM	0.8203	0.5138	0.8165	0.7998
RF	0.8060	0.6168	0.8148	0.8051
NN	0.7966	0.8074	0.8159	0.8006
RGB data				
SVM	0.7921	0.7190	0.5792	0.7821
RF	0.8076	0.7207	0.8128	0.7890
NN	0.8073	0.7313	0.8152	0.7795

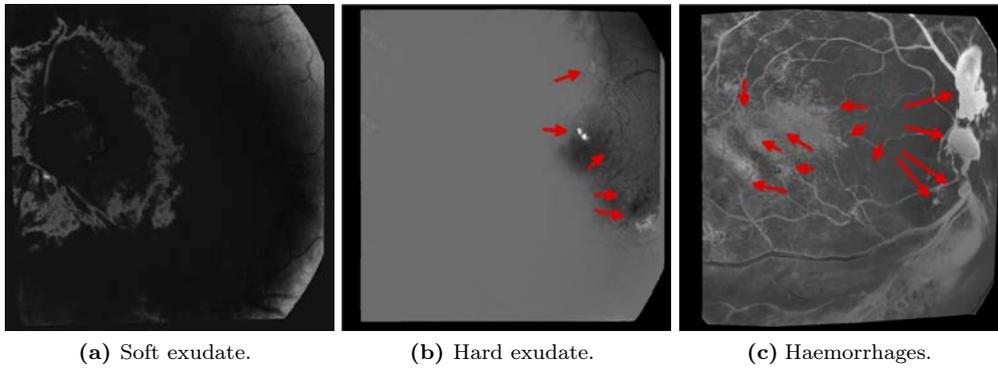


Figure 4.3: Neural network (NN) classifier scores for different lesions (trained with spectral data). Intensity value depicts the estimated likelihood of the pixel belonging to the lesion class.

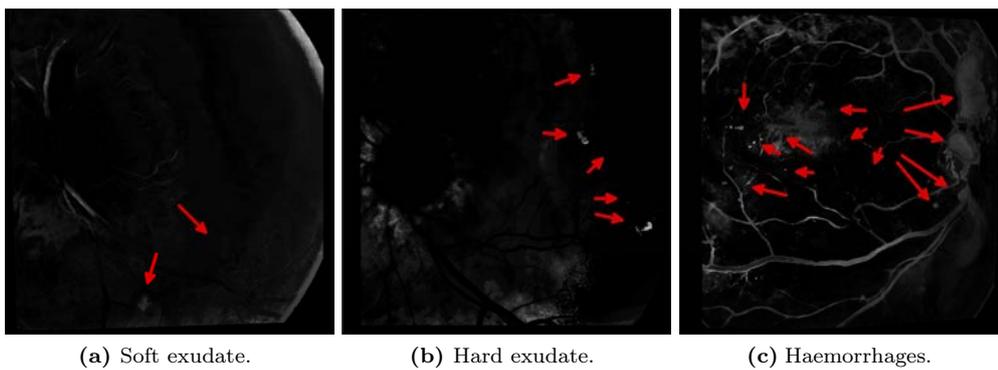


Figure 4.4: Random forest (RF) classifier scores for different lesions (trained with spectral data). Intensity value depicts the estimated likelihood of the pixel belonging to the lesion class.

statistical knowledge of the retinal parameters, and for the choroidal parameters no prior distributions are available.

While the accuracy of the histological parameter maps is difficult to evaluate, some retinal structure provide reference for the evaluation. The significant majority of the retinal haemoglobin should be found in the retinal blood vessels. Furthermore, macular pigment should appear only in the macula, and the macula should show an increase in the amount of RPE melanin. These factors can be used to guide the qualitative evaluation of the derived parameter maps.

A set of parameter maps were derived from the image data of 7 spectral images using k-nearest neighbour (kNN) to select the nearest quotient in the model for each image quotient. The images were selected from DiaRetDB2 based on perceived health of the retina of the subject (the model is defined for healthy tissue) and image quality. The

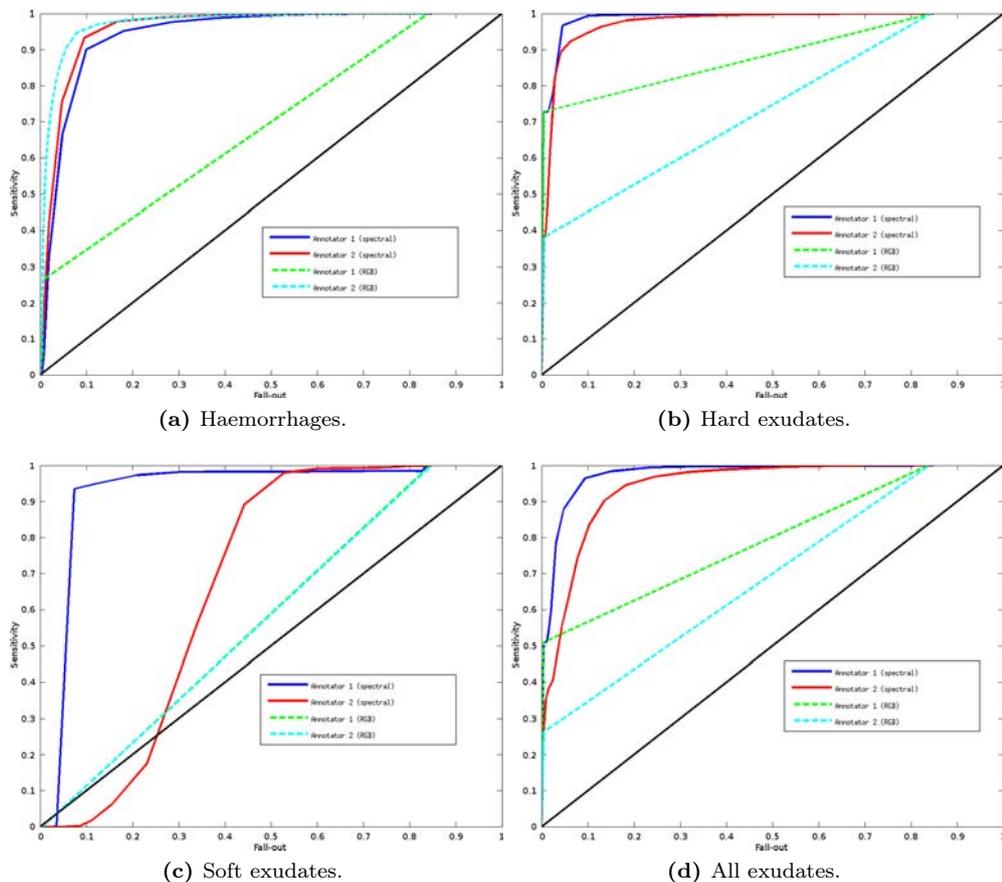


Figure 4.5: Support vector machine (SVM) classifier ROC curves.

number of suitable spectral images was limited by the fact that the majority of the subjects in DiaRetDB2 are diabetics. The maps were derived for the model after performing the alignment described in Section 4.3.2.

For one of the images, the image foveal region data do not have the expected relationship with the rest of the image data. Figure 4.11 shows an example of a failed alignment due to a strong illumination field bias affecting the image foveal region. Examples of the derived parameter maps are shown in Figures 4.12 and 4.13.

The illumination field bias present in the spectral images is assumed to affect the inversion result. Straightforward channelwise correction of the illumination field bias drastically changes the shape of the point cloud of image data (see Figure 4.14), causing the alignment approach to fail and the inversion to produce unrealistic parameter maps.

Applying the illumination correction approach proposed in Section 2.7 retains the shape of the image data point cloud and parameter maps derived from the corrected data show improved correspondence to the histological prior. The parameter maps derived from

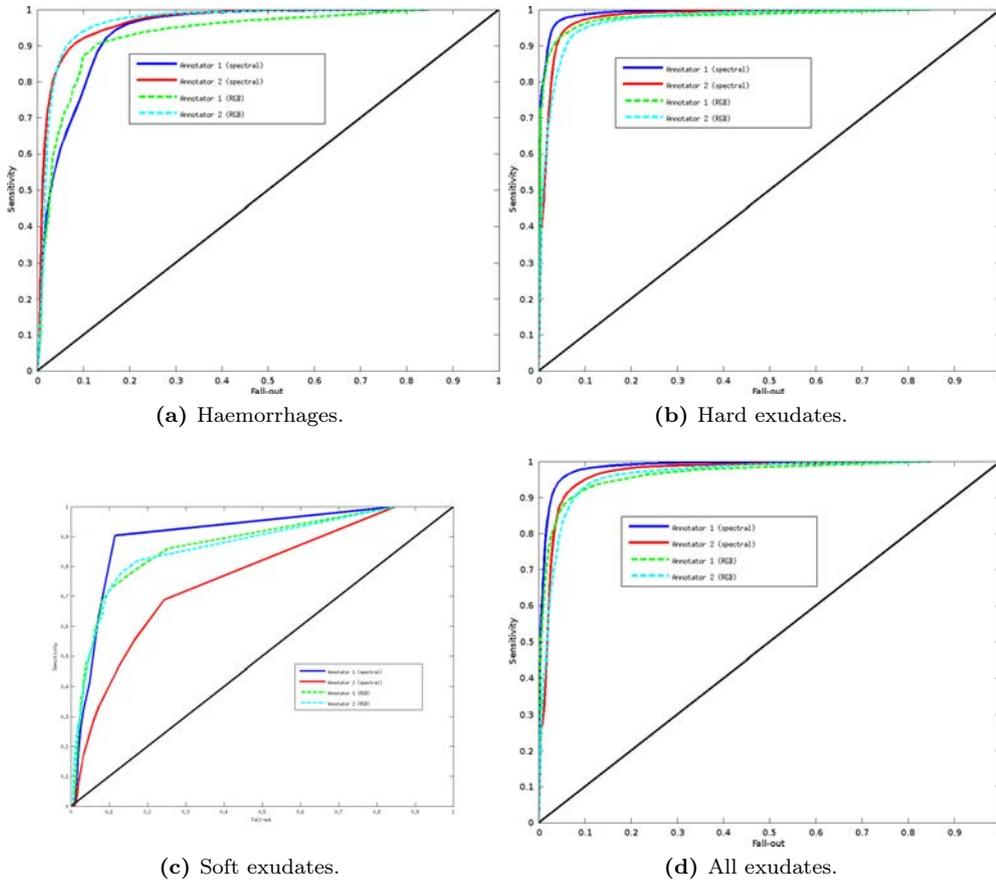


Figure 4.6: Random forest (RF) classifier ROC curves.

the corrected image data are shown in Figures 4.15 and 4.16.

For the most part, the parameter maps follow distributions expected from the retinal structures. Retinal haemoglobin is found mostly in the blood vessels. Increased haemoglobin is also found at haemorrhages, and reduced haemoglobin is found in the avascular region at the macula. As expected, macular pigment is concentrated at the macula. The macula also shows an increased amount of RPE melanin, which is consistent with the histological prior.

Parameter maps derived from the images where the alignment failed followed the expected distributions poorly (see Figure 4.13). It is not known whether the poor alignment is due to errors in the spectra during the acquisition, changes in the retina due to diabetes, or some other factor not included in the model.

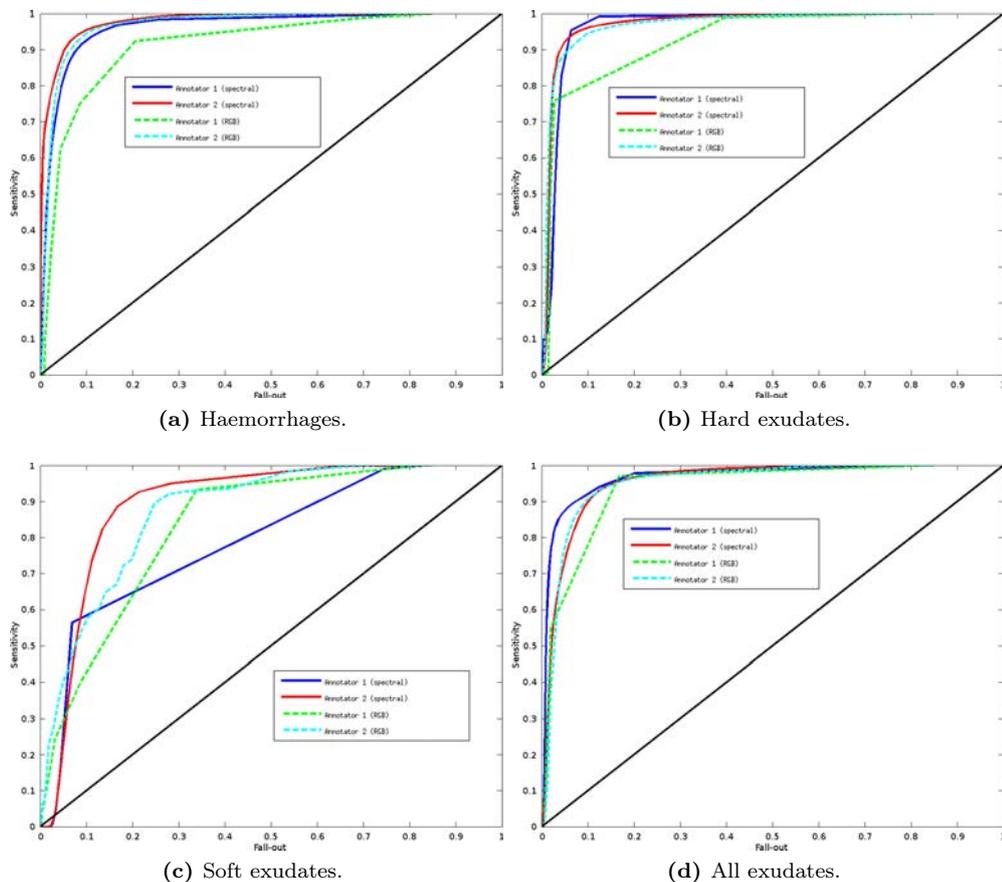


Figure 4.7: Neural network (NN) classifier ROC curves.

4.5 Discussion

Even with the direct use of unprocessed image spectra as training data, the performance of the classifiers in the detection of retinal lesions was promising, especially for hard exudates. The lesion detection performance could likely be improved significantly by adding heuristics, optimising the acquisition wavelengths, normalising the inter-person variance of the spectral data or applying dimensionality reduction.

A major challenge for the lesion detection method development is the lack of accurate ground truth for the lesion locations. Annotation refinement can be an intermediate solution, but accurate annotations by medical experts would be required for proper method development and evaluation. A possible approach would be to implement the annotation refinement as a part of a semi-automatic annotation framework, where a medical expert would verify and correct the results of the automatic refinement.

Based on the experiments on illumination correction, the uneven illumination fields in the channel images can have a significant effect on the parameter map estimation. However,

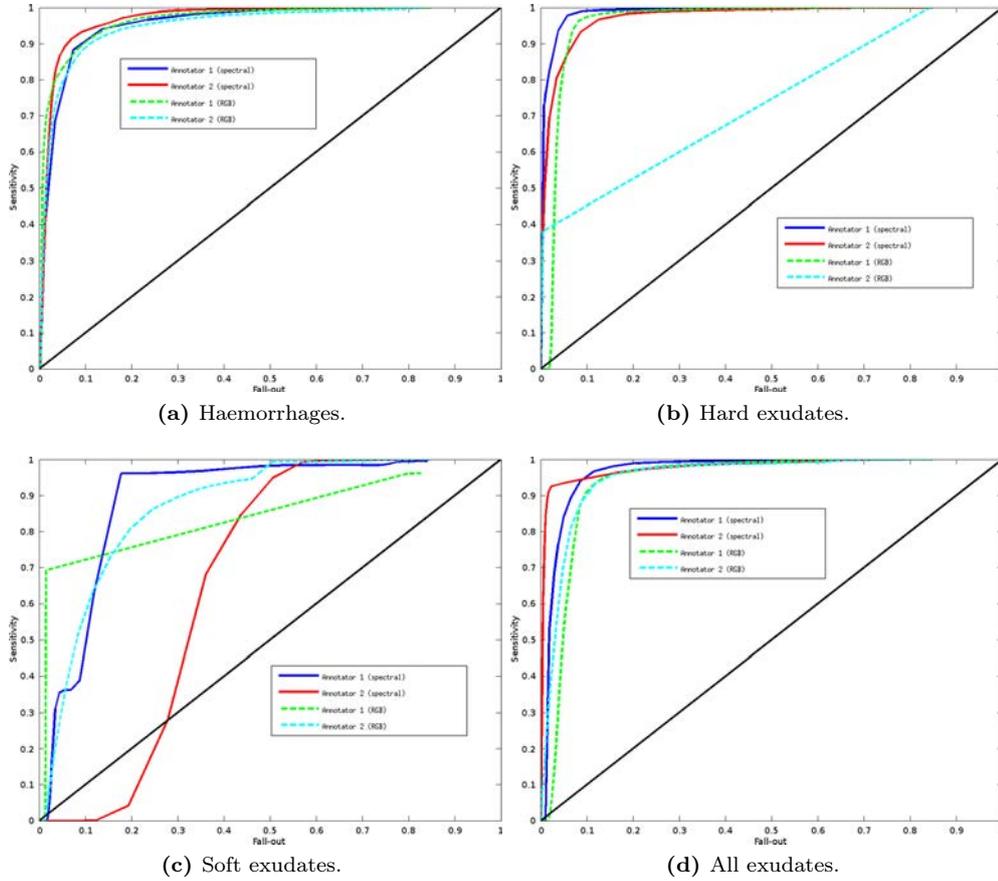


Figure 4.8: Support vector machine (SVM) classifier ROC curves (refined annotations).

the loss in the visual quality of an image due to moderate uneven illumination fields may not be very significant. Especially if the quality of an acquired spectral image is estimated channel-wise (which is likely to be the case for systems requiring a separate alignment post-processing step), it can be difficult to avoid the channel-wise illumination field bias.

Channel-wise illumination correction is not a viable option as it can affect the consistency of the spectra and consequently cause significant error in the inversion. If the illumination field cannot be compensated for in the imaging phase, the performance of the illumination correction method proposed in Section 2.7 suggests that using the retinal structures to guide the illumination field estimation is a viable approach to spectral retinal image illumination correction.

The development of the model of light interaction with the retinal tissue and the inversion method would benefit from a dedicated data set of healthy subjects with emphasis on the quality and consistency of the image spectra. With fewer sources of error, the evaluation

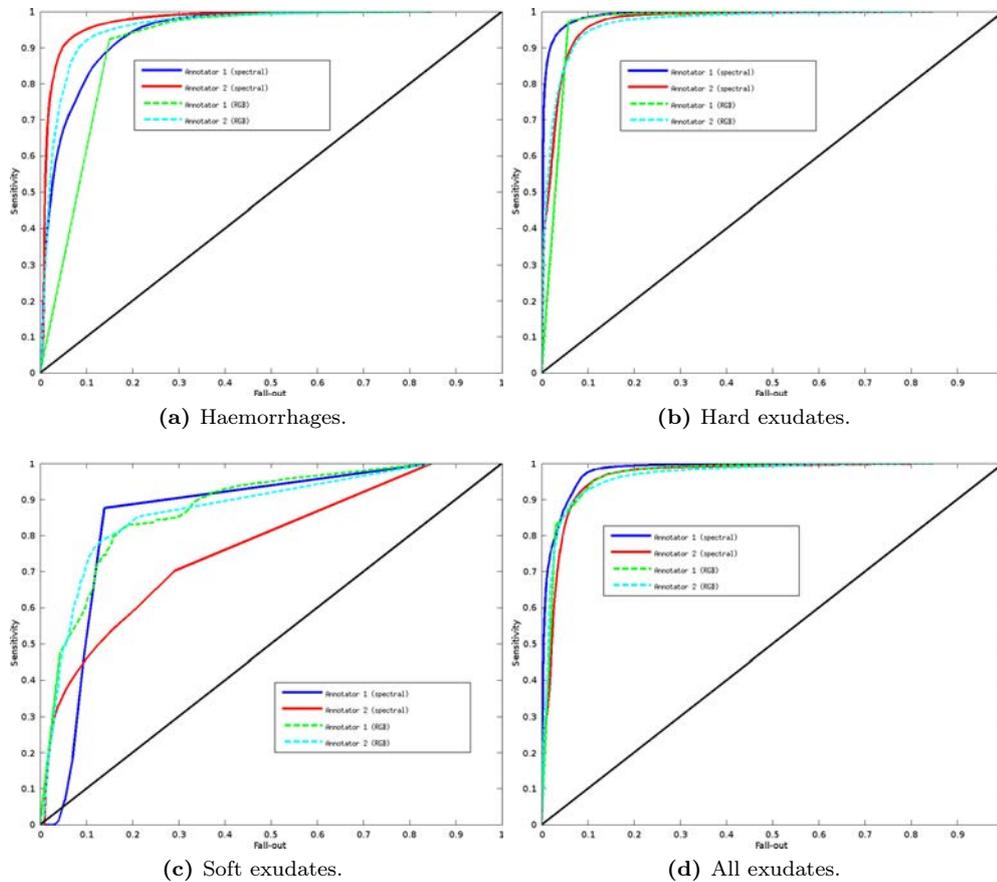


Figure 4.9: Random forest (RF) classifier ROC curves (refined annotations).

of the effects of changes in the model or inversion would be facilitated.

