Comprehensive Retinal Image Analysis:

Image Processing and Feature Extraction Techniques Oriented to the Clinical Task.

Andrés G. Marrugo Hernández

Memoria presentada para optar al grado de doctor por la UPC

María S. Millán García-Varela
Directora

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Terrassa, 2013

Comprehensive Retinal Image Analysis: Image Processing and Feature Extraction Techniques Oriented to the Clinical Task

Andrés G. Marrugo Hernández

Ph.D. in Optical Engineering

Departament d'Òptica i Optometria
Universitat Politècnica de Catalunya

September 26, 2013

Doña María Sagrario Millán y García-Varela, Catedrática de Escuela Universitaria de la Universidad Politécnica de Cataluña

CERTIFICA

que Don Andrés Guillermo Marrugo Hernández, Ingeniero en Mecatrónica, ha realizado bajo su dirección y en el Departamento de Óptica y Optometría de la Universidad Politécnica de Cataluña, el trabajo "Comprehensive Retinal Image Analysis: Image Processing and Feature Extraction Techniques Oriented to the Clinical Task", que se recoge en este compendio de publicaciones y memoria para optar al grado de Doctor por la UPC.

Y para que conste de acuerdo con la legislación vigente, firma este certificado

Dra. María Sagrario Millán y García-Varela

Terrassa, 18 de Julio de 2013

A Nico y Alejo

Agradecimientos/Acknowledgements

Se dice que estamos conectados con cualquier otra persona del planeta por sólo seis grados de separación. Sea cierto o no, la verdad es que no hay construcción humana en que la colaboración no haya sido fundamental para el éxito—incluso en las empresas más abstractas del conocimiento humano el trabajo escrito de los que vinieron antes nos sirve de guía y representan colaboradores distantes, pero nadie *nunca comienza con la hoja en blanco.*

Hoy después de haber dedicado un tiempo importante de mi vida a este proyecto del doctorado estoy obligado a dar las gracias a todos aquellos que me apoyaron a lo largo del camino.

A mi directora de tesis la Dra. María Sagrario Millán por su dedicación y empeño para con este proyecto. Siempre tuvo confianza en mí y con su experiencia supo guiarme. Su motivación para la investigación y la académia me sirven de inspiración para continuar mi camino.

A mi gran amigo el Dr. Héctor Abril con quien pude contar en cualquier momento y para cualquier cosa.

A la Dra. Elisabet Pérez Cabré por su amistad y sus consejos. A las familia Pérez-Cabré y Sisquella-Cabré por su colaboración como sujetos de estudio para los experimentos.

A mis compañeros Curro y Juan Manuel con quienes compartí momentos inolvidables, discusiones productivas y otras intrascendentes. Gracias por todo y especialmente por todos los favores en mi ausencia.

A los doctores Fidel Vega, Jesús Armengol, Jaume Escofet, y demás integrantes del GOAPI. Sus ideas, sus aportes y sus consejos están de una manera u otra plasmados aquí.

Al Dr. Edison Valencia, a quien aún no he conocido en persona, pero cuya tesis marcó el punto de partida de este trabajo.

Al Dr. Gabriel Cristóbal por sus consejos y sus aportes durante el tiempo que estuve en el Instituto de Óptica en Madrid e incluso tiempo después.

Al Dr. Salvador Gabarda por su colaboración en todo lo relacionado con calidad de imagen.

A Isabel Martínez de la óptica Arán por su colaboración en las fases iniciales de esta tesis.

Al Dr. Jordi Monés del institut de la màcula i la retina por facilitarnos imágenes para su análisis.

A Juan Luís Fuentes y el servicio de oftalmología del hospital universitario Miguel Servet por proveernos de imágenes y problemas asociados a la adquisición. Una parte de esta tesis es nuestro itento por atacar esos problemas.

A toda la gente de la escuela de óptica, del departamento de óptica y optometría y el centro universitario de la visión.

To all my Czech friends. My time in Prague was a meaningful experience from which I learnt many important things for both my life and my academic career. Thanks to Dr. Jan Flusser for receiving me at the Institute of Information Theory and Automation (UTIA).

To Dr. Filip Šroubek and Dr. Michal Šorel for your willingness to collaborate with me and work on the puzzling problems surrounding retinal images. Your input and your advice is an important part of this thesis.

To Dr. Barbara Zitova. You received me at your department and you welcomed me at your home. Thank you for your willingness to review this work and for your suggestions.