In addition to more suitable data, a more sophisticated inversion method can potentially provide significant improvement to the parameter map estimation. The kNN algorithm is simple, relatively fast, and guarantees the inversion result from the set of simulated model parameters (instead of, e.g., interpolated parameter values). However, kNN has a number of drawbacks. The nearest neighbour depends on the choice of distance measure, and while Euclidean distance is sufficient for many applications, it may give sub-optimal results on the high-dimensional and non-linear model. The implementation of a more refined inversion method would likely provide more accurate parameter maps.

4.6 Summary

The spectral fundus image data can directly be used for both training and testing of supervised classifiers for the detection of lesions related to retinal pathologies. Using

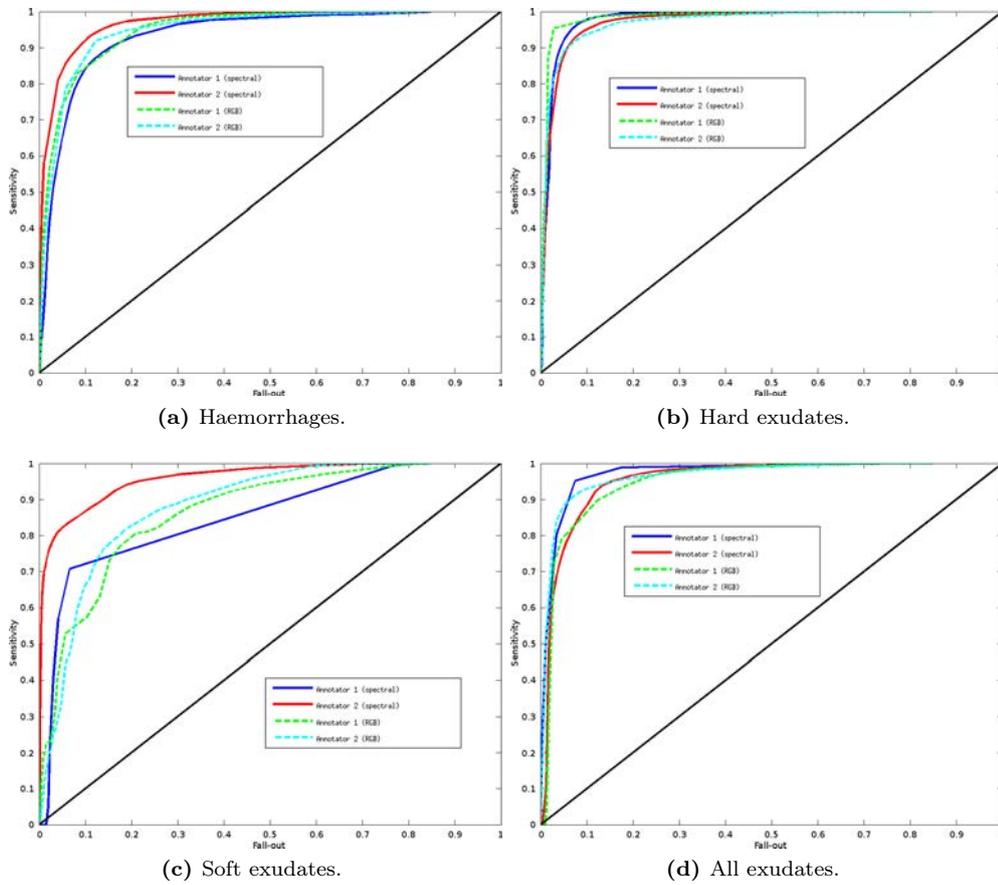


Figure 4.10: Neural network (NN) classifier ROC curves (refined annotations).

the DiaRetDB2 database with coarse expert annotations for training and testing, the performance of several supervised classifiers in the detection of lesions was evaluated. Of the exudates, the soft exudate class was more often misclassified. Excluding soft exudates, the classifiers trained with the unprocessed image spectra outperformed the classifiers trained with RGB features.

Another set of classifiers was trained using spectra and RGB values extracted from the ROI defined by the expert annotations after refinement using the method proposed in Section 3.5. Using training data extracted from the refined regions did not improve the performance of the classifiers using spectral data. For the RGB data, however, a significant increase the performance of the SVM classifier and a slight improvement in the NN classifier performance was observed.

An approach to the estimation of the molecular concentrations of retinal haemoglobin, RPE melanin, choroidal melanin, macular pigment and choroidal haemoglobin from the image spectra using a model of light interaction with retinal tissue was presented. The presented approach to aligning the model and image data was able to find reasonable

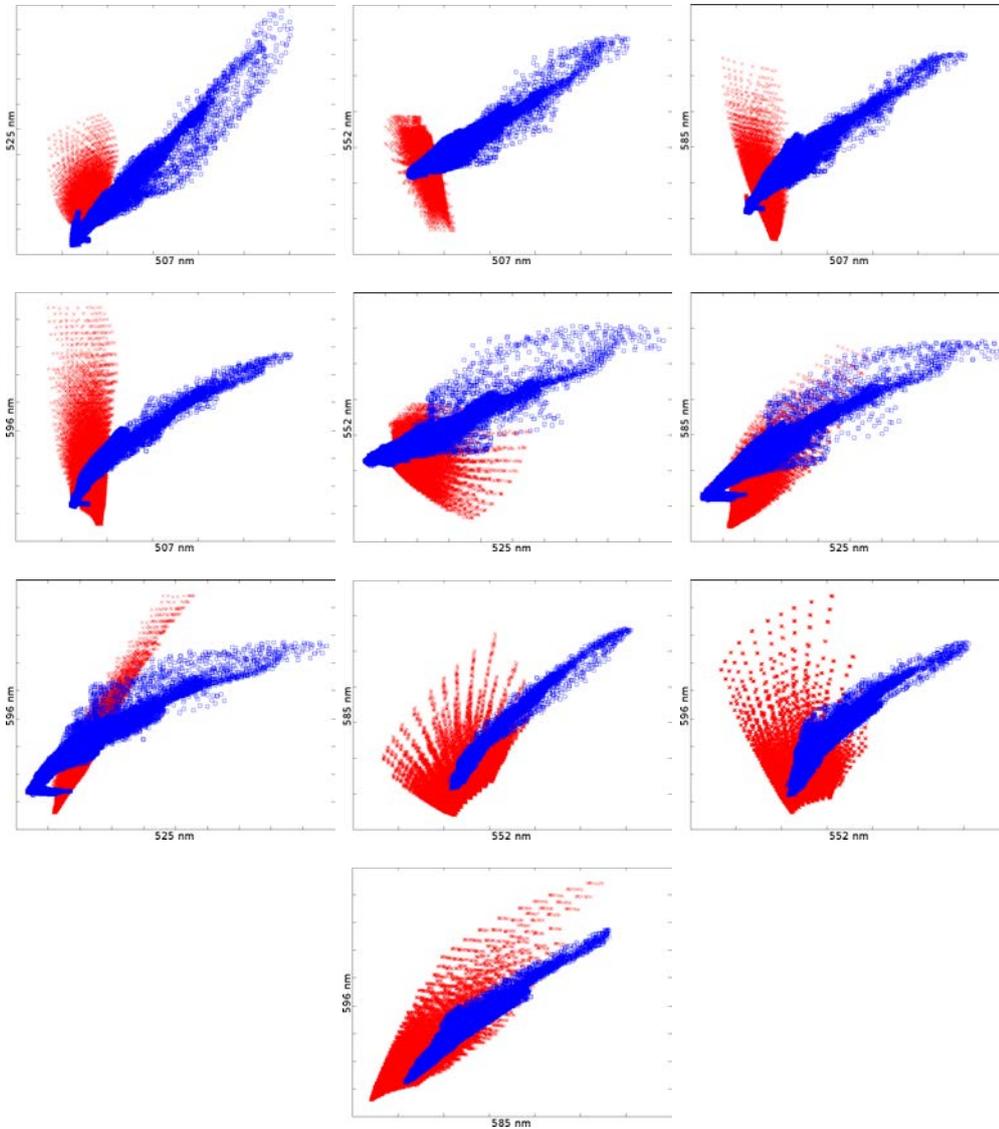


Figure 4.11: Poorly aligned model and image data (image 7). The dimensions of the five-dimensional data are plotted pairwise. The model values are plotted in red and the image data in blue.

translations and rotations, but returned too small scaling factors.

The qualitative evaluation show the parameter maps to follow distributions expected from the retinal structures. Retinal haemoglobin was found mostly in the blood vessels. Increased haemoglobin was also found at haemorrhages, and decreased haemoglobin was found in the avascular region at the macula. Macular pigment was generally concentrated

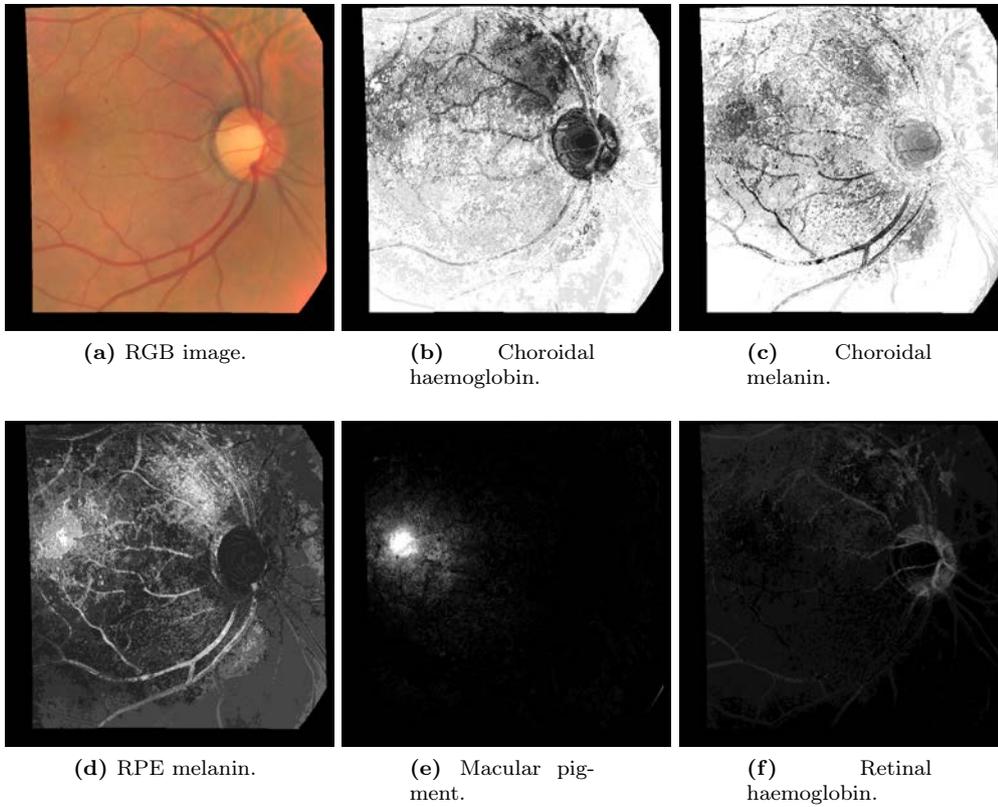


Figure 4.12: The parameter maps derived from image 5.

at the macula. The macula also showed an increased amount of RPE melanin.

Channel-wise correction of illumination field bias caused changes in the image data that resulted in unrealistic parameter maps after inversion. The illumination correction method proposed in Section 2.7 seemed to improve the inversion results (macular pigment became more concentrated to the macula region and blood vessels showed as more defined on the retinal haemoglobin channel).

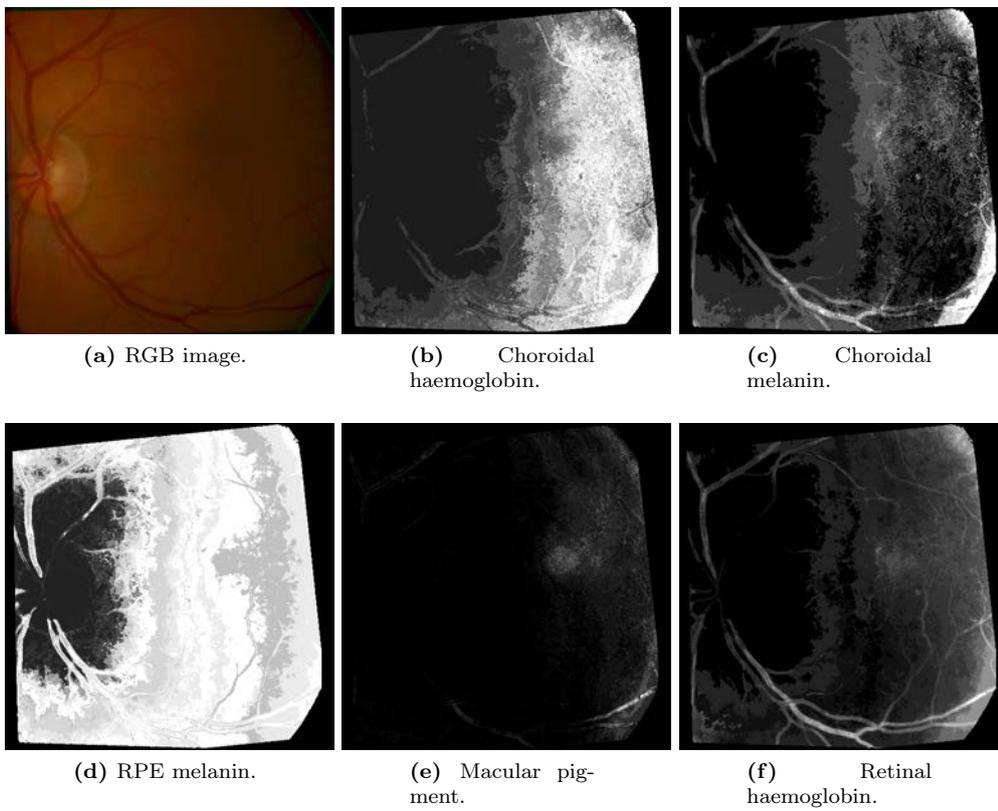


Figure 4.13: The parameter maps derived from image 7.

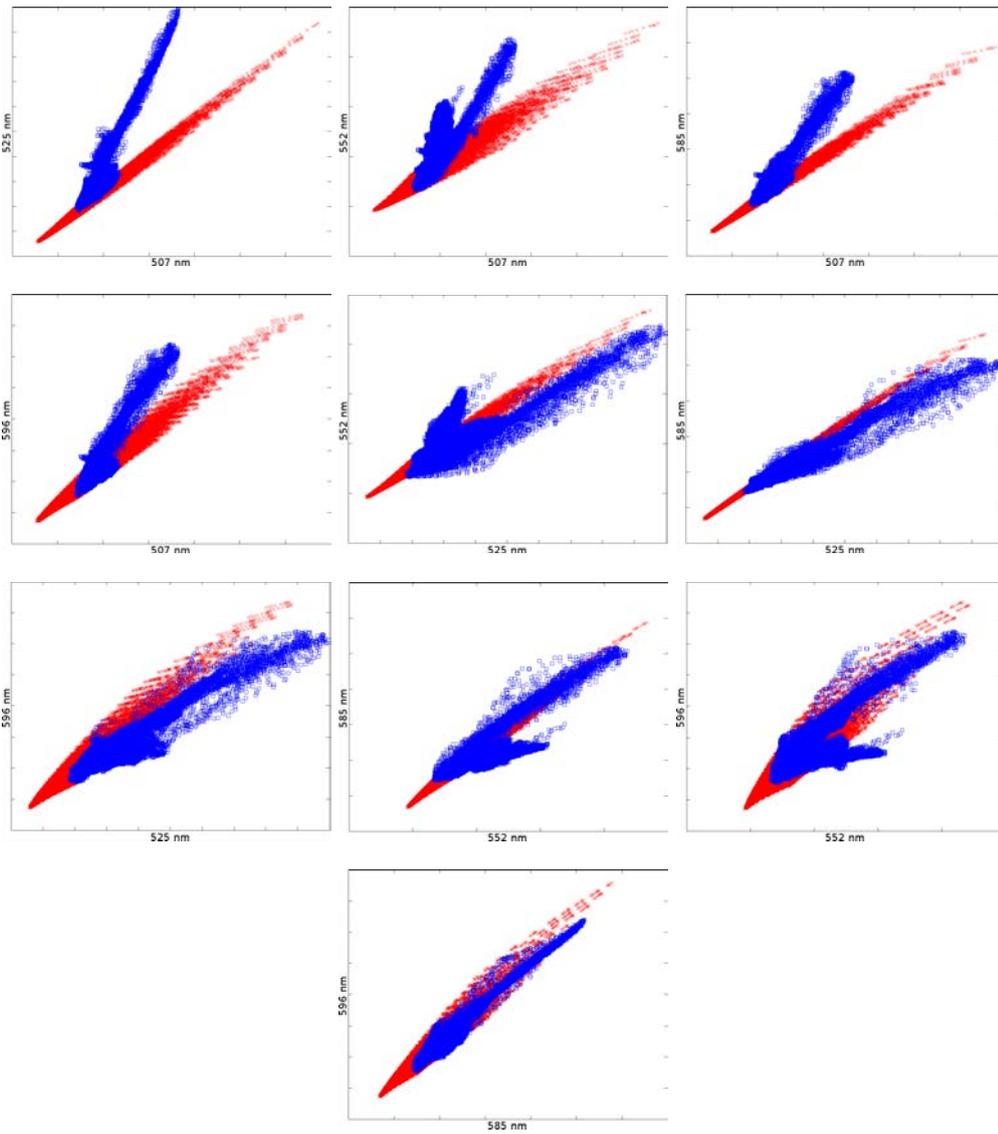


Figure 4.14: Aligned model and image data after channel-wise illumination field correction (image 5). The dimensions of the five-dimensional data are plotted pairwise. The model values are plotted in red and the image data in blue.

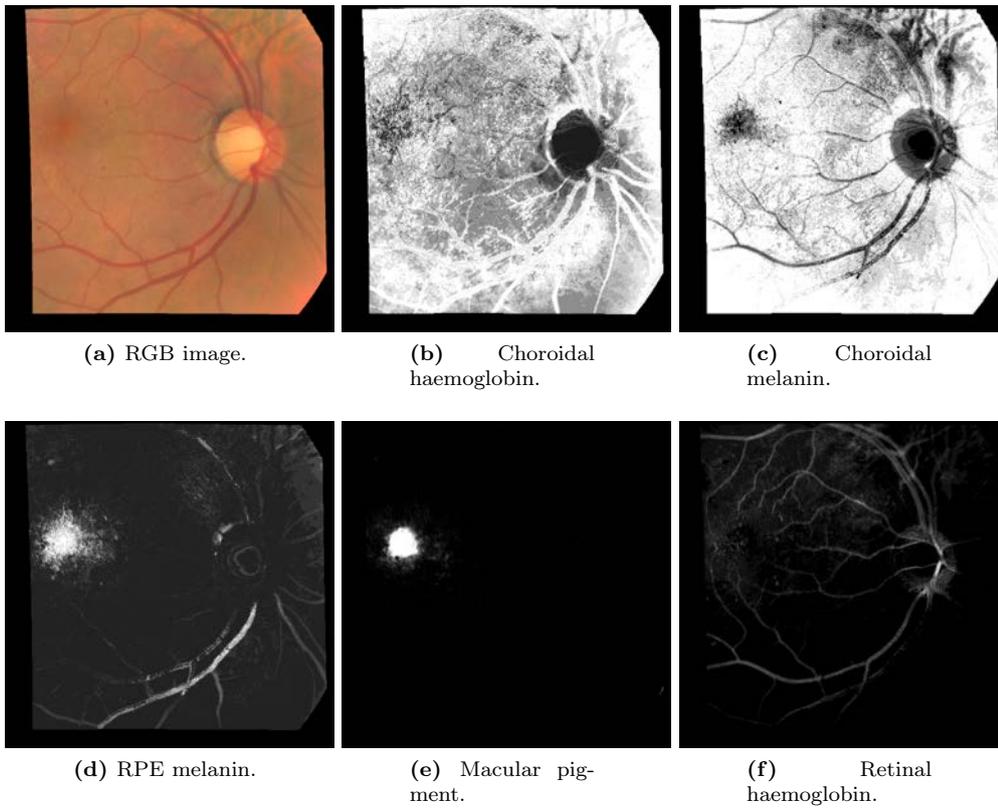


Figure 4.15: The parameter maps derived from image 5 after proposed illumination correction.