To Dr. Tomaš Suk. Thanks for sharing your anecdotes and your thoughts in the field of image processing. And to the rest of people at UTIA I value the time we spent together and the things we shared.

To Dr. Peter Schelkens at Vrije Universiteit Brussel for your insight and extensive review of the work presented here.

A mis amigos de Terrassa y Barcelona: Flor María por tu cariño y por todas las atenciones que tuviste para conmigo y mi familia. A Pipe, Nicolás, Julio, Anita, Alexandra, Juan José, Jenny, Fabio, Jorge, Harold Saavedra, Harold Zuluaga, Tomasz, Francisco, Kostas, Míkel, Reza y Ramin.

A mis suegros Marina y Ulpiano que desde la distancia nos apoyaron a mí, a mi mujer y a mis hijos. Que soportaron la tortura de la burocracia y los visados para venir a acompañarnos.

A mis padres que atravesaron el océano Atlántico un sin fin de veces por echarme una mano a mí y a mi familia. Gracias por acompañarme en esos momentos decisivos.

A mis hermanos Maria Emma y Juan Javier por sus palabras de apoyo desde la distancia.

A mi mujer y amiga Lenny Alexandra. Este es un camino que he caminado a tu lado. Sólo tu sabes todo por lo que he pasado. Gracias por tus palabras, por tu entendimiento, por tus consejos, pero por encima de todo, gracias por contagiarme de tu actitud. Tu actitud por la vida, sólo valoras lo importante y siempre desechas lo demás. Nunca te quedas en el pasado, siempre miras más allá.

Gracias.

Financiación

Para la realización de este trabajo se ha contado con la financiación de diferentes instituciones, especialmente:

El Ministerio de Ciencia e Innovación y Fondos FEDER a través de los proyectos de investigación del Plan Nacional de Desarrollo y Producción Industrial DPI2006-05479 y DPI2009-08879.

Red de Procesado de Imagen y Señal Multidimensional (PRISMA).

La Universidad Politécnica de Cataluña por la concesión de una beca predoctoral UPC Recerca a la cual tuve que renunciar poco tiempo después por la concesión de una beca predoctoral para la formación de profesorado universitario (FPU) del Ministerio de Educación.

Contents

List of Figures

List of Tables

Abstract

Medical digital imaging has become a key element of modern health care procedures. It provides a visual documentation, a permanent record for the patients, and most importantly the ability to extract information about many diseases. Ophthalmology is a field that is heavily dependent on the analysis of digital images because they can aid in establishing an early diagnosis even before the first symptoms appear. This dissertation contributes to the digital analysis of such images and the problems that arise along the *imaging pipeline*, a field that is commonly referred to as *retinal image analysis*. We have dealt with and proposed solutions to problems that arise in retinal image acquisition and longitudinal monitoring of retinal disease evolution. Specifically, non-uniform illumination, poor image quality, automated focusing, and multichannel analysis. However, there are many unavoidable situations in which images of poor quality, like blurred retinal images because of aberrations in the eye, are acquired. To address this problem we have proposed two approaches for blind deconvolution of blurred retinal images. In the first approach, we consider the blur to be space-invariant and later in the second approach we extend the work and propose a more general space-variant scheme.

For the development of the algorithms we have built preprocessing solutions that have enabled the extraction of retinal features of medical relevancy, like the segmentation of the optic disc and the detection and visualization of longitudinal structural changes in the retina. Encouraging experimental results carried out on real retinal images coming from the clinical setting demonstrate the applicability of our proposed solutions.

Resumen

La captura de imagen digital es una parte fundamental de los procedimientos médicos modernos. Provee de documentación visual, un registro permanente para los pacientes y más importante la posibilidad de extraer información sobre muchas enfermedades. La oftalmología es un rama de la medicina que depende considerablemente del análisis de imágenes digital. Por medio de la imagen digital se puede establecer un diagnóstico temprano incluso antes que los primeros síntomas comiencen a aparecer. Este trabajo contribuye al análisis digital de tales imágenes y los problemas que surgen a lo largo del proceso de formación de imagen. Este campo se le conoce como *análisis de imagen retiniana*. En esta tesis se han propuesto soluciones a problemas asociados con la adquisición de imagen retiniana y el monitoreo de cambios temporales en patologías retinianas. Específicamente, los problemas de iluminación no-uniforme, baja calidad de imagen, enfoque automático, y análisis multi-canal. Sin embargo, existen situaciones inevitables en que se adquieren imágenes de baja calidad, como imágenes emborronadas debido a las aberraciones del ojo. Este problema lo hemos abordado utilizando dos metodologías para la deconvolución ciega de imágenes emborronadas. En la primera aproximación, consideramos que el emborronamiento es invariante espacialmente y en la segunda aproximación extendimos el trabajo y propusimos un método más general espacialmente variante.

Para el desarrollo de algoritmos hemos construido soluciones de preprocesado que han permitido la extracción de características de relevancia médica, como la segmentación del disco óptico y la detección y visualización de cambios temporales en la retina. Resultados experimentales significativos obtenidos en imágenes reales provenientes de la práctica clínica demuestran la aplicabilidad de las soluciones propuestas.

Part I

Introduction and Justification of the Thematic Unit

Chapter 1

Introduction

> **Imagine the following scene:** *You're living in a small town several hundred kilometers away from the nearest city. You've been having a serious headache for the past couple of weeks. You book an appointment with your local primary care center. The day you go, a family physician gives you a physical examination and doesn't detect anything quite wrong with you, yet you are sent to get your eyes checked. The ophthalmologist comes once a month to town, thus you get checked by the person in charge of fundus examination (a non-physician). The technician takes photographs of the retina in both of your eyes and detects something out of the ordinary: a somewhat enlarged excavation of the optic nerve head. The computer confirms so and also warns of several possible micro-aneurysms and loss of retinal nerve fibers since your last visit. The software automatically sends an alert along with the images to the ophthalmologist to be reviewed on a smart-phone or mobile device. The ophthalmologist confirms the diagnosis suggested by the system in combination with the annotations from the physician and the fundus examiner. The system automatically books you an appointment for Monday morning with the ophthalmologist.*

This story portrays an ideal situation in some distant future, but the truth is that scientifically and technologically speaking we are making great efforts to bring this closer to the present day. It will certainly take time and a great deal of resources before telemedicine and computer-aided diagnosis become ubiquitous and mature enough so that we can rely on them, but the research carried out in the last decade is starting to offer true solutions to the overwhelming problem of health care. This is the ultimate goal of science and technology at the service of medicine.

Ophthalmology is a branch of medicine that is heavily dependent on images. The goal of this thesis is to contribute to the digital analysis of such images and the problems that arise along the *imaging pipeline*; a field

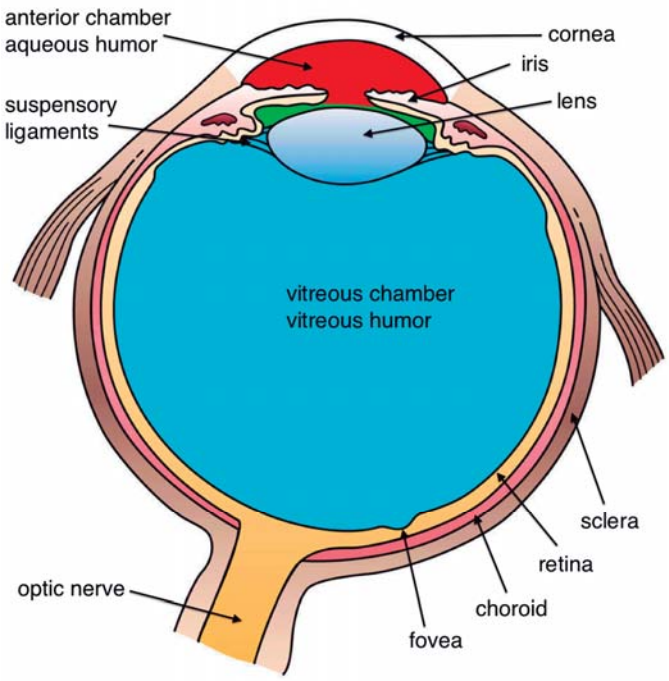


Figure 1.1: Cross section of the human eye (Illustration by Holly Fischer cc).

that is commonly referred to as *retinal image analysis*. In order to provide a kind of showcase of many of the challenges involved in bringing retinal image analysis to the clinical setting, we have deliberately chosen every solution that has been portrayed in the story above so that it can be linked to a reference, whether it be a medical study or a technological development. These references are mentioned in § 1.1.3.

In this chapter we introduce the main background and key concepts regarding the approach followed in this thesis. We present a brief overview of the state of the art in the field of retinal image analysis. A more in-depth discussion with the relevant literature and related works is carried out in every chapter. We have done this so that any chapter could be read as a standalone document.