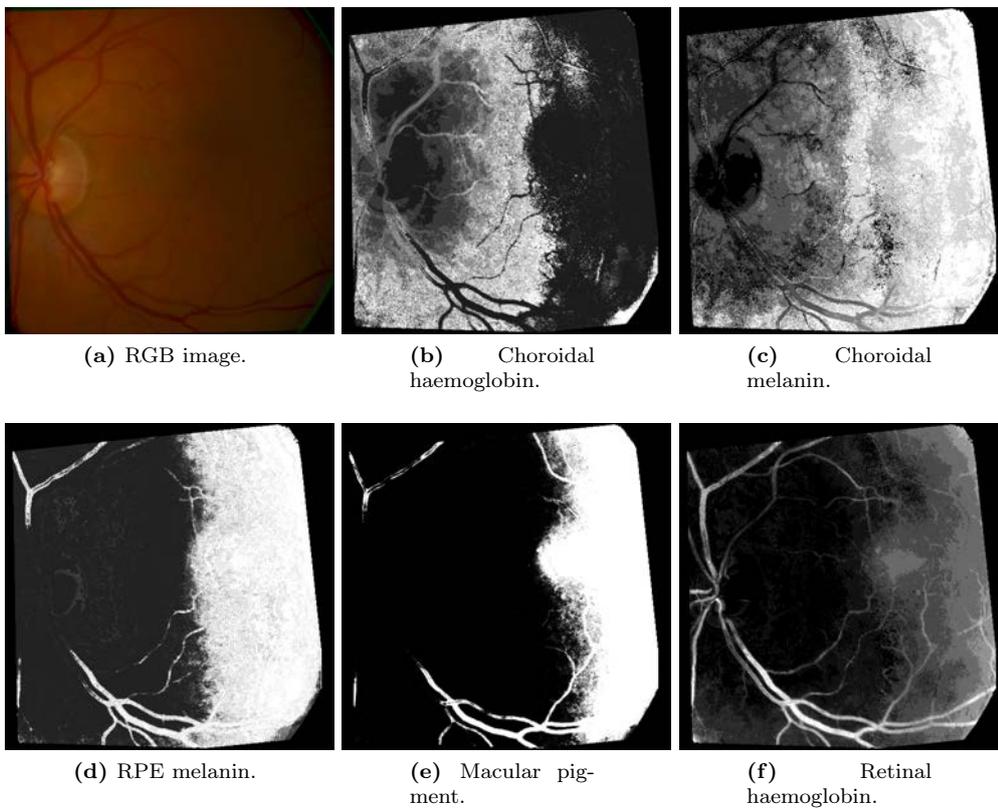


Figure 4.16: The parameter maps derived from image 7 after proposed illumination correction.

5.1 Main contributions

The methods for and considerations on composing and the post-processing of spectral retinal images presented in this thesis can potentially provide significant improvements to the quality of spectral retinal image data. The geometric and spectral calibration, correcting of uneven illumination and accounting for dirt and dust in the optics can help to remove multiple sources of errors in spectral retinal images, longitudinal data and data combined from multiple sources.

With further development and evaluation, the maps of estimated molecular concentrations of retinal haemoglobin, RPE melanin, choroidal melanin, macular pigment and choroidal haemoglobin from the image spectra may provide a useful modality for the early detection of changes in the retina or the formation of lesions. The 3D-reconstruction of the retinal shape from retinal images can be used to improve both the spectral image composition and the model of light interaction in the retina (by using the estimated retinal shape instead of uniform layers). If evaluations with medical data suggest reasonable accuracy, the reconstructions can have clinical value in the diagnosis of macular edema, glaucoma and other retinal conditions.

A supervised classifier trained with the unprocessed image spectra produced reasonable detection results for haemorrhages and hard exudates. The performance of the classifiers suggests that a lesion detection method outperforming current state-of-the-art method based on RGB data could be developed using the spectral data. With the publication of spectral retinal image database DiaRetDB2 with gold standard annotations, the development of methods relying on spectral retinal data becomes more viable, as every research group does not need to collect their own spectral retinal images and expert annotations for the development, training and evaluation.

With the ageing population and changed living habits, retinal diseases such as diabetic retinopathy (DR) and age-related macular degeneration (AMD) will increasingly burden the health care systems of the developed countries. Early detection and personalised

diagnostics directing the limited resources to the patients in most need of and most benefiting from the therapeutic interventions will be paramount in the proper management of these diseases.

Broadening the utilisation of spectral information could improve automatic and semi-automated diagnostics for large scale screening of retinal diseases and enable an early entry to a treatment program. With the improvement in the quality of the spectral data, and the development of image analysis methods based on retinal reflectance spectra, there is a potential for the improvement of the quantitative detection of retinal changes for follow-up, clinically relevant end-points for clinical studies and development of new therapeutic modalities.

5.2 Limitations of the study

One of the major challenges in developing and comparing methods for composing and processing spectral fundus images is the difficulty of obtaining meaningful reference measurements. Without a proper reference, method development becomes difficult as the performance of a method cannot be properly evaluated or compared to other approaches. *In vitro* measurements do not properly capture the characteristics of a living eye, and while techniques such as OCT and spectroscopy can be used to accurately measure the retinal structures and reflectance *in vivo*, the alignment and comparison with the fundus image spectra is challenging.

The alignment of the retinal image data and the model of light interaction in tissue remains challenging. While the foveal region data of both the model and the images produce similar distributions, the exact range of the parameter values that can be expected to be present in the macular region of the image is not known.

After the alignment, a number of images showed a significant portion of the data outside the model. Some of the data outside the model is due to retinal structures that are not included in the model, e.g., the optic disk. Another potential source for disparity is pathological changes in the retina. Problems with the image acquisition and illumination field bias can also affect the image data. Further studies are needed to determine the source (or sources) of the disparities between the model and the image data.

5.3 Future work

While the GDB-ICP registration method showed good performance on the synthetic and semisynthetic test data, the composition of a spectral retinal image from a set of channel images remains a challenging problem. A method designed specifically for the registration of spectral retinal data would likely be required for improved alignment. As general registration methods do not take into account the change in information content between the channels, the space of transformations realistic to correct for the inter-channel eye movements, and the potentially large number of images to be registered, there is room for improvement in the registration of channel images.

The use of synthetic data can provide quantitative measurements on an approximation of the *in vivo* measurements. While complete correspondence to measurements from the

true medical data should not be expected, results from quantitative evaluation using synthetic data can be more representative than comparisons derived from measurements from other modalities (e.g., sparse spectroscopy measurements or mean values based on patient information). With a sufficiently accurate synthetic data as a reference, the development and evaluation of the composition and processing of spectral fundus images would be significantly facilitated. A commonly accepted synthetic benchmark would enable proper comparison of methods by different researchers.

The stereoscopic reconstruction of the fundus can be improved by considering the spectral dimension of the data. As the depth of tissue penetration is dependent on the wavelength of the light entering the fundus, the information content of the individual channels represents the structures at different levels of the fundus. Taking into account the possible light paths in the retinal interaction would very likely benefit the reconstruction. In theory, prior knowledge (e.g., from simulation) of the expected penetration depth could be used to refine the reconstruction. With proper characterisation of both the reflectance spectra and the retinal tissue, possibly even (a number of) the individual layers of the retina could be reconstructed. A reconstruction containing both spatial and spectral information of the retina open the possibility of a subject specific optical model of the retina. Furthermore, replacing the very simplified pinhole camera model used by the current reconstruction approach with a more sophisticated model based on the fundus camera could significantly improve the reconstruction result.

While no histological ground truth for the distributions of the molecules in the retina can be achieved, prior knowledge of average distributions could be used to guide the alignment of the model of light interaction with the retina with the spectral image data. In practice this could be implemented by iterating the transformation parameters aligning the model and the image data, and computing a cost function based on the average distributions versus the inversion result. In addition, more sophisticated methods than kNN can be used for more reasonable inversion. As no histological ground truth is available for performance evaluation, the quantification of error in the intermediate steps (e.g., model generation, model alignment and inversion) becomes important in evaluating the uncertainty of the results. By estimating the error within the inversion process (e.g., by distance between image spectra and closest model reference spectra), some understanding of the uncertainty of the inversion process can be achieved.

Conclusion

In this thesis, the image processing techniques required for composing spectral retinal images with accurate reflection spectra were studied. A protocol for calibrating a fundus camera was presented, including geometric and spectral calibration, determining the spatial resolution, consideration for correcting uneven illumination and vignetting, and accounting for dirt and dust in the optics. The qualitative evaluation results suggest that significant improvements can be achieved related to the quality of spectral retinal images, longitudinal data and data from multiple sources.

The wavelength channel image registration was quantitatively evaluated using synthetic and semisynthetic data. Of the compared methods, GDB-ICP outperformed the rest. The experiments on synthetic and semi-synthetic data showed that the registration error increases with increasing wavelength difference between the floating and base image. GDB-ICP outperformed the other evaluated methods in both the number of successful registrations and registration accuracy.

A method for the 3D-reconstruction of the retinal shape from retinal images based on the method by Lin and Medioni [139] was presented. The proposed method achieved lower mean registration error than the original method, and was able to produce a reasonable reconstruction where the original method failed.

A method for correcting the uneven illumination field in spectral images was presented. The method was able to reduce the illumination field bias in the spectral image while retaining the consistency of the spectral dimension.

The steps in the acquisition of images and gold standards of lesion locations for composing the spectral fundus image database DiaRetDB2 were presented. The spectral information coupled with gold standards for multiple lesion types provide a good basis for method development in automatic detection and classification of retinal lesions.

The feasibility of the spectral retinal image data for the detection of lesions related to retinal pathologies was studied using supervised classifiers trained with the spectral data. Of the exudates, the soft exudate class was more often misclassified. While some of

the classifiers struggled with the high dimensionality of the spectral data, the classifiers trained with the unprocessed image spectra outperformed the classifiers trained with RGB features.

An approach to the estimation of the molecular concentrations of retinal haemoglobin, RPE melanin, choroidal melanin, macular pigment and choroidal haemoglobin from the image spectra using a model of light interaction with retinal tissue was presented. The density maps for retinal haemoglobin and macular pigment extracted from the image spectra correspond to the histological prior (i.e., haemoglobin in the veins and a concentration of macular pigment in the macula).

The methods and experimental results presented in this thesis suggest that spectral retinal images can be a cost-effective and valuable addition to the diagnosis and screening of eye diseases. The acquisition of accurate retinal spectra remains challenging, but with the presented methods and further development of spectral imaging and spectral image processing, the spectral retinal data can be significantly improved.

- [1] ABRÀMOFF, M., GARVIN, M., AND SONKA, M. Retinal imaging and image analysis. *IEEE Reviews in Biomedical Engineering* 3 (2010), 169–208.
- [2] ACHARYA, R., CHUA, C., NG, E., YU, W., AND CHEE, C. Application of higher order spectra for the identification of diabetes retinopathy stages. *Journal of Medical Systems* 32, 6 (2008), 481–488.
- [3] ACHARYA, U., DUA, S., DU, X., SREE, S., AND CHUA, C. Automated diagnosis of glaucoma using texture and higher order spectra features. *IEEE Transactions on Information Technology in Biomedicine* 15, 3 (2011), 449–455.
- [4] ACHARYA, U., NG, E.-K., TAN, J.-H., SREE, S., AND NG, K.-H. An integrated index for the identification of diabetic retinopathy stages using texture parameters. *Journal of Medical Systems* 36, 3 (2012), 2011–2020.
- [5] AHONEN, T., MATAS, J., HE, C., AND PIETIKÄINEN, M. Rotation invariant image description with local binary pattern histogram fourier features. In *Image Analysis*. Springer, 2009, pp. 61–70.
- [6] AIBINU, A., IQBAL, M., NILSSON, M., AND SALAMI, M. A new method of correcting uneven illumination problem in fundus images. In *Proceedings of International Conference on Robotics, Vision, Information, and Signal Processing* (Penang, Malaysia, November 2007), pp. 445–449.
- [7] AL-DIRI, B., HUNTER, A., STEEL, D., HABIB, M., HUDAIB, T., AND BERRY, S. Review-a reference data set for retinal vessel profiles. In *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society* (Vancouver, Canada, August 2008), IEEE, pp. 2262–2265.
- [8] ALAHI, A., ORTIZ, R., AND VANDERGHEYNST, P. Freak: Fast retina keypoint. In *Proceedings of Computer Vision and Pattern Recognition* (2012), IEEE, pp. 510–517.
- [9] ALM, A., AND BILL, A. Ocular and optic nerve blood flow at normal and increased intraocular pressures in monkeys (macaca irus): a study with radioactively labelled microspheres including flow determinations in brain and some other tissues. *Experimental Eye Research* 15, 1 (1973), 15–29.

- [10] ANDERSON, R., AND PARRISH, J. The optics of human skin. *Journal of Investigative Dermatology* 77, 1 (1981), 13–19.
- [11] AQUINO, A., GEGÚNDEZ-ARIAS, M., AND MARIN, D. Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. *IEEE Transactions on Medical Imaging* 29, 11 (2010), 1860–1869.
- [12] ASSOCIATION, W. M. World medical association declaration of helsinki. ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization* 79, 4 (2001), 373.
- [13] BALCI, S., GOLLAND, P., AND WELLS, W. Non-rigid groupwise registration using b-spline deformation model. *Open Source and Open Data for MICCAI* (2007), 105–121.
- [14] BANKMAN, I. *Handbook of Medical Image Processing and Analysis*, second ed. Academic Press, 2009.
- [15] BASHKATOV, A., GENINA, E., KOCHUBEY, V., AND TUCHIN, V. Estimation of wavelength dependence of refractive index of collagen fibers of scleral tissue. In *Proceedings of EOS/SPIE European Biomedical Optics Week* (2000), International Society for Optics and Photonics, pp. 265–268.
- [16] BAY, H., ESS, A., TUYTELAARS, T., AND GOOL, L. V. Speeded-up robust features (surf). *Computer Vision and Image Understanding* 110, 3 (2008), 346–359.
- [17] BEACH, J., SCHWENZER, K., SRINIVAS, S., KIM, D., AND TIEDEMAN, J. Oximetry of retinal vessels by dual-wavelength imaging: calibration and influence of pigmentation. *Journal of Applied Physiology* 86, 2 (1999), 748–758.
- [18] BEACH, J., TIEDEMAN, J., HOPKINS, M., AND SABHARWAL, Y. Multispectral fundus imaging for early detection of diabetic retinopathy. In *Proceedings of International Biomedical Optics Symposium* (1999), International Society for Optics and Photonics, pp. 114–121.
- [19] BERENDSCHOT, T., DELINT, P., AND VAN NORREN, D. Fundus reflectance—historical and present ideas. *Progress in Retinal and Eye Research* 22, 2 (2003), 171–200.
- [20] BERENDSCHOT, T., AND VAN NORREN, D. Objective determination of the macular pigment optical density using fundus reflectance spectroscopy. *Archives of Biochemistry and Biophysics* 430, 2 (2004), 149–155.
- [21] BERENDSCHOT, T., WILLEMSE-ASSINK, J., BASTIAANSE, M., JONG, P. D., AND VAN NORREN, D. Macular pigment and melanin in age-related maculopathy in a general population. *Investigative Ophthalmology and Visual Science* 43, 6 (2002), 1928–1932.

- [22] BERENDSCHOT, T. T. J. M., DELINT, P. J., AND VAN NORREN, D. Fundus reflectance – historical and present ideas. *Progress in Retinal and Eye Research* 22 (2003).
- [23] BERNARDES, R., SERRANHO, P., AND LOBO, C. Digital ocular fundus imaging: a review. *Ophthalmologica* 226, 4 (2011), 161–181.
- [24] BEZDEK, J. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, 1981.
- [25] BHATIA, K., ALJABAR, P., BOARDMAN, J., SRINIVASAN, L., MURGASOVA, M., COUNSELL, S., RUTHERFORD, M., HAJNAL, J., EDWARDS, A., AND RUECKERT, D. Groupwise combined segmentation and registration for atlas construction. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention*. Springer, 2007, pp. 532–540.
- [26] BHATIA, K., HAJNAL, J., HAMMERS, A., AND RUECKERT, D. Similarity metrics for groupwise non-rigid registration. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention*. Springer, 2007, pp. 544–552.
- [27] BOCK, R., MEIER, J., MICHELSON, G., NYÚL, L., AND HORNEGGER, J. Classifying glaucoma with image-based features from fundus photographs. *Pattern Recognition* (2007), 355–364.
- [28] BOETTNER, E., AND WOLTER, J. Transmission of the ocular media. *Investigative Ophthalmology and Visual Science* 1, 6 (1962), 776–783.
- [29] BONE, R. A., BRENER, B., AND GIBERT, J. C. Macular pigment, photopigments, and melanin: Distributions in young subjects determined by four-wavelength reflectometry. *Vision Research* 47 (2007), 3259–3268.
- [30] BOUGUET, J.-Y. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/, 2013. Accessed: 2015-01-07.
- [31] BOYKOV, Y., VEKSLER, O., AND ZABIH, R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20, 12 (2001), 1222–1239.
- [32] BREIMAN, L. Random forests. *Machine Learning* 45, 1 (2001), 5–32.
- [33] BROWN, D., KAISER, P., MICHELS, M., SOUBRANE, G., HEIER, J., KIM, R., SY, J., AND SCHNEIDER, S. Ranibizumab versus verteporfin for neovascular age-related macular degeneration. *New England Journal of Medicine* 355, 14 (2006), 1432–1444.
- [34] BUADES, A., COLL, B., AND MOREL, J.-M. A non-local algorithm for image denoising. In *Proceedings of Computer Vision and Pattern Recognition* (San Diego, California, USA, June 2005), vol. 2, IEEE, pp. 60–65.
- [35] BUDAI, A., ODSTRCILIK, J., KOLAR, R., HORNEGGER, J., JAN, J., KUBENA, T., AND MICHELSON, G. A public database for the evaluation of fundus image segmentation algorithms. *Investigative Ophthalmology and Visual Science* 52, 14 (2011), 1345–1345.

- [36] CALCAGNI, A., GIBSON, J. M., STYLES, I. B., CLARIDGE, E., AND ORIHUELA-ESPINA, F. Multispectral retinal image analysis: a novel non-invasive tool for retinal imaging. *Eye* 25, 12 (2011).
- [37] CHANDRAKER, M., AGARWAL, S., KRIEGMAN, D., AND BELONGIE, S. Globally optimal algorithms for stratified autocalibration. *International Journal of Computer Vision* 90, 2 (2010), 236–254.
- [38] CHANWIMALUANG, T., FAN, G., YEN, G., AND FRANSEN, S. 3-d retinal curvature estimation. *IEEE Transactions on Information Technology in Biomedicine* 13, 6 (2009), 997–1005.
- [39] CHERIYAN, J., MENON, H., AND NARAYANANKUTTY, K. 3d reconstruction of human retina from fundus image—a survey. *International Journal on Modern Engineering Research* 2 (2012), 3089–3092.
- [40] CHIU, S. Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems* 2, 3 (1994), 267–278.
- [41] CHOREMIS, J., AND CHOW, D. Use of telemedicine in screening for diabetic retinopathy. *Canadian Journal of Ophthalmology* 38, 7 (2003), 575–579.
- [42] CHOUDHARY, T., BALL, D., FERNANDEZ-RAMOS, J., MCNAUGHT, A., AND HARVEY, A. Assessment of acute mild hypoxia on retinal oxygen saturation using snapshot retinal oximetry. *Investigative Ophthalmology and Visual Science* 54, 12 (2013), 7538–7543.
- [43] CLARIDGE, E. personal communication.
- [44] COHEN, B., AND DINSTEIN, I. New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition* 35, 2 (2002), 455–463.
- [45] CRUM, W., HARTKENS, T., AND HILL, D. Non-rigid image registration: theory and practice. *The British Journal of Radiology* 77 (2004), S140–S153.
- [46] DEGUCHI, K., KAWAMATA, D., MIZUTANI, K., HONTANI, H., AND WAKABAYASHI, K. 3d fundus shape reconstruction and display from stereo fundus images. *IEICE Transactions on Information and Systems* 83, 7 (2000), 1408–1414.
- [47] DEGUCHI, K., NOAMI, J., AND HONTANI, H. 3d fundus pattern reconstruction and display from multiple fundus images. In *Proceedings of International Conference on Pattern Recognition* (2000), vol. 4, IEEE, pp. 94–97.
- [48] DELORI, F. Noninvasive technique for oximetry of blood in retinal vessels. *Applied Optics* 27, 6 (1988), 1113–1125.
- [49] DELORI, F. Spectrophotometer for noninvasive measurement of intrinsic fluorescence and reflectance of the ocular fundus. *Applied Optics* 33, 31 (1994), 7439–7452.
- [50] DELORI, F., AND BURNS, S. Fundus reflectance and the measurement of crystalline lens density. *Journal of the Optical Society of America A* 13, 2 (1996), 215–226.

- [51] DELORI, F., AND PFLIBSEN, K. Spectral reflectance of the human ocular fundus. *Applied Optics* 28, 6 (1989), 1061–1077.
- [52] DUDA, R., HART, P., AND STORK, D. *Pattern Classification*. John Wiley & Sons, 2001.
- [53] DUONG, T. Magnetic resonance imaging of the retina: a brief historical and future perspective. *Saudi Journal of Ophthalmology* 25, 2 (2011), 137–143.
- [54] EVERDELL, N., STYLES, I., CALCAGNI, A., GIBSON, J., HEBDEN, J., AND CLARIDGE, E. Multispectral imaging of the ocular fundus using light emitting diode illumination. *Review of Scientific Instruments* 81, 9 (2010), 093706.
- [55] FÁLT, P., HILTUNEN, J., HAUTA-KASARI, M., SORRI, I., KALESNYKIENE, V., PIETILÄ, J., AND UUSITALO, H. Spectral imaging of the human retina and computationally determined optimal illuminants for diabetic retinopathy lesion detection. *Journal of Imaging Science and Technology* 55, 3 (2011), 030509–1–030509–10.
- [56] FÁLT, P., HILTUNEN, J., HAUTA-KASARI, M., SORRI, I., KALESNYKIENE, V., AND UUSITALO, H. Extending diabetic retinopathy imaging from color to spectra. In *Proceedings of Scandinavian Conference on Image Analysis* (Oslo, Norway, June 2009), pp. 149–158.
- [57] FÁLT, P., YAMAGUCHI, M., MURAKAMI, Y., LAAKSONEN, L., LENSU, L., CLARIDGE, E., HAUTA-KASARI, M., AND UUSITALO, H. Multichannel spectral image enhancement for visualizing diabetic retinopathy lesions. In *Image and Signal Processing*. Springer, 2014, pp. 52–60.
- [58] FAUGERAS, O., LUONG, Q.-T., AND MAYBANKS, S. Camera self-calibration: Theory and experiments. In *Proceedings of European Conference on Computer Vision* (Santa Margherita Ligure, Italy, May 1992), pp. 321–334.
- [59] FAUST, O., ACHARYA, R., NG, E., NG, K.-H., AND SURI, J. Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review. *Journal of Medical Systems* 36, 1 (2012), 145–147.
- [60] FAWZI, A., LEE, N., ACTON, J., LAINE, A., AND SMITH, R. Recovery of macular pigment spectrum in vivo using hyperspectral image analysis. *Journal of Biomedical Optics* 16, 10 (2011), 106008–106008.
- [61] FERRARA, N., DAMICO, L., SHAMS, N., LOWMAN, H., AND KIM, R. Development of ranibizumab, an anti-vascular endothelial growth factor antigen binding fragment, as therapy for neovascular age-related macular degeneration. *Retina* 26, 8 (2006), 859–870.
- [62] FISCHLER, M., AND BOLLES, R. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24, 6 (1981), 381–395.
- [63] FORACCHIA, M., GRISAN, E., AND RUGGERI, A. Luminosity and contrast normalization in retinal images. *Medical Image Analysis* 9, 3 (2005), 179–190.

- [64] FRAZ, M., REMAGNINO, P., HOPPE, A., UYYANONVARA, B., RUDNICKA, A., OWEN, C., AND BARMAN, S. Blood vessel segmentation methodologies in retinal images – a survey. *Computer Methods and Programs in Biomedicine* 108, 1 (2012), 407–433.
- [65] FRICK, K., GOWER, E., KEMPEN, J., AND WOLFF, J. Economic impact of visual impairment and blindness in the united states. *Archives of Ophthalmology* 125, 4 (2007), 544–550.
- [66] FUMERO, F., ALAYÓN, S., SANCHEZ, J., SIGUT, J., AND GONZALEZ-HERNANDEZ, M. Rim-one: An open retinal image database for optic nerve evaluation. In *Proceedings of International Symposium on Computer-Based Medical Systems* (Porto, Portugal, June 2011), IEEE, pp. 1–6.
- [67] FUSIELLO, A., BENEDETTI, A., FARENZENA, M., AND BUSTI, A. Globally convergent autocalibration using interval analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 12 (2004), 1633–1638.
- [68] F.W. FITZKE III, F. W. Refractive index of the human corneal epithelium and stroma. *Journal of Refractive Surgery* 11, 2 (1995), 100.
- [69] GAO, L., SMITH, R., AND TKACZYK, T. Snapshot hyperspectral retinal camera with the image mapping spectrometer (ims). *Biomedical Optics Express* 3, 1 (January 2012), 48–54.
- [70] GARCÍA, M., LÓPEZ, M., ÁLVAREZ, D., AND HORNERO, R. Assessment of four neural network based classifiers to automatically detect red lesions in retinal images. *Medical Engineering and Physics* 32, 10 (2010), 1085–1093.
- [71] GARCÍA, M., SÁNCHEZ, C., LÓPEZ, M., ABÁSOLO, D., AND HORNERO, R. Neural network based detection of hard exudates in retinal images. *Computer Methods and Programs in Biomedicine* 93, 1 (2009), 9–19.
- [72] GARDNER, G., KEATING, D., WILLIAMSON, T., AND ELLIOTT, A. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *The British Journal of Ophthalmology* 80, 11 (1996), 940–944.
- [73] GIANCARDO, L., MERIAUDEAU, F., KARNOWSKI, T., JR, K. T., GRISAN, E., FAVARO, P., RUGGERI, A., AND CHAUM, E. Textureless macula swelling detection with multiple retinal fundus images. *IEEE Transactions on Biomedical Engineering* 58, 3 (2011), 795–799.
- [74] GIANCARDO, L., MERIAUDEAU, F., KARNOWSKI, T., LI, Y., GARG, S., TOBIN, K., AND CHAUM, E. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. *Medical Image Analysis* 16, 1 (2012), 216–226.
- [75] GONZALEZ, R., AND WOODS, R. *Digital image processing*, second ed. Prentice-Hall Inc., 2001.

- [76] GRISAN, E., GIANI, A., CESERACCIU, E., AND RUGGERI, A. Model-based illumination correction in retinal images. In *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2006), IEEE, pp. 984–987.
- [77] GROUP, D. R. S. R. Photocoagulation treatment of proliferative diabetic retinopathy: the second report of diabetic retinopathy study findings. *Ophthalmology* 85, 1 (1978), 82–106.
- [78] GROUP, E. D. P. R. Causes and prevalence of visual impairment among adults in the united states. *Archives of Ophthalmology* 122, 4 (2004), 477–485.
- [79] GROUP, E. T. D. R. S. R. Treatment techniques and clinical guidelines for photocoagulation of diabetic macular edema: Early treatment diabetic retinopathy study report number 2. *Ophthalmology* 94, 7 (1987), 761–774.
- [80] GUO, J.-M., LIU, Y.-F., CHANG, C.-H., AND NGUYEN, H.-S. Improved hand tracking system. *IEEE Transactions on Circuits and Systems for Video Technology* 22, 5 (2012), 693–701.
- [81] GUO, Y., YAO, G., LEI, B., AND TAN, J. Monte carlo model for studying the effects of melanin concentrations on retina light absorption. *Journal of the Optical Society of America A* 25, 2 (2008), 304–311.
- [82] HAMM, J., YE, D., VERMA, R., AND DAVATZIKOS, C. Gram: A framework for geodesic registration on anatomical manifolds. *Medical Image Analysis* 14, 5 (2010), 633–642.
- [83] HAMMER, M., ROGGAN, A., SCHWEITZER, D., AND MULLER, G. Optical properties of ocular fundus tissues—an in vitro study using the double-integrating-sphere technique and inverse monte carlo simulation. *Physics in Medicine and Biology* 40, 6 (1995), 963.
- [84] HAMMER, M., AND SCHWEITZER, D. Quantitative reflection spectroscopy at the human ocular fundus. *Physics in Medicine and Biology* 47, 2 (2002), 179–191.
- [85] HAMMER, M., SCHWEITZER, D., THAMM, E., AND KOLB, A. Optical properties of ocular fundus tissues determined by optical coherence tomography. *Optics Communications* 186, 1 (2000), 149–153.
- [86] HARALICK, R., SHANMUGAM, K., AND DINSTEN, I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 6 (1973), 610–621.
- [87] HARDARSON, S., HARRIS, A., KARLSSON, R., HALLDORSSON, G., KAGEMANN, L., RECHTMAN, E., ZOEGA, G., EYSTEINSSON, T., BENEDIKTSSON, J., THORSTEINSSON, A., THORSTEINSSON, A., JENSEN, P., BEACH, J., AND STEFÁNSSON, E. Automatic retinal oximetry. *Investigative Ophthalmology and Visual Science* 47, 11 (2006), 5011–5016.
- [88] HARVEY, A., FLETCHER-HOLMES, D., GORMAN, A., ALTENBACH, K., ARLT, J., AND READ, N. Spectral imaging in a snapshot. In *Proceedings of Biomedical Optics* (Changchun, China, August 2005), International Society for Optics and Photonics, pp. 110–119.

- [89] HASAN, M., JIA, X., ROBLES-KELLY, A., ZHOU, J., AND PICKERING, M. R. Multi-spectral remote sensing image registration via spatial relationship analysis on sift keypoints. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium* (2010), IEEE, pp. 1011–1014.
- [90] HAYREH, S. Segmental nature of the choroidal vasculature. *British Journal of Ophthalmology* 59, 11 (1975), 631–648.
- [91] HEIKKIÄ, J., AND SILVÉN, O. A four-step camera calibration procedure with implicit image correction. In *Proceedings of Computer Vision and Pattern Recognition* (San Juan, Puerto Rico, June 1997), pp. 1106–1112.
- [92] HEMMINKI, V., KÄHÖNEN, M., TUOMISTO, M. T., TURJANMAA, V., AND UUSITALO, H. Determination of retinal blood vessel diameters and arteriovenous ratios in systemic hypertension: Comparison of different calculation formulae. *Graefe's Archive for Clinical and Experimental Ophthalmology* 245, 1 (2007), 8–17.
- [93] HILL, D., BATCHELOR, P., HOLDEN, M., AND HAWKES, D. Medical image registration. *Physics in Medicine and Biology* 46, 3 (2001), R1.
- [94] HIPWELL, J., STRACHAN, F., OLSON, J., MCHARDY, K., SHARP, P., AND FORRESTER, J. Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool. *Diabetic Medicine* 17 (2000), 588–594.
- [95] HIROHARA, Y., OKAWA, Y., MIHASHI, T., YAMAGUCHI, T., NAKAZAWA, N., TSURUGA, Y., AOKI, H., MAEDA, N., UCHIDA, I., AND FUJIKADO, T. Validity of retinal oxygen saturation analysis: Hyperspectral imaging in visible wavelength with fundus camera and liquid crystal wavelength tunable filter. *Optical Review* 14, 3 (2007), 151–158.
- [96] HOOVER, A., AND GOLDBAUM, M. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Transactions on Medical Imaging* 22, 8 (2003), 951–958.
- [97] HOOVER, A., KOUZNETSOVA, V., AND GOLDBAUM, M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging* 19, 3 (2000), 203–210.
- [98] HORECKER, B. The absorption spectra of hemoglobin and its derivatives in the visible and near infra-red regions. *Journal of Biological Chemistry* 148, 1 (1943), 173–183.
- [99] HUANG, D., KAISER, P., LOWDER, C., AND TRABOULSI, E. *Retinal imaging*, first ed. Elsevier Inc., 2006.
- [100] HUNOLD, W., AND MALESSA, P. Spectrophotometric determination of the melanin pigmentation of the human ocular fundus in vivo. *Ophthalmic Research* 6, 5-6 (1974), 355–362.
- [101] HUNT, R. *The Reproduction of Colour*, sixth ed. John Wiley & Sons, 2008.

- [102] JAFAR, H., NANDI, A., AND AL-NUAIMY, W. Detection of exudates in retinal images using pure splitting technique. In *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society* (Buenos Aires, Argentina, August 2010), pp. 6745–6748.
- [103] JANNIN, P., KRUPINSKI, E., AND WARFIELD, S. Validation in medical image processing. *IEEE Transactions on Medical Imaging* 25, 11 (2006), 1405–1409.
- [104] JAVITT, J., CANNER, J., AND SOMMER, A. Cost effectiveness of current approaches to the control of retinopathy in type i diabetics. *Ophthalmology* 96, 2 (1989), 255–264.
- [105] JIA, H., WU, G., WANG, Q., AND SHEN, D. Absorb: Atlas building by self-organized registration and bundling. *NeuroImage* 51, 3 (2010), 1057–1070.
- [106] JOHNSON, W., WILSON, D., FINK, W., HUMAYUN, M., AND BEARMAN, G. Snapshot hyperspectral imaging in ophthalmology. *Journal of Biomedical Optics* 12, 1 (2007), 014036–014036.
- [107] JOHNSON, W. R., WILSON, D. W., FINK, W., HUMAYUN, M., AND BEARMAN, G. Snapshot hyperspectral imaging in ophthalmology. *Journal Biomedical Optics* 12, 1 (2007), 014036.
- [108] JOSHI, S., DAVIS, B., JOMIER, M., AND GERIG, G. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage* 23 (2004), S151–S160.
- [109] KAHN, H., AND HILLER, R. Blindness caused by diabetic retinopathy. *American Journal of Ophthalmology* 78, 1 (1974), 58–67.
- [110] KAKUMANU, P., MAKROGIANNIS, S., AND BOURBAKIS, N. A survey of skin-color modeling and detection methods. *Pattern Recognition* 40, 3 (2007), 1106–1122.
- [111] KÄMÄRÄINEN, J. Mvprmatlab. <http://www.optovue.com/ifusion://bitbucket.org/kamarain/mvprmatlab>, 2015. Accessed: 2016-02-23.
- [112] KANDE, G., SUBBAIAH, V., AND SAVITHRI, S. Segmentation of exudates and optic disc in retinal images. In *Indian Conference on Computer Vision, Graphics and Image Processing* (Bhubaneswar, India, December 2008), pp. 535–542.
- [113] KANG, S., AND WEISS, R. Can we calibrate a camera using an image of a flat, textureless lambertian surface? In *Proceedings of European Conference on Computer Vision*. Springer, 2000, pp. 640–653.
- [114] KANKANAHALLI, S., BURLINA, P., WOLFSON, Y., FREUND, D., AND BRESSLER, N. Automated classification of severity of age-related macular degeneration from fundus photographs. *Investigative Ophthalmology and Visual Science* 54, 3 (2013), 1789–1796.
- [115] KAUPPI, T. *Eye Fundus Image Analysis for Automatic Detection of Diabetic Retinopathy*. PhD thesis, Lappeenranta University of Technology, 2010.

- [116] KAUPPI, T., KALESNYKIENE, V., KÄMÄRÄINEN, J.-K., LENSU, L., SORRI, I., RANINEN, A., VOUTILAINEN, R., UUSITALO, H., KÄLVIÄINEN, H., AND PIETILÄ, J. The diaretdb1 diabetic retinopathy database and evaluation protocol. In *Proceedings of British Machine Vision Conference* (Warwick, UK, September 2007), pp. 15.1–15.10.
- [117] KAUPPI, T., KÄMÄRÄINEN, J.-K., LENSU, L., KALESNYKIENE, V., SORRI, I., UUSITALO, H., AND KÄLVIÄINEN, H. Constructing benchmark databases and protocols for medical image analysis: Diabetic retinopathy. *Computational and Mathematical Methods in Medicine 2013* (2013), 1–15.
- [118] KAYA, S., WEIGERT, G., PEMP, B., SACU, S., WERKMEISTER, R., DRAGOSTINOFF, N., GARHÖFER, G., SCHMIDT-ERFURTH, U., AND SCHMETTERER, L. Comparison of macular pigment in patients with age-related macular degeneration and healthy control subjects—a study using spectral fundus reflectance. *Acta Ophthalmologica* 90, 5 (2012), e399–e403.
- [119] KEANE, P., AND SADDA, S. Retinal imaging in the twenty-first century: state of the art and future directions. *Ophthalmology* 121, 12 (2014), 2489–2500.
- [120] KHOUBEHI, B., BEACH, J. M., AND KAWANO, H. Hyperspectral imaging for measurement of oxygen saturation in the optic nerve head. *Investigative Ophthalmology and Visual Science* 45 (2004), 1464–1472.
- [121] KILBRIDE, P., ALEXANDER, K., FISHMAN, M., AND FISHMAN, G. Human macular pigment assessed by imaging fundus reflectometry. *Vision Research* 29, 6 (1989), 663–674.
- [122] KLEIN, J.-C., MENARD, M., CAZUGUEL, G., FERNANDEZ-MALOIGNE, C., SCHAEFER, G., GAIN, P., COCHENER, B., MASSIN, P., LAY, B., AND CHAR-TON, B. Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology. <http://messidor.crihan.fr/index-en.php>. Accessed: 2015-09-16.
- [123] KNIGHTON, R. Quantitative reflectometry of the ocular fundus. *IEEE Engineering in Medicine and Biology Magazine* 14, 1 (1995), 43–51.
- [124] KNIGHTON, R., JACOBSON, S., AND ROMAN, M. Specular reflection from the surface of the retina. In *Proceedings of Optoelectronics and Laser Applications in Science and Engineering* (Los Angeles, California, USA, January 1989), International Society for Optics and Photonics, pp. 10–17.
- [125] KOHNER, E., PATEL, V., AND RASSAM, S. Role of blood flow and impaired autoregulation in the pathogenesis of diabetic retinopathy. *Diabetes* 44, 6 (1995), 603–607.
- [126] KOLAR, R., ODSTRČILIK, J., JAN, J., AND HARABIS, V. Illumination correction and contrast equalization in colour fundus images. In *Proceedings of European Signal Processing Conference* (Barcelona, Spain, August 2011), IEEE, pp. 298–302.

- [127] KUBECKA, L., JAN, J., AND KOLAR, R. Retrospective illumination correction of retinal images. *Journal of Biomedical Imaging 2010* (2010), 11.
- [128] LAAKSONEN, L., CLARIDGE, E., FÄLT, P., HAUTA-KASARI, M., UUSITALO, H., AND LENSU, L. Comparison of image registration methods for composing spectral retinal images. *Biomedical Signal Processing and Control*. Submitted for review 6.5.2016.
- [129] LAAKSONEN, L., CLARIDGE, E., FÄLT, P., HAUTA-KASARI, M., UUSITALO, H., AND LENSU, L. Comparison of image registration methods for composing spectral retinal images. In *Proceedings of International Workshop on Ophthalmic Medical Image Analysis* (Boston, Massachusetts, USA, September 2014), pp. 57–64.
- [130] LAAKSONEN, L., HANNUKSELA, A., CLARIDGE, E., FÄLT, P., HAUTA-KASARI, M., LENSU, L., AND UUSITALO, H. Significance and refinement of spatial accuracy of ground truth in the detection of retinal lesions. *Computerized Medical Imaging and Graphics*. Submitted for review 31.1.2016.
- [131] LAAKSONEN, L., HERTTUAINEN, J., UUSITALO, H., AND LENSU, L. Refining coarse manual segmentations with stable probability regions. In *Proceedings of International Workshop on Ophthalmic Medical Image Analysis* (Munich, Germany, October 2015).
- [132] LABSPHERE. Labsphere: Spectralon targets. <http://www.labsphere.com/products/reflectance-standards-and-targets/reflectance-targets/spectralon-targets.aspx>. Accessed: 2015-01-10.
- [133] LALIBERTÉ, F., GAGNON, L., AND SHENG, Y. Registration and fusion of retinal images - an evaluation study. *IEEE Transactions on Medical Imaging* 22, 5 (2003), 661–673.
- [134] LAM, B., AND YAN, H. A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields. *IEEE Transactions on Medical Imaging* 27, 2 (2008), 237–246.
- [135] LASKO, T., BHAGWAT, J., ZOU, K., AND OHNO-MACHADO, L. The use of receiver operating characteristic curves in biomedical informatics. *Journal of Biomedical Informatics* 38, 5 (2005), 404–415.
- [136] LEE, S., ABRÀMOFF, M., AND REINHARDT, J. Feature-based pairwise retinal image registration by radial distortion correction. In *Medical Imaging* (2007), International Society for Optics and Photonics, pp. 651220–651220.
- [137] LEUTENEGGER, S., CHLI, M., AND SIEGWART, R. Brisk: Binary robust invariant scalable keypoints. In *Proceedings of International Conference on Computer Vision* (2011), IEEE, pp. 2548–2555.
- [138] LIANG, Z., WONG, D., LIU, J., CHAN, K., AND WONG, T. Towards automatic detection of age-related macular degeneration in retinal fundus images. In *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society* (2010), IEEE, pp. 4100–4103.

- [139] LIN, Y., AND MEDIONI, G. Retinal image registration from 2d to 3d. In *Proceedings of Computer Vision and Pattern Recognition* (2008), IEEE, pp. 1–8.
- [140] LIU, D., WOOD, N., XU, X., WITT, N., HUGHES, A., AND THOM, S. 3d reconstruction of the retinal arterial tree using subject-specific fundus images. In *Advances in Computational Vision and Medical Image Processing*. Springer, 2009, pp. 187–201.
- [141] LUJAN, B., WANG, F., GREGORI, G., ROSENFELD, P., KNIGHTON, R., PULIAFITO, C., DANIS, R., HUBBARD, L., CHANG, R., BUDENZ, D., SEIDER, M., AND KNIGHT, O. Calibration of fundus images using spectral domain optical coherence tomography. *Ophthalmic Surgery, Lasers and Imaging* 39, 4 (2008), 15–20.
- [142] MABERLEY, D., WALKER, H., KOUSHIK, A., AND CRUESS, A. Screening for diabetic retinopathy in james bay, ontario: a cost-effectiveness analysis. *Canadian Medical Association Journal* 168, 2 (2003), 160–164.
- [143] MAINTZ, J., AND VIERGEVER, M. A survey of medical image registration. *Medical Image Analysis* 2, 1 (1998), 1–36.
- [144] MARKELJ, P., TOMAŽEVIČ, D., LIKAR, B., AND PERNUŠ, F. A review of 3d/2d registration methods for image-guided interventions. *Medical Image Analysis* 16, 3 (2012), 642–661.
- [145] MARTINELLO, M., FAVARO, P., NIETO, G. M., HARVEY, A., GRISAN, E., SCARPA, F., AND RUGGERI, A. 3-d retinal surface inference: Stereo or monocular fundus camera? In *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society* (2007), IEEE, pp. 896–899.
- [146] MARTINEZ-PEREZ, M., AND ESPINOSA-ROMERO, A. 3d reconstruction of retinal blood vessels from two views. In *Proceedings of Indian Conference on Computer Vision, Graphics & Image Processing* (Kolkata, India, December 2004), pp. 258–263.
- [147] MATAS, J., CHUM, O., URBAN, M., AND PAJDLA, T. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing* 22, 10 (2002), 761–767.
- [148] MORA, A., VIEIRA, P., MANIVANNAN, A., AND FONSECA, J. Automated drusen detection in retinal images using analytical modelling algorithms. *Biomedical Engineering Online* 10, 1 (2011), 59.
- [149] MORDANT, D., AL-ABBOUD, I., MUYO, G., GORMAN, A., SALLAM, A., RITCHIE, P., HARVEY, A., AND MCNAUGHT, A. Spectral imaging of the retina. *Eye* 25, 3 (2011), 309–320.
- [150] MORDANT, D., AL-ABBOUD, I., MUYO, G., GORMAN, A., SALLAM, A., RODMELL, P., CROWE, J., MORGAN, S., RITCHIE, P., HARVEY, A., AND MCNAUGHT, A. Validation of human whole blood oximetry, using a hyperspectral fundus camera with a model eye. *Investigative Ophthalmology and Visual Science* 52, 5 (2011), 2851–2859.