1.1 Background and Motivation

1.1.1 Fundus Imaging

When alluding to the eye, the posterior segment, or all regions behind the crystalline lens, we refer to it as the ocular fundus (Fig. 1.1). The retina lies at the rear surface of the globe, and in ophthalmology, is an area of primary concern if visual loss is an issue not traceable to refractive error or problems in the cornea or the lens.

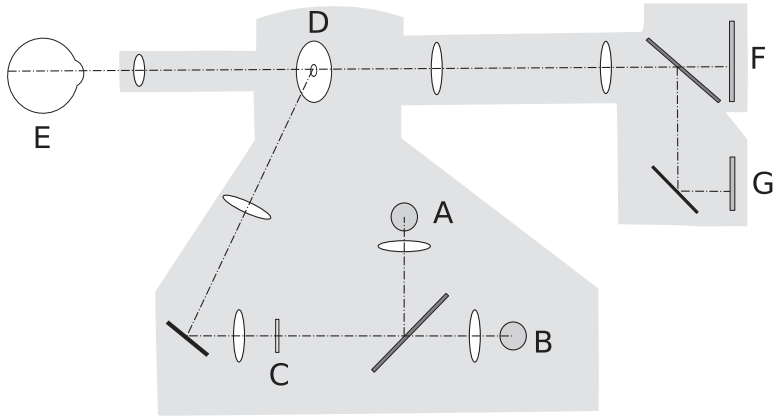


Figure 1.2: The optical pathway of the fundus camera. Light generated from either (A) the viewing lamp or (B) the electronic flash is projected through a (C) set of filters and lenses onto a mirror. This mirror reflects the light into a series of lenses that focus the light. The light is reflected by a (D) round mirror with a central aperture, exits the camera through the objective lens, and proceeds into the (E) eye through the cornea. The back-scattered light from the retina exits the cornea and continues through the central aperture of the previously described mirror, through the diopter compensation lenses, and then reaches the camera system: (F) the sensor and (G) the viewing screen.

The eye fundus has been observed by physicians as early as 1850 with the invention of the ophthalmoscope by the German physician Hermann Von Helmholtz (Saine & Tyler, 2002). This was an instrument that enabled the examination of the retina by using a bright light near the eye and shining it into the patient's pupil. However, it was not until the mid 1920s that the Carl Zeiss Company made available the first commercial fundus camera. Many were the limitations in clinical use of fundus photography in the 1930s and 1940s which can be attributed to the difficulty in obtaining quality images (Saine & Tyler, 2002). Since then, many advances came forth both in quality and specific-purpose developments of techniques for imaging the eye fundus such as: fluorescent angiography, modern electronic fundus photography, stereo fundus photography, confocal laser opthalmoscopy, among others (Abramoff et al., 2010).

All of the modern systems, including the aforementioned and others like optical coherence tomography (OCT) and scanning laser polarimetry, are mostly oriented toward the diagnosis of specific diseases related with certain fundus structures best imaged by a determined technique. In other words, most of these techniques are capable of imaging the eye fundus under a specific modality, and therefore usually work in a complimentary way. However, fundus photography is the most commonly found in medical centers because it provides a more general fundus examination with little patient intervention (Bennett & Barry, 2009).

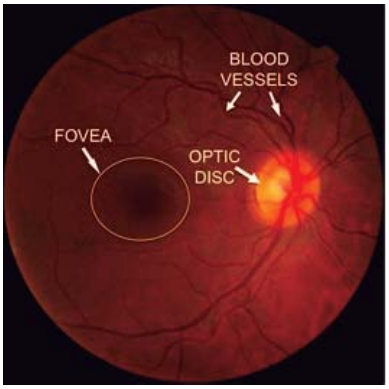


Figure 1.3: A retinal image.

To photograph the retina, it is usually required that the iris be dilated with pharmaceuticals and photographed with a *mydriatic* fundus camera (mydriatic means the iris must be dilated). A fundus camera (Fig. 1.2) is a specialized low-power microscope with an attached camera designed to photograph the interior of the eye in association with the optical system of the eye. Retinal cameras can also be *non-mydriatic* in that the patient's natural dilation in a dark room is used. These cameras use near infrared light to focus and a white light arc lamp to produce a flash that illuminates the retina. The fundus images are captured using a conventional digital camera, attached to the retinal camera body designed to image the eye fundus in association with the optical system of the eye. For further details on fundus imaging the reader is referred to (Bennett & Barry, 2009, Bernardes et al., 2011, Abramoff et al., 2010).