- [151] MORTENSEN, E., DENG, H., AND SHAPIRO, L. A sift descriptor with global context. In *Proceedings of Computer Vision and Pattern Recognition* (San Diego, California, USA, June 2005), pp. 184–190.
- [152] MYRONENKO, A., AND SONG, X. Intensity-based image registration by minimizing residual complexity. *IEEE Transactions on Medical Imaging* 29, 11 (2010), 1882–1891.
- [153] MYRONENKO, A., SONG, X., AND SAHN, D. Maximum likelihood motion estimation in 3d echocardiography through non-rigid registration in spherical coordinates. *Functional Imaging and Modeling of the Heart 5528* (2009), 427–436.
- [154] NARASIMHA-IYER, H., CAN, A., ROYSAM, B., STEWART, C., TANENBAUM, H., MAJEROVICS, A., AND SINGH, H. Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy. *IEEE Transactions on Biomedical Engineering* 53, 6 (June 2006), 1084–1098.
- [155] NAYAK, J., ACHARYA, R., BHAT, P., SHETTY, N., AND LIM, T.-C. Automated diagnosis of glaucoma using digital fundus images. *Journal of Medical Systems* 33, 5 (2009), 337–346.
- [156] NIEMEIJER, M., ABRÀMOFF, M., AND GINNEKEN, B. V. Segmentation of the optic disc, macula and vascular arch in fundus photographs. *IEEE Transactions on Medical Imaging* 26, 1 (2007), 116–127.
- [157] NIEMEIJER, M., ABRÀMOFF, M., AND VAN GINNEKEN, B. Fast detection of the optic disc and fovea in color fundus photographs. *Medical Image Analysis* 13, 6 (2009), 859–870.
- [158] NIEMEIJER, M., STAAL, J., VAN GINNEKEN, B., LOOG, M., AND ABRAMOFF, M. Comparative study of retinal vessel segmentation methods on a new publicly available database. In *Proceedings of SPIE Medical Imaging* (2004), International Society for Optics and Photonics, pp. 648–656.
- [159] NIEMEIJER, M., VAN GINNEKEN, B., J, M. C., MIZUTANI, A., QUELLEC, G., SÁNCHEZ, C., ZHANG, B., HORNERO, R., LAMARD, M., MURAMATSU, C., ET AL. Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs. *IEEE Transactions on Medical Imaging* 29, 1 (2010), 185–195.
- [160] NIEMEIJER, M., VAN GINNEKEN, B., RUSSELL, S., AND ABRÀMOFF, M. Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. *Investigative Ophthalmology and Visual Science* 48, 5 (2007), 2260–2267.
- [161] NIEMEIJER, M., VAN GINNEKEN, B., RUSSELL, S., SUTTORP-SCHULTEN, M., AND ABRAMOFF, M. Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for early diagnosis of diabetic retinopathy. *Investigative Ophthalmology and Visual Science* 48, 5 (2007), 2260–2267.

- [162] NIEMEIJER, M., VAN GINNEKEN, B., STAAL, J., SUTTORP-SCHULTEN, M., AND ABRÀMOFF, M. Automatic detection of red lesions in digital color fundus photographs. *IEEE Transactions on Medical Imaging* 24, 5 (2005), 584–592.
- [163] NIEMEIJER, M., VAN GINNEKEN, B., STAAL, J., SUTTORP-SCHULTEN, M., AND ABRÀMOFF, M. Automatic detection of red lesions in digital color fundus photographs. *IEEE Transactions on Medical Imaging* 24, 5 (2005), 584–592.
- [164] NING, J., ZHANG, L., ZHANG, D., AND WU, C. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition* 43, 2 (2010), 445–456.
- [165] NORONHA, K., AND NAYAK, K. A review of fundus image analysis for the automated detection of diabetic retinopathy. *Journal of Medical Imaging and Health Informatics* 2, 3 (2012), 258–265.
- [166] NOURRIT, V., DENNISS, J., MUQIT, M., SCHIESSL, I., FENERTY, C., STANGA, P., AND HENSON, D. High-resolution hyperspectral imaging of the retina with a modified fundus camera. *Journal français d’ophtalmologie* 33, 10 (2010), 686–692.
- [167] ODSTRCILIK, J., KOLAR, R., BUDAI, A., HORNEGGER, J., JAN, J., GAZAREK, J., KUBENA, T., CERNOSEK, P., SVOBODA, O., AND ANGELOPOULOU, E. Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. *IET Image Processing* 7, 4 (2013), 373–383.
- [168] OHKUBO, Y., KISHIKAWA, H., ARAKI, E., MIYATA, T., ISAMI, S., MOTOYOSHI, S., KOJIMA, Y., FURUYOSHI, N., AND SHICHIRI, M. Intensive insulin therapy prevents the progression of diabetic microvascular complications in japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. *Diabetes Research and Clinical Practice* 28, 2 (1995), 103–117.
- [169] OLIVEIRA, F., AND TAVARES, J. Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering* 17, 2 (2014), 73–93.
- [170] OPTICS, E. Edmund optics: Chrome on glass, 3 frequency grid distortion target. <http://www.edmundoptics.com/testing-targets/test-targets/distortion-test-targets/multi-frequency-grid-distortion-targets/46250/>. Accessed: 2015-05-21.
- [171] OPTOVUE. ifusion - sd-oct & digital fundus photography. <http://www.optovue.com/ifusion/>, 2016. Accessed: 2016-02-18.
- [172] OSAREH, A., MIRMEHDI, M., THOMAS, B., AND MARKHAM, R. Automatic recognition of exudative maculopathy using fuzzy c-means clustering and neural networks. In *Proceedings of Medical Image Understanding and Analysis Conference* (Birmingham, United Kindom, July 2001), pp. 49–52.
- [173] OSAREH, A., MIRMEHDI, M., THOMAS, B., AND MARKHAM, R. Automated identification of diabetic retinal exudates in digital colour images. *The British Journal of Ophthalmology* 87, 10 (2003), 1220–1223.

- [174] OSAREH, A., SHADGAR, B., AND MARKHAM, R. A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images. *IEEE Transactions on Information Technology in Biomedicine* 13, 4 (2009), 535–545.
- [175] OTSU, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 1 (1979), 62–66.
- [176] OWEN, C., RUDNICKA, A., MULLEN, R., BARMAN, S., MONEKOSSO, D., WHINCUP, P., NG, J., AND PATERSON, C. Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (caiar) program. *Investigative Ophthalmology and Visual Science* 50, 5 (2009), 2004–2010.
- [177] PATTON, N., ASLAM, T., MACGILLIVRAY, T., DEARY, I., DHILLON, B., EIKELBOOM, R., YOGESAN, K., AND CONSTABLE, I. Retinal image analysis: Concepts, applications and potential. *Progress in Retinal and Eye Research* 25, 1 (2006), 99–127.
- [178] PFINGSTHORN, M., BIRK, A., SCHWERTFEGER, S., BÜLOW, H., AND PATHAK, K. Maximum likelihood mapping with spectral image registration. In *Proceedings of IEEE International Conference on Robotics and Automation* (2010), IEEE, pp. 4282–4287.
- [179] POKORNY, J., SMITH, V., AND LUTZE, M. Aging of the human lens. *Applied Optics* 26, 8 (1987), 1437–1440.
- [180] POLLEFEYS, M., AND GOOL, L. V. A stratified approach to metric self-calibration. In *Proceedings of Computer Vision and Pattern Recognition* (San Juan, Puerto Rico, June 1997), IEEE, pp. 407–412.
- [181] PREECE, S., AND CLARIDGE, E. Monte carlo modelling of the spectral reflectance of the human eye. *Physics in Medicine and Biology* 47, 16 (2002), 2863–2877.
- [182] RAMELLA-ROMAN, J., MATHEWS, S., KANDIMALLA, H., NABILI, A., DUNCAN, D., D’ANNA, S., SHAH, S., AND NGUYEN, Q. Measurement of oxygen saturation in the retina with a spectroscopic sensitive multi aperture camera. *Optics Express* 16, 9 (2008), 6170–6182.
- [183] RANKIN, S., WALMAN, B., BUCKLEY, A., AND DRANCE, S. Color doppler imaging and spectral analysis of the optic nerve vasculature in glaucoma. *American Journal of Ophthalmology* 119, 6 (1995), 685–693.
- [184] RAPANTZIKOS, K., ZERVAKIS, M., AND BALAS, K. Detection and segmentation of drusen deposits on human retina: Potential in the diagnosis of age-related macular degeneration. *Medical Image Analysis* 7, 1 (2003), 95–108.
- [185] RAVISHANKAR, S., JAIN, A., AND MITTAL, A. Automated feature extraction for early detection of diabetic retinopathy in fundus images. In *Proceedings of Computer Vision and Pattern Recognition* (Miami, Florida, USA, June 2009), pp. 210–217.

- [186] RESNIKOFF, S., PASCOLINI, D., ETYA'ALE, D., KOCUR, I., PARARAJASEGARAM, R., POKHARELAL, G., AND MARIOTTI, S. Global data on visual impairment in the year 2002. *Bulletin of the World Health Organization* 82, 11 (2004), 844–851.
- [187] Retinopathy online challenge. <http://webeye.opth.uiowa.edu/ROC/>. Accessed: 2015-09-16.
- [188] REZA, A., ESWARAN, C., AND HATI, S. Automatic tracing of optic disc and exudates from color fundus images using fixed and variable thresholds. *Journal of Medical Systems* 33, 1 (2009), 73–80.
- [189] RICCI, E., AND PERFETTI, R. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions on Medical Imaging* 26, 10 (2007), 1357–1365.
- [190] RIGAMONTI, R., AND LEPETIT, V. Accurate and efficient linear structure segmentation by leveraging ad hoc features with learned filters. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention*. Springer, 2012, pp. 189–197.
- [191] RODMELL, P., CROWE, J., GORMAN, A., HARVEY, A., MUYO, G., MORDANT, D., MCNAUGHT, A., AND MORGAN, S. Light path-length distributions within the retina. *Journal of Biomedical Optics* 19, 3 (2014), 036008–036008.
- [192] ROHEN, J. Anatomie und embryologie. *Augenheilkunde in Klinik und Praxis* 1 (1977), 1.1–1.5.
- [193] ROSENFELD, P., MOSHFEGHI, A., AND PULIAFITO, C. Optical coherence tomography findings after an intravitreal injection of bevacizumab (avastin) for neovascular age-related macular degeneration. *Ophthalmic Surgery, Lasers and Imaging* 36, 4 (2004), 331–335.
- [194] SALYER, D., DENNINGHOFF, K., BEAUDRY, N., BASAVANTHAPPA, S., PARK, R., AND CHIPMAN, R. Diffuse spectral fundus reflectance measured using subretinally placed spectralon. *Journal of Biomedical Optics* 13, 4 (2008), 044004–044004.
- [195] SÁNCHEZ, C., HORNERO, R., LÓPEZ, M., ABOY, M., POZA, J., AND ABÁSOLO, D. A novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis. *Medical Engineering and Physics* 30, 3 (2008), 350–357.
- [196] SÁNCHEZ, C., NIEMEIJER, M., ABRÀMOFF, M., AND VAN GINNEKEN, B. Active learning for an efficient training strategy of computer-aided diagnosis systems: Application to diabetic retinopathy screening. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention*. Springer, 2010, pp. 603–610.
- [197] SAVAGE, G., JOHNSON, C., AND HOWARD, D. A comparison of noninvasive objective and subjective measurements of the optical density of human ocular media. *Optometry & Vision Science* 78, 6 (2001), 386–395.

- [198] SBEH, Z., COHEN, L., MIMOUN, G., AND COSCAS, G. A new approach of geodesic reconstruction for drusen segmentation in eye fundus images. *IEEE Transactions on Medical Imaging* 20, 12 (2001), 1321–1333.
- [199] SCANLON, P. The english national screening programme for sight-threatening diabetic retinopathy. *Journal of Medical Screening* 15, 1 (2008), 1–4.
- [200] SCHWEITZER, D., HAMMER, M., KRAFT, J., THAMM, E., KÖNIGSDÖRFFER, E., AND STROBEL, J. In vivo measurement of the oxygen saturation of retinal vessels in healthy volunteers. *IEEE Transactions on Biomedical Engineering* 46, 12 (1999), 1454–1465.
- [201] SCHWEITZER, D., KLEIN, S., GUENTHER, S., AND HAMMER, M. Early diagnosis of glaucoma by means of fundus reflectometry. *New Trends in Ophthalmology* 7 (1992), 241–241.
- [202] SCHWEITZER, D., LEISTRITZ, L., HAMMER, M., SCIBOR, M., BARTSCH, U., AND STROBEL, J. Calibration-free measurement of the oxygen saturation in human retinal vessels. In *Proceedings of Ophthalmic Technologies V* (San Jose, California, US, February 1995), International Society for Optics and Photonics, pp. 210–218.
- [203] SEKHAR, S., AL-NUAIMY, W., AND NANDI, A. Automated localisation of optic disk and fovea in retinal fundus images. In *Proceedings of European Signal Processing Conference* (2008), IEEE, pp. 1–5.
- [204] SHARP, P., OLSON, J., STRACHAN, F., HIPWELL, J., LUDBROOK, A., O'DONNELL, M., WALLACE, S., GOATMAN, K., GRANT, A., WAUGH, N., MCHARDY, K., AND FORRESTER, J. The value of digital imaging in diabetic retinopathy. *Health Technology Assessment* 7, 30 (2003), 1–119.
- [205] SINTHANAYOTHIN, C., BOYCE, J., COOK, H., AND WILLIAMSON, T. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. *British Journal of Ophthalmology* 83, 8 (1999), 902–910.
- [206] SINTHANAYOTHIN, C., BOYCE, J., WILLIAMSON, T., COOK, H., MENSAH, E., LAL, S., AND USHER, D. Automated detection of diabetic retinopathy on digital fundus images. *Diabetic Medicine* 19 (2002), 105–112.
- [207] SMITH, R., NAGASAKI, T., SPARROW, J., BARBAZETTO, I., KLAVER, C., AND CHAN, J. A method of drusen measurement based on the geometry of fundus reflectance. *Biomedical Engineering Online* 2, 1 (2003), 10.
- [208] SOARES, J., LEANDRO, J., JR, R. C., JELINEK, H., AND CREE, M. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging* 25, 9 (2006), 1214–1222.
- [209] SOILLE, P. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag New York, Inc., 1999.
- [210] SOMMER, C., STRAEHLE, C., KOTHE, U., AND HAMPRECHT, F. ilastik: Interactive learning and segmentation toolkit. In *Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2011), IEEE, pp. 230–233.

- [211] SOPHARAK, A., NEW, K., MOE, Y., DAILEY, M., AND UYYANONVARA, B. Automatic exudate detection with a naive bayes classifier. In *International Conference on Embedded Systems and Intelligent Technology* (Bangkok, Thailand, February 2008), pp. 139–142.
- [212] SOTIRAS, A., DAVATZIKOS, C., AND PARAGIOS, N. Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging* 32, 7 (2013), 1153–1190.
- [213] STAAL, J., ABRÀMOFF, M., NIEMEIJER, M., VIERGEVER, M., AND VAN GINNEKEN, B. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging* 23, 4 (2004), 501–509.
- [214] STEWART, C., TSAI, C.-L., AND ROYSAM, B. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transactions on Medical Imaging* 22, 11 (2003), 1379–1394.
- [215] STONE, H., AND WOLPOV, R. Blind cross-spectral image registration using pre-filtering and fourier-based translation detection. *IEEE Transactions on Geoscience and Remote Sensing* 40, 32 (2002), 637–650.
- [216] STYLES, I., CALCAGNI, A., CLARIDGE, E., ORIHUELA-ESPINA, F., AND GIBSON, J. Quantitative analysis of multi-spectral fundus images. *Medical Image Analysis* 10, 4 (August 2006), 578–597.
- [217] SWAMINATHAN, R., AND NAYAR, S. Nonmetric calibration of wide-angle lenses and polycameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 10 (2000), 1172–1178.
- [218] TANG, L., KWON, Y., ALWARD, W., GREENLEE, E., LEE, K., GARVIN, M., AND ABRÀMOFF, M. 3d reconstruction of the optic nerve head using stereo fundus images for computer-aided diagnosis of glaucoma. In *SPIE Medical Imaging* (2010), International Society for Optics and Photonics, pp. 76243D–76243D.
- [219] TAYLOR, R. *Handbook of Retinal Screening in Diabetes*, first ed. John Wiley & Sons, 2006.
- [220] TENG, T., LEFLEY, M., AND CLAREMONT, D. Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy. *Medical and Biological Engineering and Computing* 40, 1 (2002), 2–13.
- [221] THIRION, J.-P. Image matching as a diffusion process: an analogy with maxwell’s demons. *Medical Image Analysis* 2.
- [222] TOPCON. 3d oct-2000, optical coherence tomography. <http://www.topcon-medical.eu/eu/products/32-3d-oct-2000-optical-coherence-tomography.html#description>, 2015. Accessed: 2016-02-18.

- [223] TRACKING, R. T. A. B. H. M. donoser and h. bischof. In *Proceedings of International Conference on Pattern Recognition* (Tampa, Florida, December 2008), pp. 1–4.
- [224] TRUCCO, E., RUGGERI, A., KARNOWSKI, T., GIANCARLO, L., CHAUM, E., HUBSCHMAN, J., AL-DIRI, B., CHEUNG, C., WONG, D., ABRAMOFF, M., LIM, G., KUMAR, D., BURLINA, P., BRESSLER, N., JELINEK, H., MAIAUDEAU, F., QUELLEC, G., MACGILLIVRAY, T., AND DHILLON, B. Validating retinal fundus image analysis algorithms: Issues and a proposal. *Investigative Ophthalmology and Visual Science* (2013), 3546–3559.
- [225] TSAI, C.-L., LI, C.-Y., YANG, G., AND LIN, K.-S. The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence. *IEEE Transactions on Medical Imaging* 29, 3 (2010), 636 – 649.
- [226] TWINING, C., COOTES, T., MARSLAND, S., PETROVIC, V., SCHESTOWITZ, R., AND TAYLOR, C. A unified information-theoretic approach to groupwise non-rigid registration and model building. In *Information Processing in Medical Imaging* (2005), Springer, pp. 1–14.
- [227] USHER, D., DUMSKYJ, M., HIMAGA, M., WILLIAMSON, T., NUSSEY, S., AND BOYCE, J. Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. *Diabetic Medicine* 21, 1 (2004), 84–90.
- [228] VAIIOPOULOS, A. Developing matlab scripts for image analysis and quality assessment. In *SPIE Remote Sensing* (2011), International Society for Optics and Photonics, pp. 81810B–81810B.
- [229] VAN DE KRAATS, J., BERENDSCHOT, T., AND VAN NORREN, D. The pathways of light measured in fundus reflectometry. *Vision Reserch* 36, 15 (1996), 2229–2247.
- [230] VAN DER MEER, F. The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *International Journal of Applied Earth Observation and Geoinformation* 8, 1 (2006), 3–17.
- [231] VAN NORREN, D., AND TIEMEIJER, L. Spectral reflectance of the human eye. *Vision Reserch* 26, 2 (1986), 313–320.
- [232] VAN NORREN, D., AND VAN DE KRAATS, J. Retinal densitometer with the size of a fundus camera. *Vision Reserch* 29, 3 (1989), 369–374.
- [233] VERCAUTEREN, T., PENNEC, X., PERCHANT, A., AND AYACHE, N. Non-parametric diffeomorphic image registration with the demons algorithm. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention* (Brisbane, Australia, October 2007), pp. 319–326.
- [234] VERCAUTEREN, T., PENNEC, X., PERCHANT, A., AND AYACHE, N. Symmetric log-domain diffeomorphic registration: a demons-based approach. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention* (New York, New York, USA, September 2008), pp. 754–761.

- [235] VEZHENEVETS, V., AND KONOUCHE, V. Growcut: Interactive multi-label nd image segmentation by cellular automata. In *Proceedings of Graphicon* (Novosibirsk, Russia, June 2005), pp. 150–156.
- [236] VIOLA, P., AND WELLS III, W. Alignment by maximization of mutual information. *International Journal of Computer Vision* 24, 2 (1997), 137–154.
- [237] WALTER, T., KLEIN, J.-C., MASSIN, P., AND ERGINAY, A. A contribution of image processing to the diagnosis of diabetic retinopathy. *IEEE Transactions on Medical Imaging* 21, 10 (2002), 1236–1243.
- [238] WANG, H., HSU, W., GOH, K., AND LEE, M. An effective approach to detect lesions in color retinal images. In *Proceedings of Computer Vision and Pattern Recognition* (Columbus, Ohio, USA, June 2000), vol. 2, IEEE, pp. 181–186.
- [239] WANG, L., JACQUES, S., AND ZHENG, L. Monte carlo modeling of light transport in multi-layered tissues. *Computer Methods and Programs in Biomedicine* 47, 2 (1995), 131–146.
- [240] WANG, Y., JIANG, G., AND JIANG, C. Mean shift tracking with graph cuts based image segmentation. In *Proceedings of IEEE International Congress on Image and Signal Processing* (Chongqing, China, October 2012), pp. 675–679.
- [241] WANGSA-WIRAWAN, N., AND LINSSENMEIER, R. Retinal oxygen: fundamental and clinical aspects. *Archives of Ophthalmology* 121, 4 (2003), 547–557.
- [242] WELFER, D., SCHARCANSKI, J., KITAMURA, C., PIZZOL, M. D., LUDWIG, L., AND MARINHO, D. Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach. *Computers in Biology and Medicine* 40, 2 (February 2010), 124–137.
- [243] WONG, I., IU, L., KOIZUMI, H., AND LAI, W. The inner segment/outer segment junction: what have we learnt so far? *Current Opinion in Ophthalmology* 23, 3 (2012), 210–218.
- [244] WU, G., WANG, Q., JIA, H., AND SHEN, D. Feature-based groupwise registration by hierarchical anatomical correspondence detection. *Human Brain Mapping* 33, 2 (2012), 253–271.
- [245] WYAWAHARE, M., PATIL, P., AND ABHYANKAR, H. Image registration techniques: An overview. *Signal Processing, Image Processing and Pattern Recognition* 2, 3 (2009), 11–28.
- [246] XIAOHUI, Z., AND CHUTATAPE, O. Detection and classification of bright lesions in color fundus images. In *Proceedings of IEEE International Conference on Image Processing* (Singapore, Singapore, October 2004), vol. 1, IEEE, pp. 139–142.
- [247] XU, J., AND CHUTATAPE, O. Comparative study of two calibration methods on fundus camera. In *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society* (2003), vol. 1, IEEE, pp. 576–579.

- [248] XU, L., AND LUO, S. Support vector machine based method for identifying hard exudates in retinal images. In *IEEE Youth Conference on Information, Computing and Images* (Beijing, China, September 2009), pp. 138–141.
- [249] XUE, Z., WONG, K., AND WONG, S. Joint registration and segmentation of serial lung ct images for image-guided lung cancer diagnosis and therapy. *Computerized Medical Imaging and Graphics* 34, 1 (2010), 55–60.
- [250] YANG, G., STEWART, C., SOFKA, M., AND TSAI, C.-L. Registration of challenging image pairs: Initialization, estimation and decision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 11 (2007), 1973 – 1989.
- [251] ZAGERS, N., VAN DE KRAATS, J., BERENDSCHOT, T., AND VAN NORREN, D. Simultaneous measurement of foveal spectral reflectance and cone-photoreceptor directionality. *Applied Optics* 41, 22 (2002), 4686–4696.
- [252] ZHANG, X., AND CHUTATAPE, O. Detection and classification of bright lesions in color fundus images. In *Proceedings of IEEE International Conference on Image Processing* (Singapore, October 2004), pp. 139–142.
- [253] ZHANG, Z. Flexible camera calibration by viewing a plane from unknown orientations. In *Proceedings of International Conference on Computer Vision* (Kerkyra, Greece, September 1999), vol. 1, pp. 666–673.
- [254] ZHANG, Z., YIN, F., LIU, J., WONG, W., TAN, N., LEE, B., CHENG, J., AND WONG, T. Origa-light: An online retinal fundus image database for glaucoma analysis and research. In *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society* (Buenos Aires, Argentina, August 2010), IEEE, pp. 3065–3068.
- [255] ZHENG, J., TIAN, J., DENG, K., DAI, X., ZHANG, X., AND XU, M. Salient feature region: a new method for retinal image registration. *IEEE Transactions on Information Technology in Biomedicine* 15, 2 (2011), 221–232.
- [256] ZITOVA, B., AND FLUSSER, J. Image registration methods: A survey. *Image and Vision Computing* 21, 11 (2003), 977–1000.

Spectral image composition

I.1 Registration method parameters

Table I.1: Method parameters for synthetic test set (RW = regularization weight, SW = similarity weight, C = correlation parameter, UFR = update field regularization, DFR = deformation field regularization)

	Window size	RW	Scaling	Alpha	Corr	Bins
CC	50	0.01	-	-	-	-
CD2	50	0.01	0.01	-	-	-
MI	50	0.01	-	-	-	90
MS	50	0.1	0.01	-	-	-
RC	50	0.1	-	1	0.38	-
SAD	50	0.05	-	-	-	-
SSD	30	0.05	-	-	-	-
	Alpha	UFR	DFR	SW	Max step	
Demons	9	45	-	-	-	
Log-demons	-	45	45	5	5	

Table I.2: Method parameters for semisynthetic test set (RW = regularization weight, SW = similarity weight, C = correlation parameter, UFR = update field regularization, DFR = deformation field regularization)

	Window size	RW	Scaling	Alpha	Corr	Bins
CC	50	0.05	-	-	-	-
CD2	50	0.05	1	-	-	-
MI	15	0.1	-	-	-	15
MS	50	0.1	1	-	-	-
RC	10	0.01	-	1	0.38	-
SAD	10	0.1	-	-	-	-
SSD	30	0.1	-	-	-	-
	Alpha	UFR	DFR	SW	Max step	
Demons	5	20	-	-	-	
Log-demons	-	10	5	10	5	

I.2 Transformation parameter distributions used for sampling

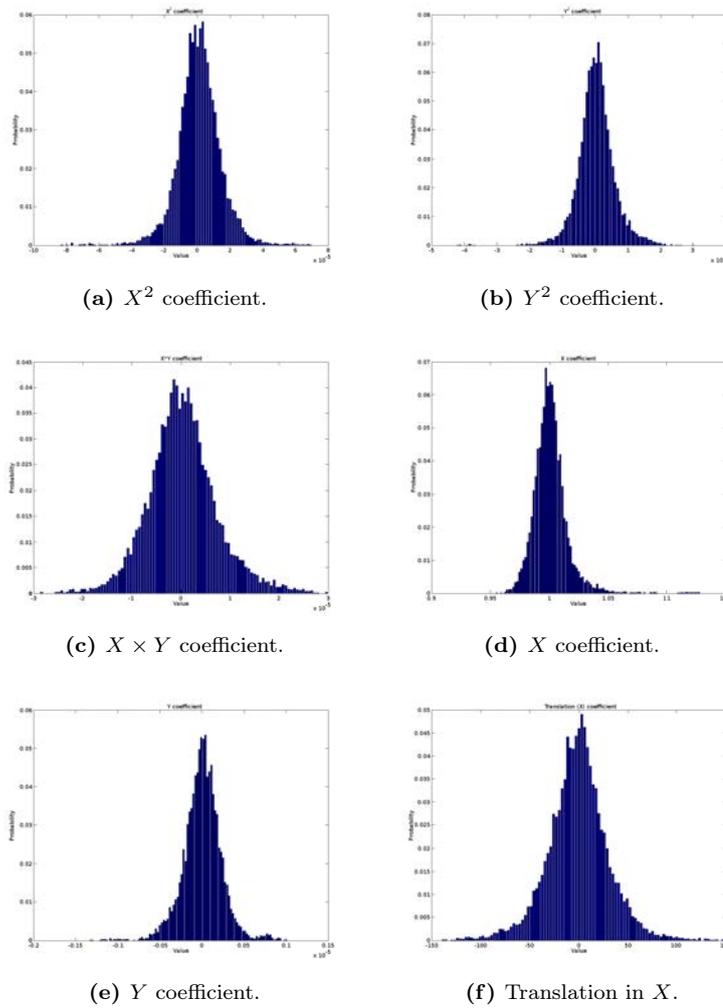


Figure I.1: Distributions for X-coordinate parameters.

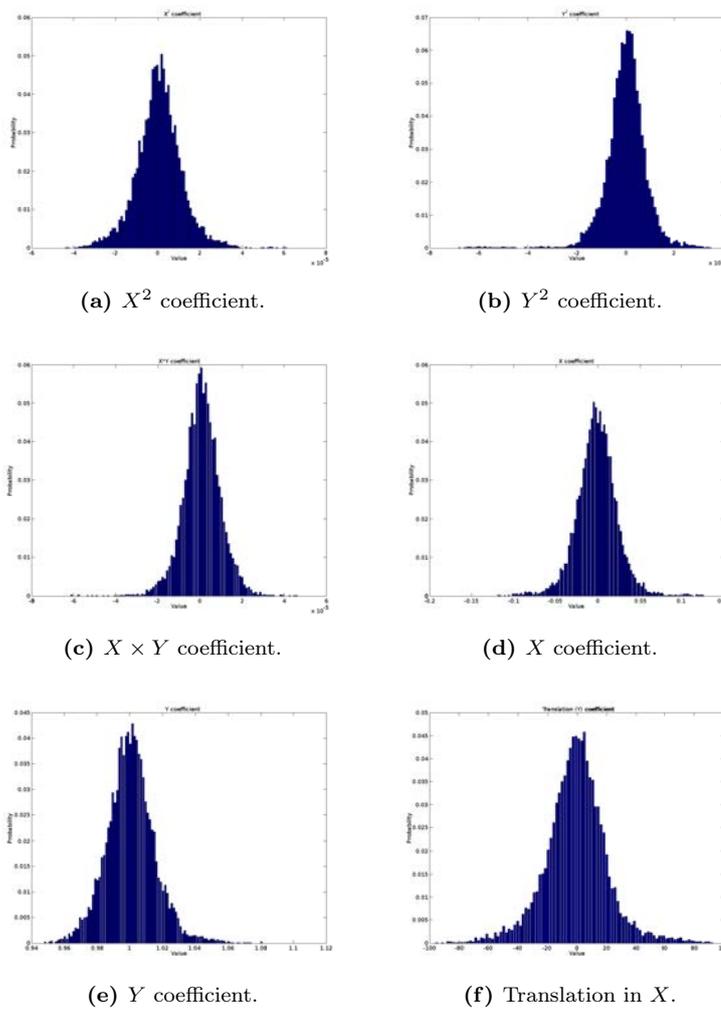


Figure I.2: Distributions for Y-coordinate parameters.

Histological parameter maps from spectral images

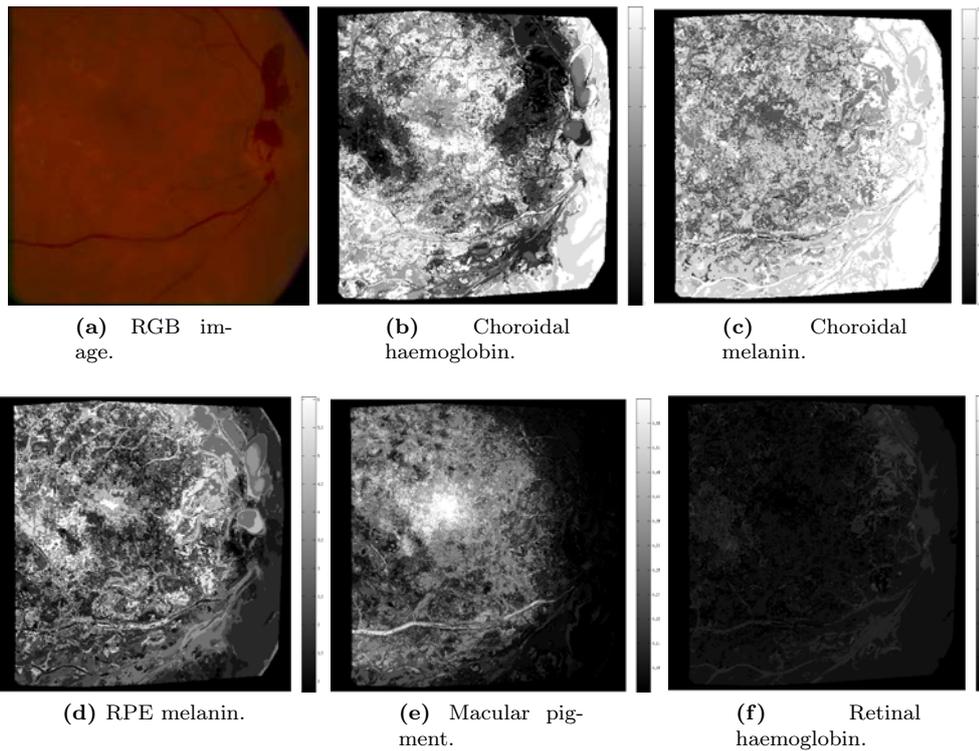


Figure II.1: The parameter maps derived from image 1.

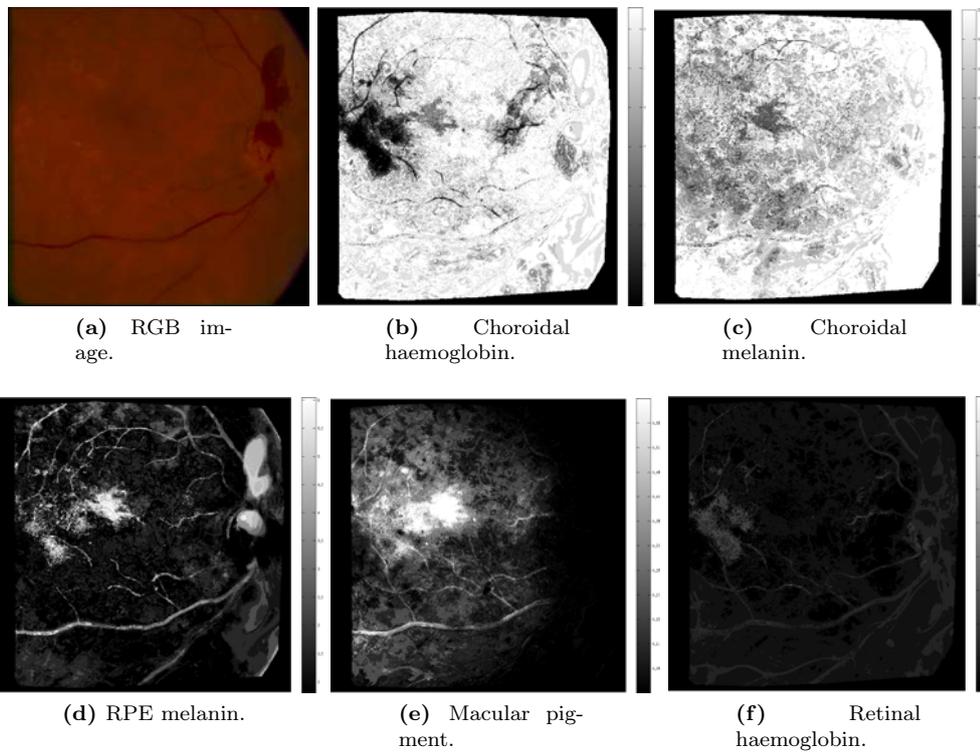


Figure II.2: The parameter maps derived from image 1 after proposed illumination correction.

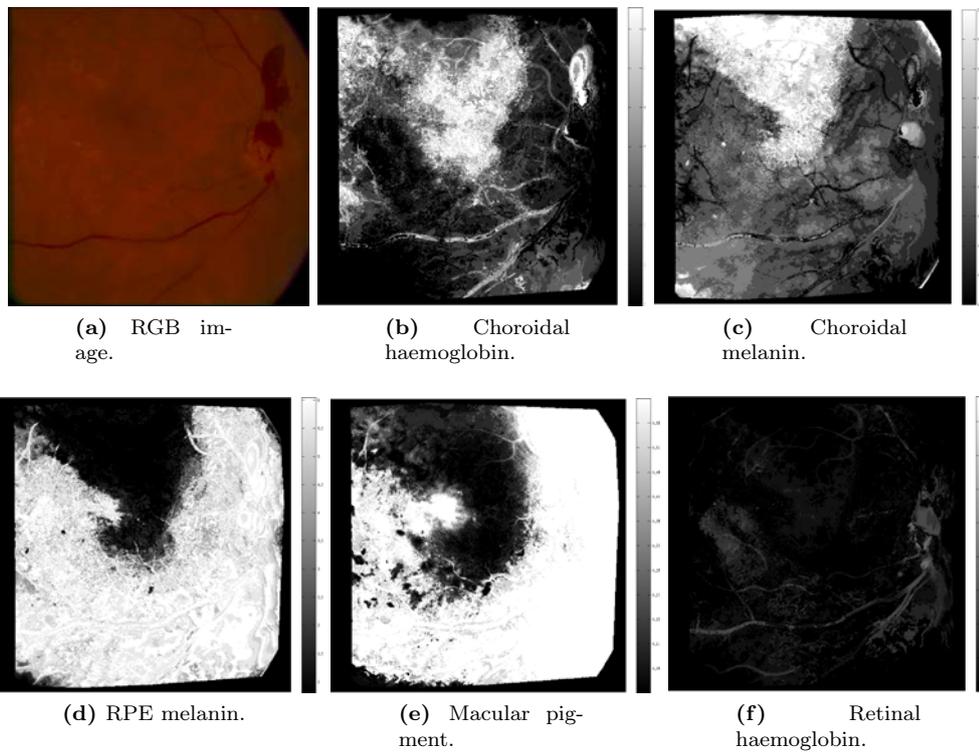


Figure II.3: The parameter maps derived from image 1 after channel-wise illumination correction.

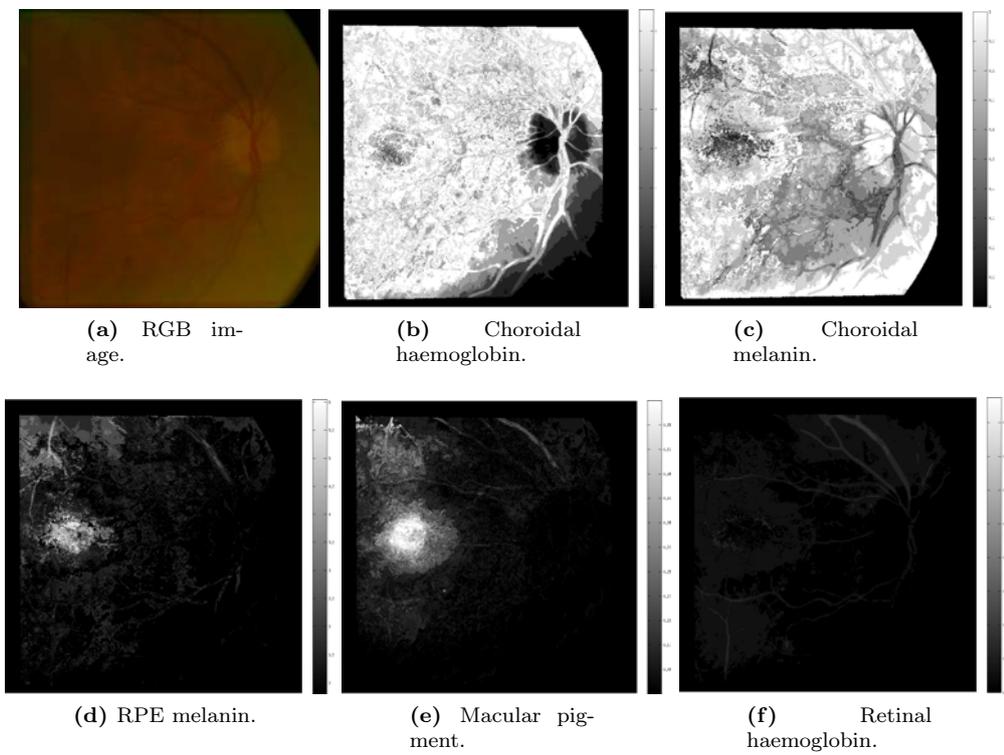


Figure II.4: The parameter maps derived from image 2.