A typical retinal image is shown in Fig. 1.3. The normal features include the optic disc, the fovea, and the blood vessels. The optic disc (or optic nerve head) is the location where ganglion cell axons exit the eye to form the optic nerve (Fig. 1.1). The fovea is the part of the retina responsible for sharp central vision. It has the highest density of photoreceptor cells and approximately 50% of the nerve fibers in the optic nerve carry information from the fovea. The blood vessels or vasculature are the circulatory system that supply blood to the different layers of the retina.

1.1.2 Retinal Manifestations of Eye and Systemic Disease

Many important diseases manifest themselves in the retina and find their origin in the eye, the brain, or the cardiovascular system. There are a number of prevalent diseases that can be studied via eye imaging and image analysis such as the following.

- *Diabetic retinopathy:* Diabetic retinopathy (DR) is a complication of diabetes mellitus and the second most common cause of blindness

and visual loss in the U.S. In the eye, hyperglycemia damages the retinal vessel walls and can lead to the growth of new blood vessels (neo-vascularization), which may bleed and cause retinal detachment (Fig. 1.4). It can also cause diabetic macular edema and damage the photoreceptors because of a breakdown of the blood-retinal barrier (Williams et al., 2004).

- *Age-related macular degeneration:* Age-related macular degeneration (AMD) is the most common cause of visual loss in the the U.S. and is a growing public health problem. The two major forms are dry and wet AMD, of which dry AMD typically leads to gradual loss of visual acuity. Wet AMD, also called choroidal neo-vascularization, is the most visually threatening form. Its natural course is a rapid deteriorating acuity, scarring of the pigment epithelium, and permanent visual loss or blindness.

- *Glaucoma:* Glaucoma is the third leading cause of blindness in the U.S., characterized by gradual damage to the optic nerve and resultant visual field loss. Early diagnose and optimal treatment have been shown to minimize the risk of the visual loss due to glaucoma (Heijl et al., 2002). The hallmark of glaucoma is cupping of the optic disc, which is the visible manifestation of the optic nerve head 3-D structure. The ratio of the optic disc cup and the neuroretinal rim surface areas in these images, called cup-to-disk-ratio, is an important structural indicator for assessing the presence of progression of glaucoma. In § 2.1 we explain further details for optic disc segmentation.

- *Vascular disease:* Cardiovascular disease manifests itself in the retina in a number of ways. Hypertension and atherosclerosis cause changes in the ratio between the diameter of retinal the arteries and veins, known as the A/V ratio. A decrease in the A/V ratio is associated with increased risk of stroke and myocardial infarction (Hubbard et al., 1999).

1.1.3 The Challenges and the Future

We started this chapter by imagining a situation in a somewhat distant future where many of the challenges and limitations that we have today have been significantly worked out. In this subsection we briefly discuss a number of advances that relate directly to the story.

As we have seen there are many diseases that manifest in the retina and as such, examination of the ocular fundus is a key component of the general physical examination and critical to the diagnosis of life- and sight-threatening medical conditions, even among patients with certain presenting

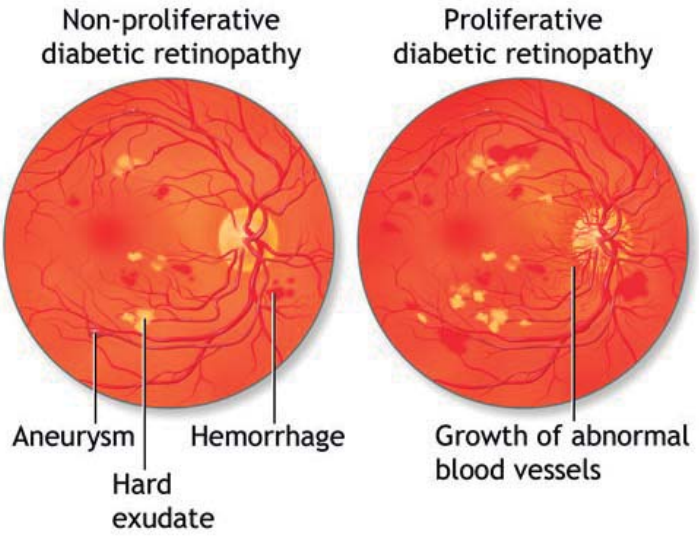


Figure 1.4: Retinal manifestations of disease: Diabetic retinopathy (Illustration by ADAM Images)

complaints such as headache (Sobri et al., 2003, Bruce et al., 2013). However, such examination is infrequently performed in most non-ophthalmic settings (Stern et al., 1995, Bruce et al., 2011). Current technology for examining the ocular fundus, like non-mydriatic fundus photography, enables non-medical personnel to assist in obtaining high-quality images (Maberley et al., 2004).