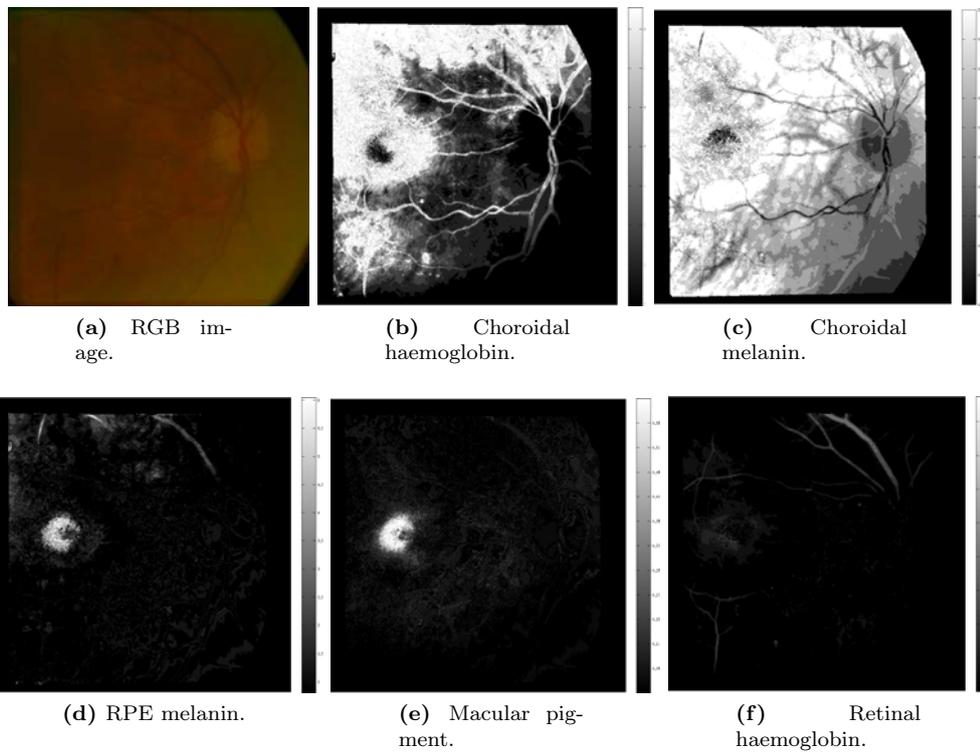


Figure II.5: The parameter maps derived from image 2 after proposed illumination correction.

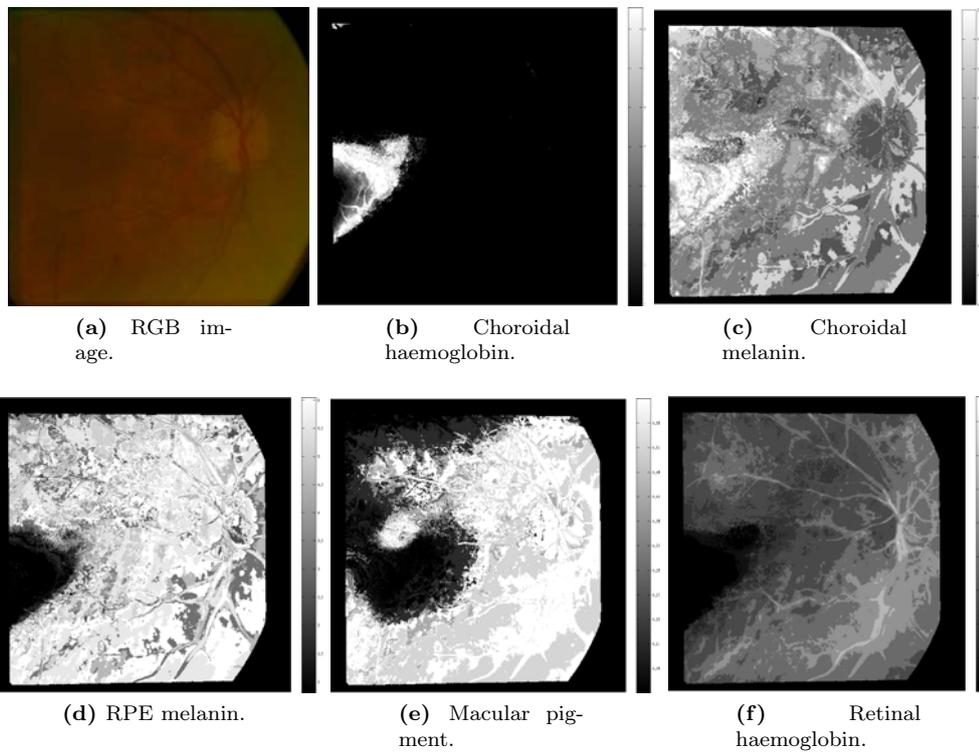


Figure II.6: The parameter maps derived from image 2 after channel-wise illumination correction.

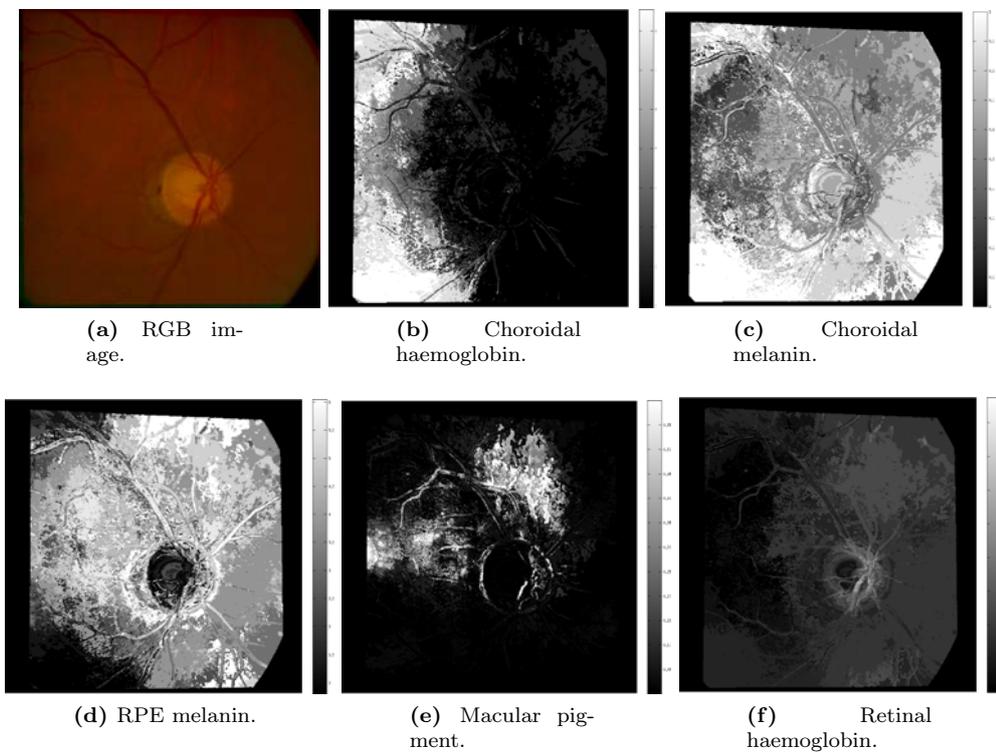


Figure II.7: The parameter maps derived from image 3.

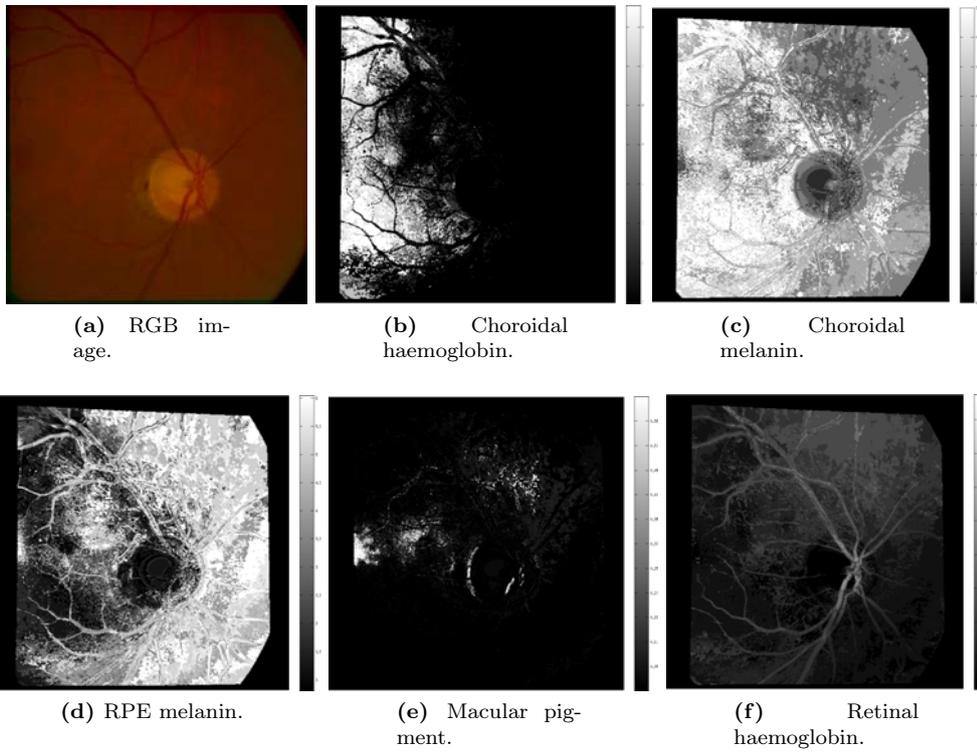


Figure II.8: The parameter maps derived from image 3 after proposed illumination correction.

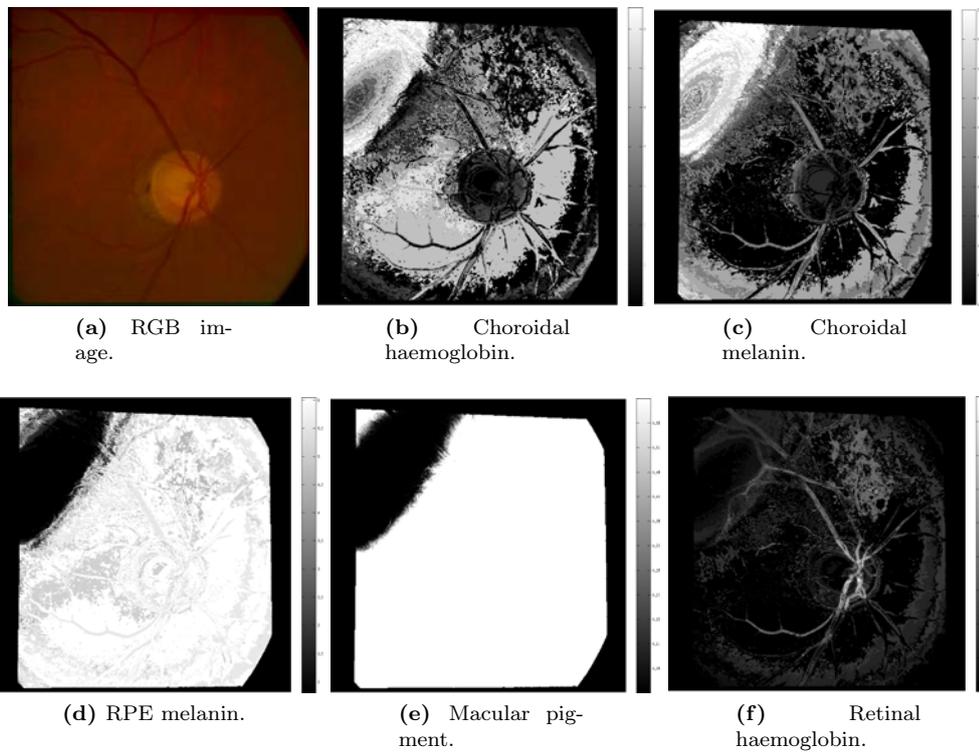


Figure II.9: The parameter maps derived from image 3 after channel-wise illumination correction.

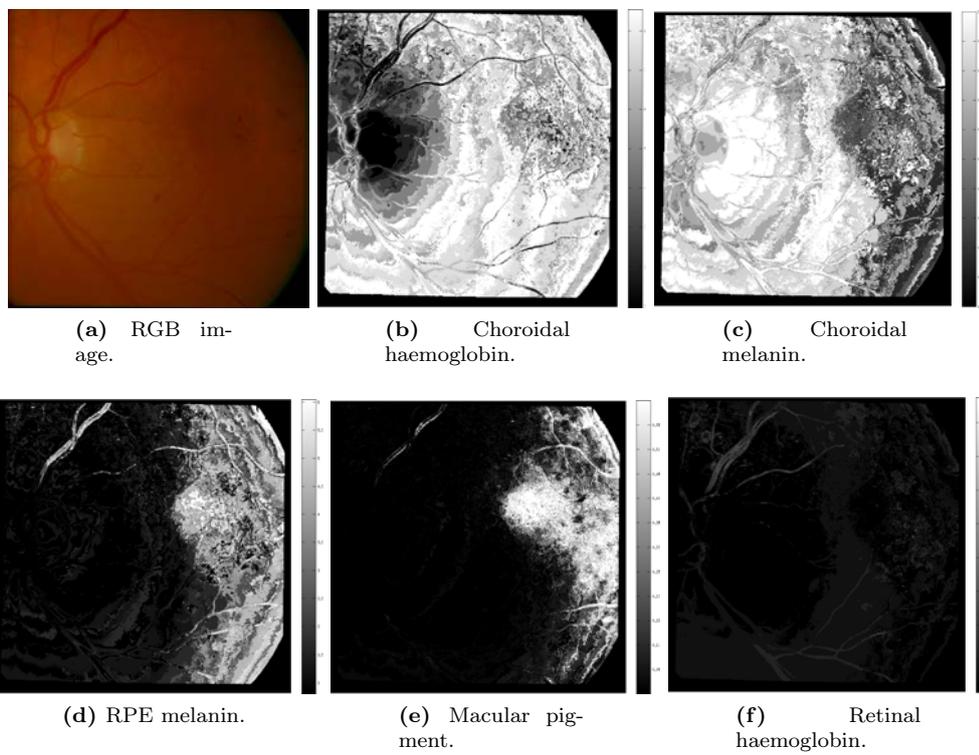


Figure II.10: The parameter maps derived from image 13.

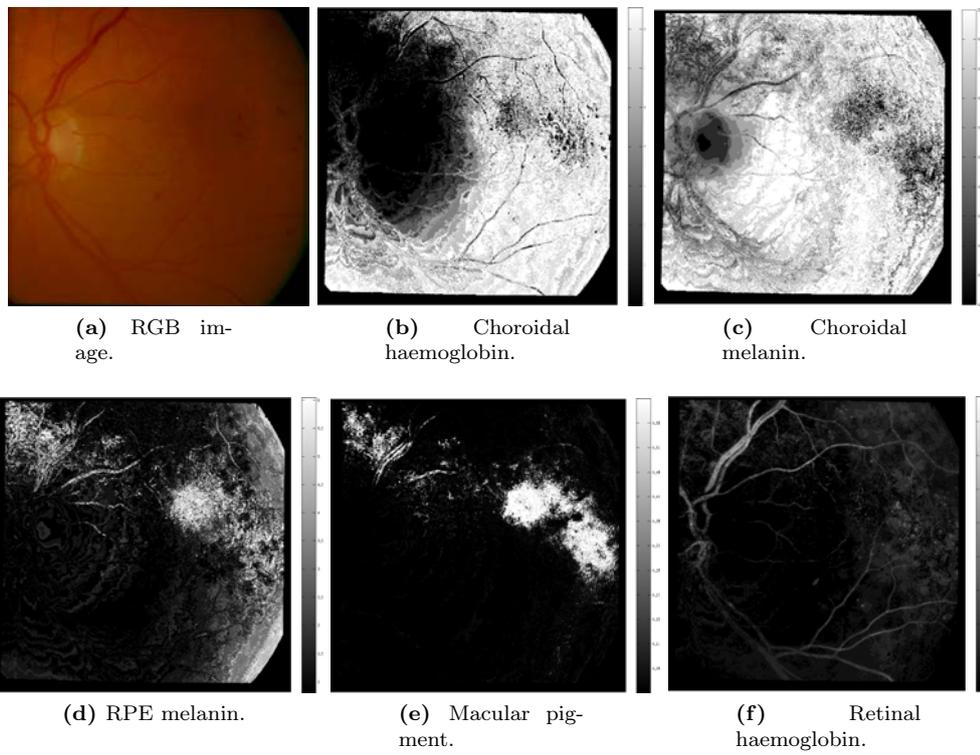


Figure II.11: The parameter maps derived from image 13 after proposed illumination correction.

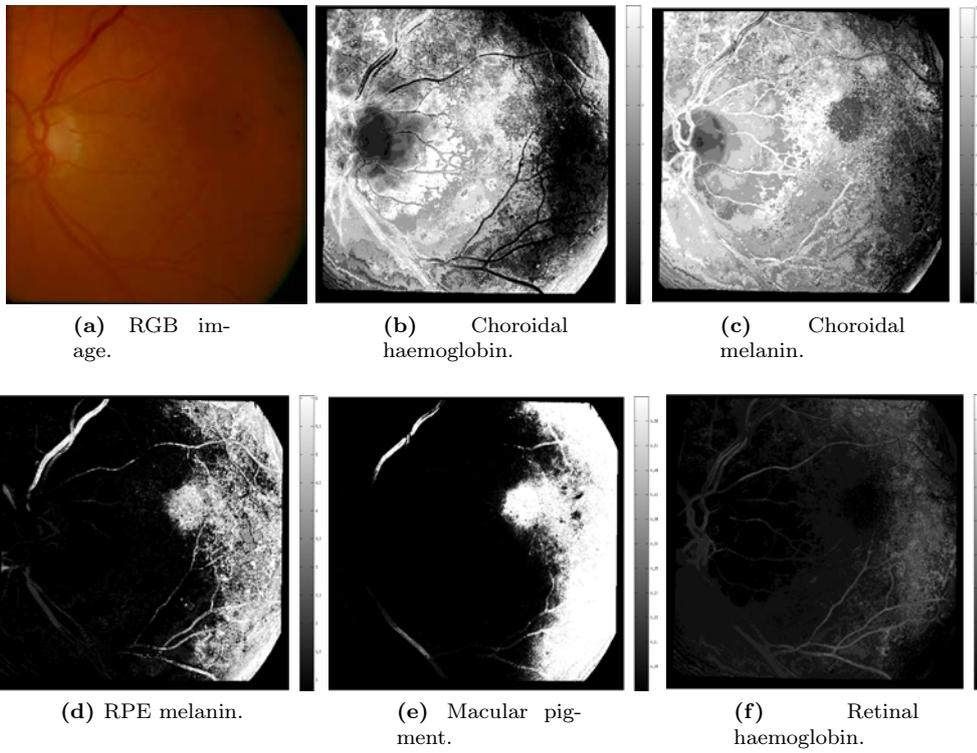


Figure II.12: The parameter maps derived from image 13 after channel-wise illumination correction.

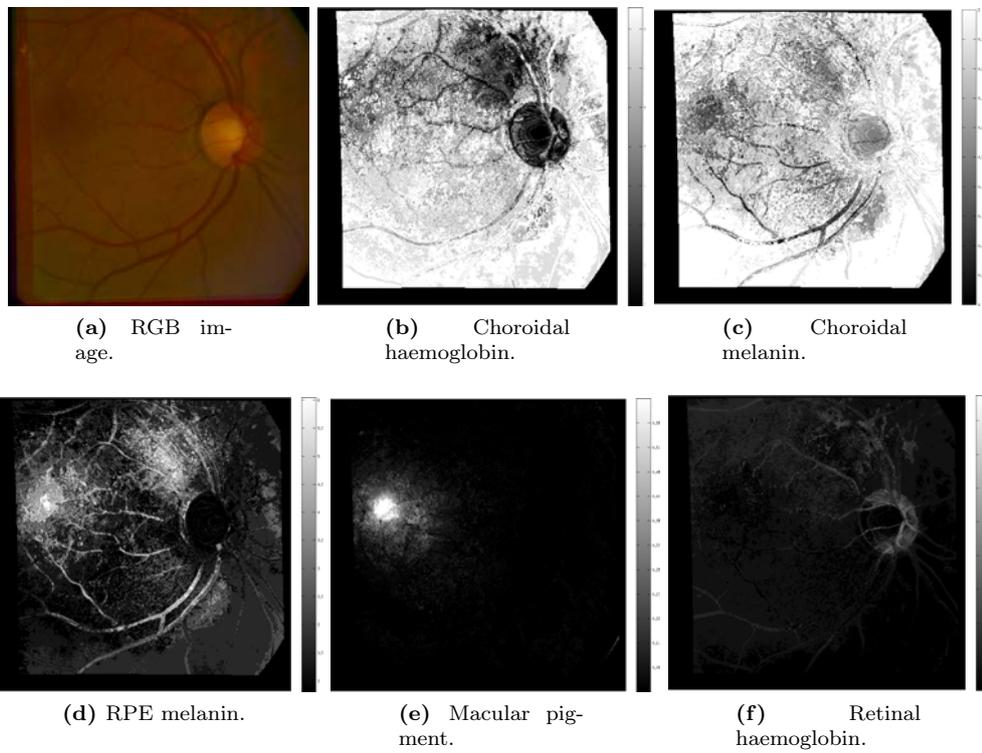


Figure II.13: The parameter maps derived from image 16.

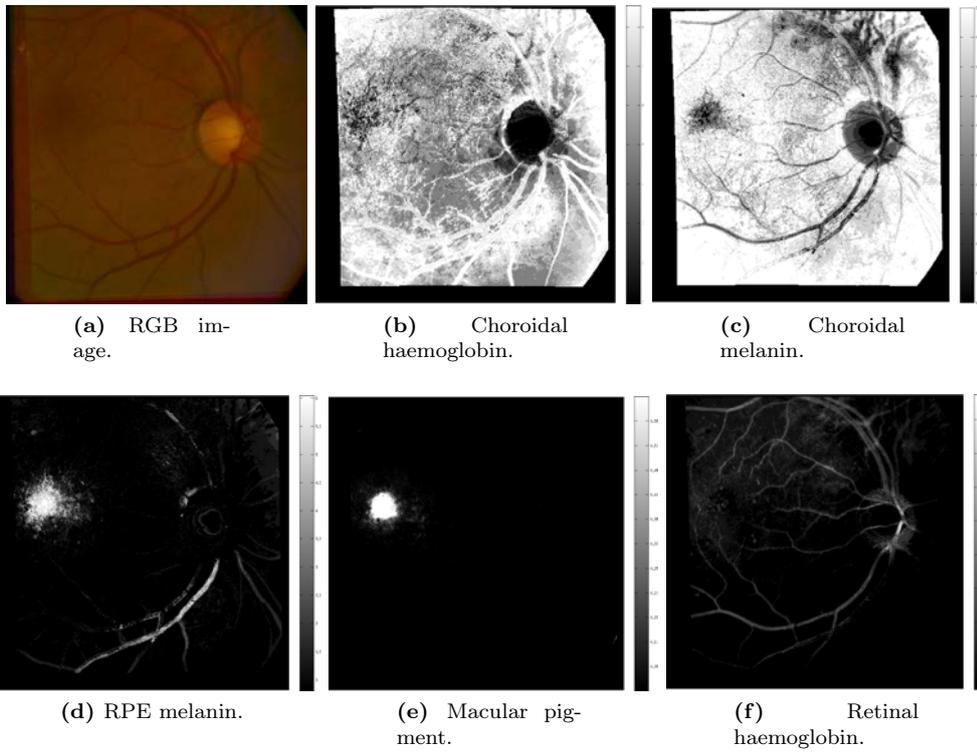


Figure II.14: The parameter maps derived from image 16 after proposed illumination correction.

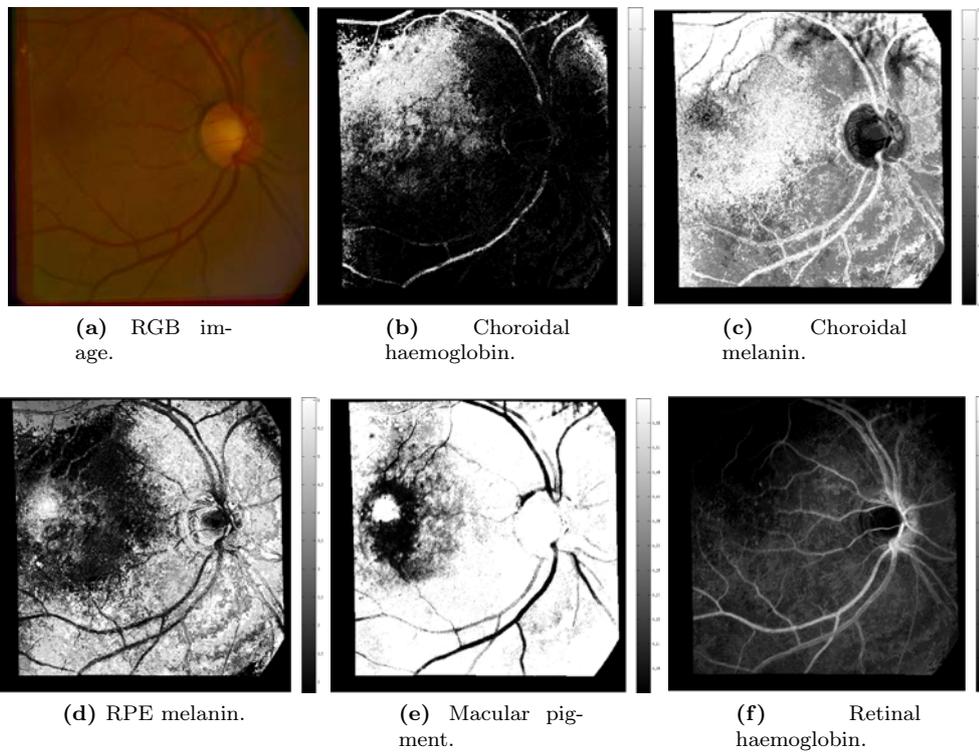


Figure II.15: The parameter maps derived from image 16 after channel-wise illumination correction.

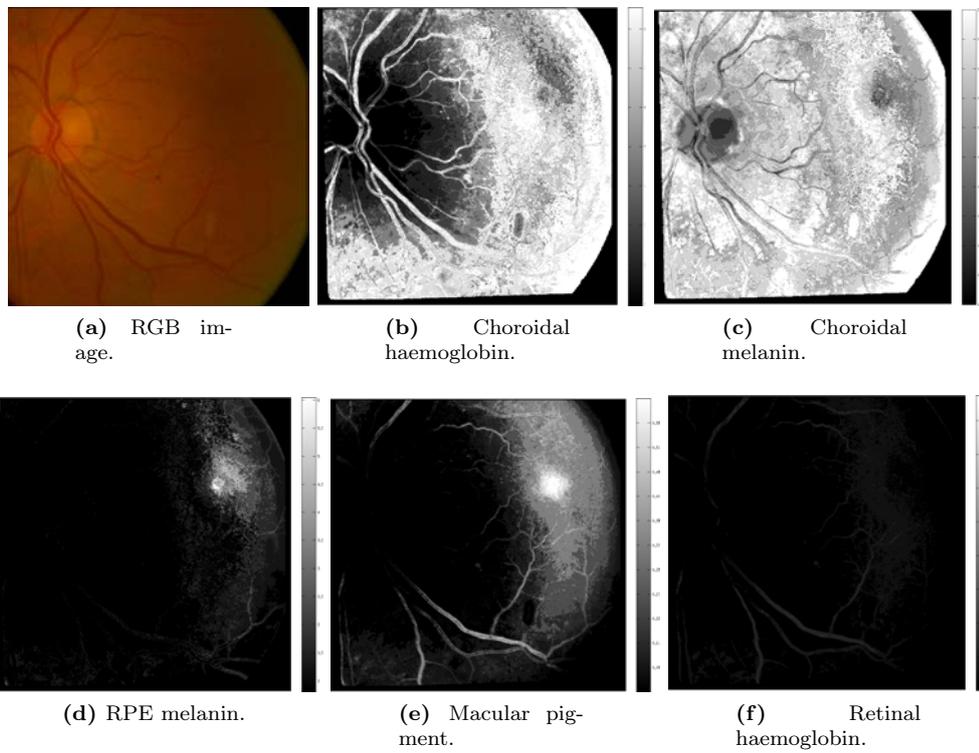


Figure II.16: The parameter maps derived from image 20.

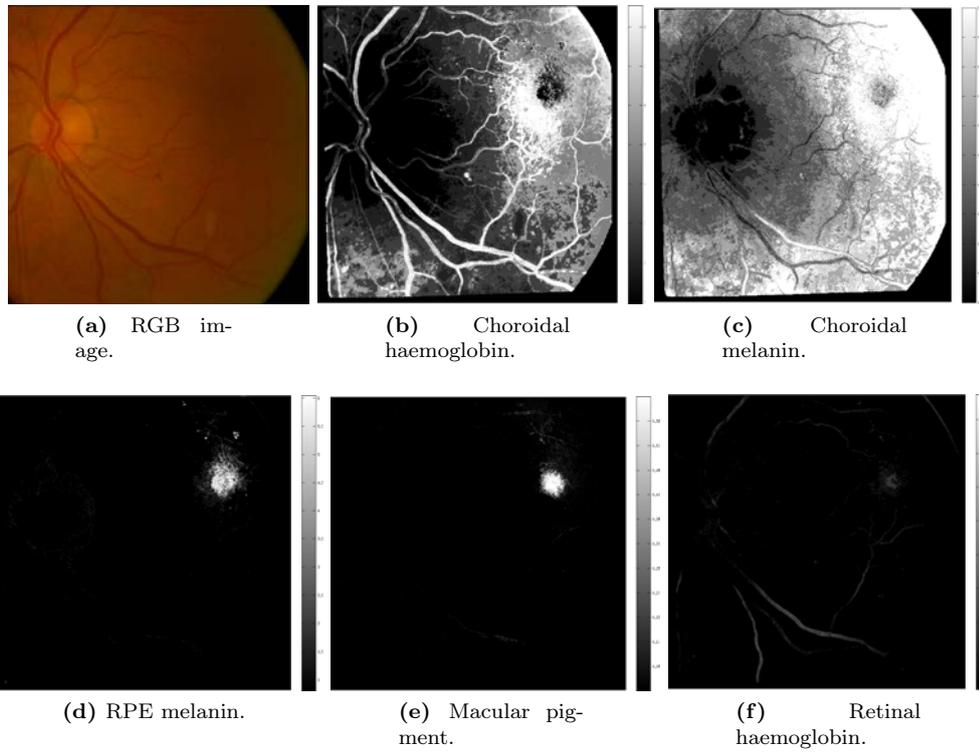


Figure II.17: The parameter maps derived from image 20 after proposed illumination correction.

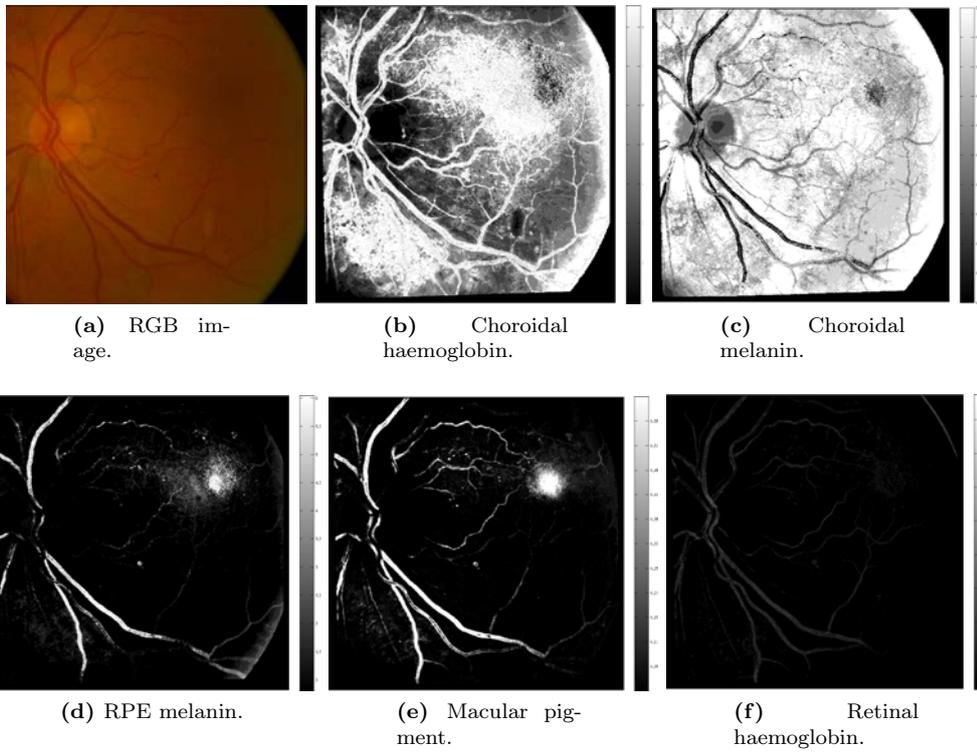


Figure II.18: The parameter maps derived from image 20 after channel-wise illumination correction.

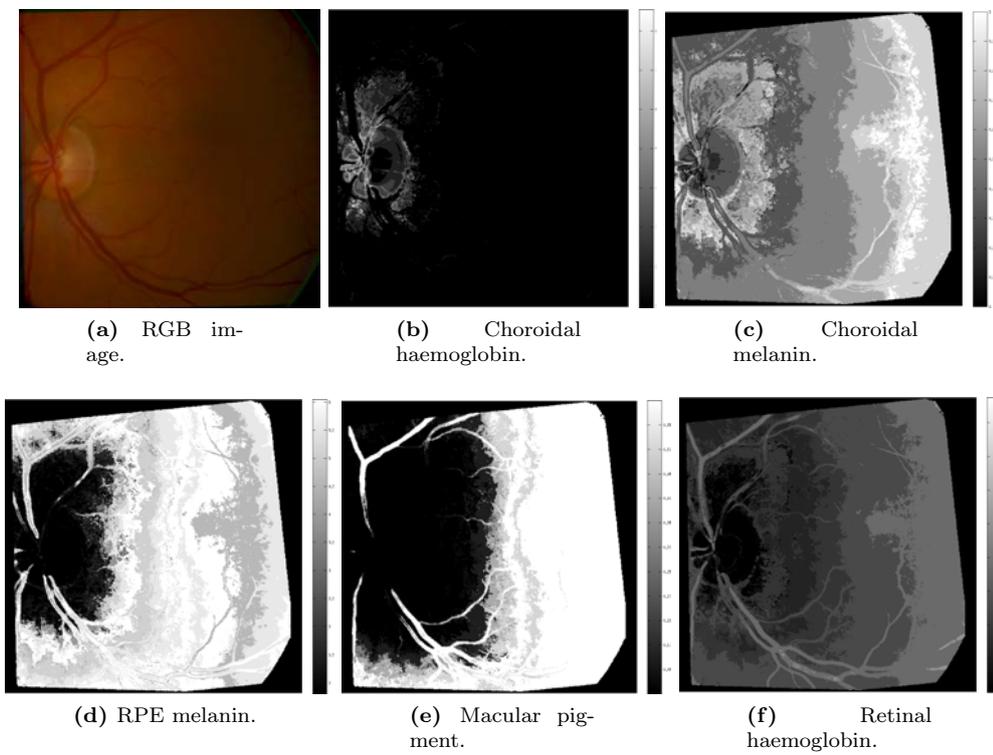


Figure II.19: The parameter maps derived from image 37.

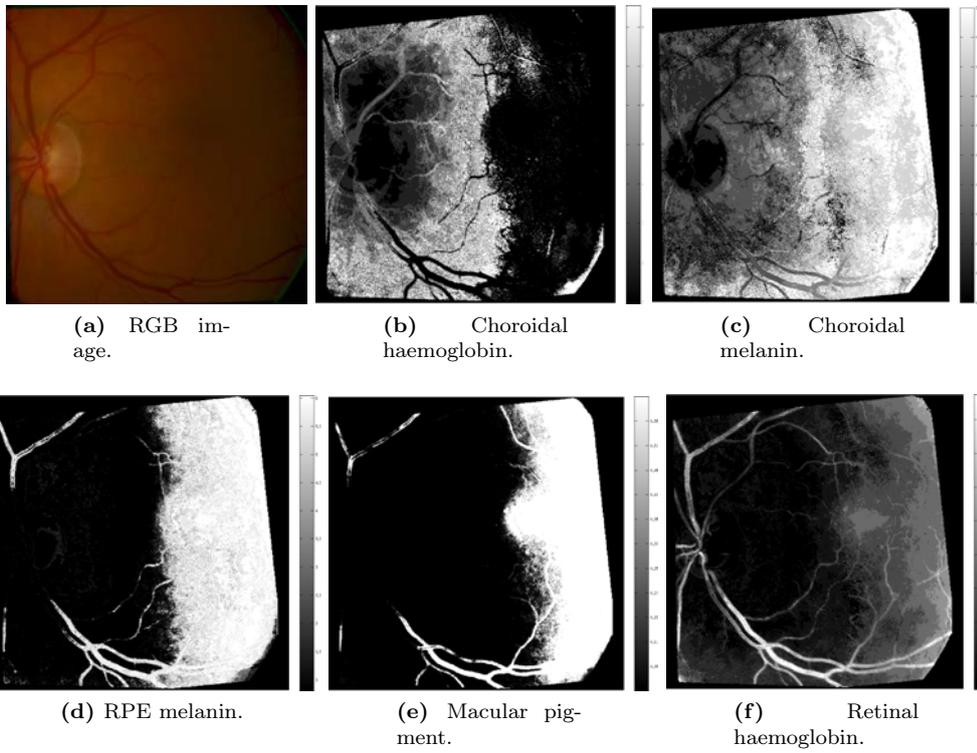


Figure II.20: The parameter maps derived from image 37 after proposed illumination correction.

ACTA UNIVERSITATIS LAPPEENRANTAENSIS

661. RIABCHENKO, EKATERINA. Generative part-based Gabor object detector. 2015. Diss.
662. ALKKIOMÄKI, VILLE. Role of service and data reuse in enterprises. 2015. Diss.
663. VÄNTSI, OLLI. Utilization of recycled mineral wool as filler in wood plastic composites. 2015. Diss.
664. KLEMOLA, KATJA. Tuottavuuden, vaikuttavuuden ja kustannusvaikuttavuuden arviointi alueellisesti integroiduissa sosiaali- ja terveyspalveluissa - palvelujen käyttöön perustuva malli ja esimerkkejä. 2015. Diss.
665. HEROLD, KRISTIINA. Impact of Word-of-Mouth on consumer decision-making: An information processing perspective in the context of a high-involvement service. 2015. Diss.
666. OLABODE, MUYIWA. Weldability of high strength aluminium alloys. 2015. Diss.
667. VANHALA, ERNO. The role of business model in computer game development organizations. 2015. Diss.
668. SALAMPASIS, DIMITRIOS. Trust-embedded open innovation: Towards a human-centric approach in the financial industry. 2015. Diss.
669. DE SMET, DIETER. Innovation ecosystem perspectives on financial services innovation. 2015. Diss.
670. PORRAS, PÄIVI. Utilising student profiles in mathematics course arrangements. 2015. Diss.
671. SALMINEN, JUHO. The role of collective intelligence in crowdsourcing innovations. 2015. Diss.
672. ROSAS, SAILA. Co-operative acquisitions - the contextual factors and challenges for co-operatives when acquiring an investor-owned firm. 2015. Diss.
673. SINKKONEN, TIINA. Item-level life-cycle model for maintenance networks - from cost to additional value. 2015. Diss.
674. TUUNANEN, JUSSI. Modelling of changes in electricity end-use and their impacts on electricity distribution. 2015. Diss.
675. MIELONEN, KATRIINA. The effect of cationic-anionic polyelectrolyte multilayer surface treatment on inkjet ink spreading and print quality. 2015. Diss.
676. OMAJENE, JOSHUA. Underwater remote welding technology for offshore structures. 2015. Diss.
677. NUUTINEN, PASI. Power electronic converters in low-voltage direct current distribution - analysis and implementation. 2015. Diss.
678. RUSATSI, DENIS. Bayesian analysis of SEIR epidemic models. 2015. Diss.
679. STRAND, ELSI. Enhancement of ultrafiltration process by pretreatment in recovery of hemicelluloses from wood extracts. 2016. Diss.

680. TANNINEN, PANU. Press forming of paperboard - advancement of converting tools and process control. 2015. Diss.
681. VALTONEN, PETRI. Distributed energy resources in an electricity retailer's short-term profit optimization. 2015. Diss.
682. FORSSTRÖM-TUOMINEN, HEIDI. Collectiveness within start up-teams - leading the way to initiating and managing collective pursuit of opportunities in organizational contexts. 2015. Diss.
683. MAGUYA, ALMASI. Use of airborne laser scanner data in demanding forest conditions. 2015. Diss.
684. PEIPPO, JUHA. A modified nominal stress method for fatigue assessment of steel plates with thermally cut edges. 2015. Diss.
685. MURASHKO, KIRILL. Thermal modelling of commercial lithium-ion batteries. 2016. Diss.
686. KÄRKKÄINEN, TOMMI. Observations of acoustic emission in power semiconductors. 2016. Diss.
687. KURVINEN, EMIL. Design and simulation of high-speed rotating electrical machinery. 2016. Diss.
688. RANTAMÄKI, JUKKA. Utilization of statistical methods for management in the forest industry. 2016. Diss.
689. PANOVA, YULIA. Public-private partnership investments in dry ports - Russian logistics markets and risks. 2016. Diss.
690. BAHARUDIN, EZRAL. Real-time simulation of multibody systems with applications for working mobile vehicles. 2016. Diss.
691. MARTIKAINEN, SOILI. Development and effect analysis of the Asteri consultative auditing process - safety and security management in educational institutions. 2016. Diss.
692. TORVINEN, PEKKA. Catching up with competitiveness in emerging markets - An analysis of the role of the firm's technology management strategies. 2016. Diss.
693. NORONTAUS, ANNUKKA. Oppisopimuskoulutus yritysten tuottamana koulutuspalveluna: tavoitteista vaikutuksiin. 2016. Diss.
694. HALMINEN, OSKARI. Multibody models for examination of touchdown bearing systems. 2016. Diss.
695. TALONPOIKA, ANNA-MARIA. Financial working capital - management and measurement. 2016. Diss.
696. INKINEN, HENRI. Intellectual capital, knowledge management practices and firm performance. 2016. Diss.
697. YANG, XIAOCHEN. Development of a welding production quality control and management system model for China. 2016. Diss.
698. LEMINEN, VILLE. Leak-proof heat sealing of press-formed paperboard trays. 2016. Diss.