The everyday cost associated with eye-care providers' decisions and the ever-increasing numbers of retinal images to be reviewed are the major motivations for the adoption of image analysis in ophthalmology (Abramoff et al., 2010). Clearly, since clinicians are costly experts, they need to optimize the time devoted to each patient. Moreover, the development of new imaging technology has resulted in rapidly increasing amounts of data collected as part of any specific retinal imaging exam. This amount of information is already exceeding the limit of clinicians' ability to fully utilize it (Abramoff et al., 2010). If we also take into account that clinicians are subjective, and their decisions suffer from inter- and intra-observer variability, the need for reliable computerized approaches to retinal image analysis is more obvious.

Although recent studies, in automated retinal image analysis, suggest that automated systems may achieve diagnostic capability levels comparable to physicians (Faust et al., 2012), this must be interpreted as an aid for the physician rather than a standalone diagnostic tool. These automated tools may help to alleviate many of the substantial challenges that remain for the widespread access of the general population to screening and prevention programs. In addition, this requires versatile imaging devices and a whole infrastructure built to satisfy the ever increasing demand for high-quality

healthcare services.

In this regard, advancements in telemedicine, particularly via nearly ubiquitous mobile devices, likely hold part of the solution to problems like reviewing images in a timely fashion. For example, Kumar et al. (2012) found that the ophthalmologists who reviewed images of patients for the telemedical diagnosis of DR had very high agreement ($\kappa = 0.9$) and gave high scores to the image quality on the iPhone4. We believe that the true benefit to society is that clinicians will end up devoting most of their time to treating the ill and increase their productivity when dealing with routine population screening. Instead of being replaced by computers (McDonnell, 2010), they will be able to do much *more* and more *accurately*—this is the ultimate goal.

1.2 State of the Art

In this section we give a broad overview of the state of the art in automated retinal image analysis. We discuss the present concerns and several of the main ideas that lead up to the contributions presented herein. A detailed discussion with the specific literature is carried out in each chapter.

1.2.1 Everything Becomes Digital

Although the first reported method for retinal image analysis was published in 1984 (Baudoin et al., 1984), it was not until the 1990s that the field dramatically changed due to the development of digital retinal imaging. Similarly to the general field of image processing, digital retinal images are usually processed in an algorithmic sequence, with the output of one stage forming the input to the next. For example, a typical sequence may consist of one or more preprocessing procedures followed by image segmentation, feature extraction and classification stages.

Alongside the development of digital imaging systems, the increase in computational power and its reduced cost have spurred a significant amount of research in the last decade. The recent reviews by Patton et al. (2006), Winder et al. (2009), Abramoff et al. (2010), Bernardes et al. (2011), and Faust et al. (2012) with nearly over a hundred citations each is evidence of the fervent field. However, recent efforts in the community are shifting from generating algorithms to detect, localize or measure retinal features and properties, validated with small sets of test data, to generating measurements of clinical and public health significance for clinicians, eye care providers and biomedical scientists and researchers, requiring larger and *real life* sets of test data (Trucco et al., in press).

In this thesis we have always pursued the application of our algorithms on real life sets of data chiefly originating from the clinical practice itself and not obtained under controlled experimental conditions. This is a much

more demanding challenge, but it is the one needed at present if automated retinal analysis is to become a reality in the clinical practice.

1.2.2 Present Concerns

The recent paper by Trucco et al. (in press) deals with the present concerns in the field. It combines input from 14 international research groups on the validation of automated retinal image analysis. This work demonstrates the international efforts being put to translating the research findings to the clinical practice in an effective way. The authors described several scenarios in which algorithms for retinal image analysis are being used.

- *Screening/monitoring,* e.g. retinal diseases like DR. The goal is to identify images showing signs of a target condition in large sets. The images (patients) selected are referred for clinical attention. It has been shown that appropriate screening of DR is cost-effective (Abràmoff et al., 2010). Diabetic retinopathy screening facilitates early detection of patients with mild stages of DR, and thus early intervention. Automated screening promises to eliminate inefficiencies within the current DR screening workflow by providing a faster, more cost-effective and accurate disease diagnosis.

- *Computer-assisted diagnosis and risk stratification.* The purpose is to detect the presence or likelihood of a disease from specific signs. Automated retinal image analysis performance must be demonstrated to be more precise than diagnosis in the absence of computer assistance, or generate richer data improving a clinician's diagnosis.

- *Biomarkers discovery.* Aimed to determine whether the occurrence of measurable features in retinal images is linked significantly (in a statistical sense) with specific outcomes or conditions that impact treatment decisions, prognosis, or diagnosis.

In addition, Trucco et al. (in press) also identify three areas that would further benefit from reliable automated retinal image analysis.

- *Longitudinal studies.* The purpose is to provide a means to study quantitatively the evolution and characterization of the disease, to assist treatment planning or gauging patient response to a treatment.

- *Computer-aided or image-guided surgery.* An emergent application of retinal image analysis algorithms (Balicki et al., 2009), e.g. vitreoretinal microsurgery, for which image analysis allows registration of intra-operative imagery with pre-operative imagery for image-guided surgical interventions (Fleming et al., 2008).

- *Tele-health.* Disease screening and monitoring could play a very important role here, e.g. in less developed countries where the incidence of diabetes is rising and screening made difficult by the lack of resources and clinicians (Wild et al., 2004). Computer-assisted telehealth programs can become a scalable method for providing expert care.

1.3 Aim of the thesis

The different scenarios in which retinal image analysis is useful lead to a number of technological and scientific challenges, many of which arise from the attempts to automate or improve a procedure, to decentralize resources, etc. If we regard retinal image analysis from an algorithmic perspective we can identify several areas or stages to which the objectives of this thesis are oriented like the following

1. *Image acquisition.* It is the most important part of the process. The goal is to search for better tools that automate and improve acquisition, e.g. autofocusing in retinal cameras. Previous works and the relevant literature are discussed in § 3.1 and § 4.2.

2. *Image quality assessment.* Assessment of image quality is a general topic of interest, however retinal imaging has its own constraints that merit approaches tailored to the specific requirements. Previous works and the relevant literature are discussed in § 3.2.

3. *Image preprocessing.* The goal is to adapt the images so that subsequent processing is more efficient or accurate. Previous works and the relevant literature to the approaches proposed herein are discussed in § 2.3.1, § 3.3, § 5.4, and § 6.3.1.

4. *Image enhancement/restoration.* Image enhancement or restoration overlap significantly and differ in that the latter is more concerned with accentuation of features, whereas the former with the minimization of the effect of some known degradation. For example deconvolution for deblurring retinal images is a topic extensively addressed in this thesis. Previous works and the relevant literature are discussed in § 5.1 and § 6.1.

5. *Image feature extraction.* It is typically defined as a procedure in which the input is an image and the output are measurements. As we have discussed previously, the segmentation of retinal features (whether normal or pathological) or the detection of longitudinal changes, are the examples of feature extraction addressed in this thesis. Previous works and the relevant literature are discussed in § 2.1 and § 5.1.

This multistage approach or pipeline changes depending on the situation at hand. To that end we have identified a number of problems or opportunities which are the subject of study in this thesis. In some cases it is the improvement of an existing solution, in others proposing an alternate and possibly better solution, and finally others in which our approach is entirely novel. The specific objectives achieved in this thesis are:

- *To determine and validate a robust technique for retinal image quality estimation.* This is a broad enough objective that we have tackled it by first analyzing several state-of-the-art no-reference image quality metrics and their performance in retinal imaging (Chapter 3). This lead to the development of a novel robust focus measure for non-mydriatic retinal imaging (Chapter 4).

- *To determine a suitable method for the compensation of non-uniform illumination in retinal images.* This was achieved both for visualization purposes (single-image compensation in Chapter 3) and for subsequent processing (multi-image compensation in Chapter 5).

- *To develop a strategy for the restoration and enhancement of blurred retinal images.* This was achieved initially by considering the space-invariant case (Chapter 5) in which the blur is modeled as caused by a single PSF. In the next step we considered the space-variant case (Chapter 6) in which we have extended the model to take into account a more general space-variant PSF.

- *To detect and identify longitudinal structural changes in the retina possibly linked to the evolution of a disease.* This objective was achieved as a prerequisite or preprocessing stage for the image enhancement of blurred retinal images by means of multi-image deconvolution (Chapter 5).

- *To design and validate a method for the segmentation of the optic disc.* This objective was the starting point of the thesis as it followed from the previous work carried out in the research group (Chapter 2). It was achieved by implementing and adapting an active contours approach to the considered problem.

1.3.1 Methodology

This work is about the digital analysis of images. In general, we use color images, although color is not critical from a colorimetric point of view. For the most part, we use images from *real* patients thus it is an experimental work. This thesis is oriented to the development of algorithms that provide a solution to a range of problems, rather than performing clinical trials with

a large number of patients and carrying out a statistical comparison with a panel of specialists.

When necessary, we have used a real reference image. We have degraded it, and later enhanced it with the considered algorithm. The result has been compared with the original before the artificial degradation. This enabled the refinement of the algorithm and served as validation.

1.3.2 Originality

The contributions described herein are not limited to previously existing solutions, rather we have developed our own algorithms built upon related approaches found in the literature. In addition, as part of the methodology of the work, the results obtained with our proposed algorithms have been compared against existing algorithms published by other authors (in certain circumstances reproducing and applying them to the images of interest).

The work developed for this thesis is also original in the way the objectives have been laid out. This is a multi-objective work in which we have considered the whole imaging pipeline, ranging from acquisition to feature extraction. To that end we have analyzed the numerous limitations while trying to acquire images and later processing them for feature extraction. This holistic approach is noteworthy in the sense that most works focus on a single stage.

1.3.3 Novelty

The topic of this thesis is novel and is clearly very much of international interest as evidenced by the significant increase of published papers in the field in the last decade. From the conception of this work we have not set to perform incremental developments, but to contribute in a much more substantial way to the different proposed objectives. In fact, this thesis is presented in the modality of compilation of publications, which gives credit to the novelty of its contributions.

1.4 Thesis Outline

This thesis is presented as compilation of publications. It is organized in the following way: Part I, which consists of chapter 1, is the introduction and justification of the thematic unit, Part II, which consists of chapters 2 to 6, is a summary and analysis of the results, and finally Part III is the compilation of publications.

In Fig. 1.5 we show a mind-map of the results presented in this thesis. If you look carefully you will notice that they are organized by the stages that we have described in the previous section. Presenting the results in this way seemed unpractical because each contribution would be isolated from the

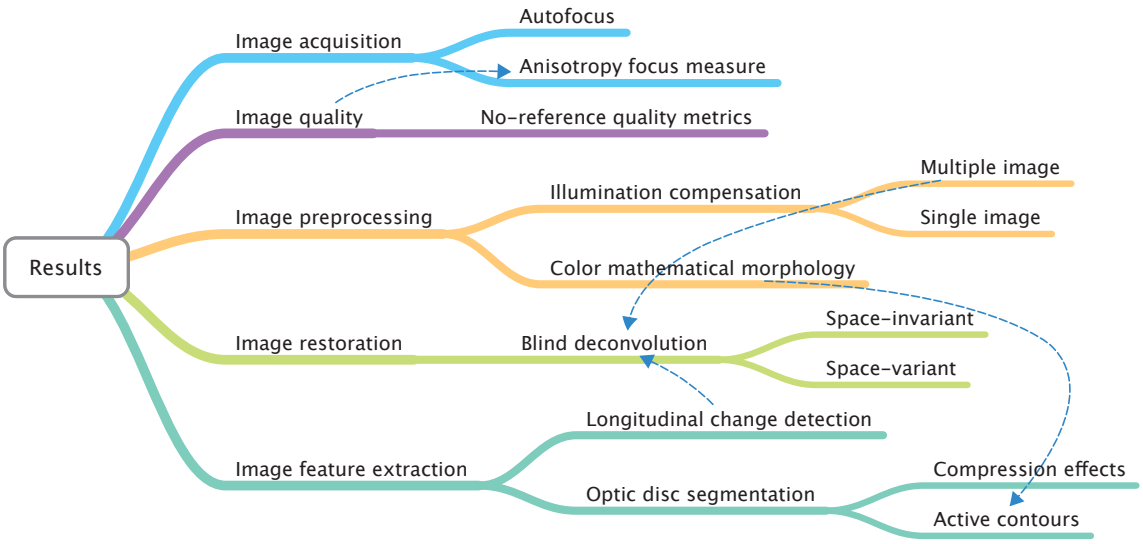


Figure 1.5: A mindmap of the results and contributions of this thesis.

underlying context and the corresponding problem. That is why there are dashed blue arrows connecting different topics. For this reason the summary and analysis of the results (which is the most extensive part) is arranged, to a great extent, in a chronological way. We believe that this will enable the reader to form a better picture of the overarching story of this thesis.

The first two chapters are the basis or preliminary work in which we set out to identify the necessities in the field of retinal image analysis and the goals worth pursuing for research. Unsurprisingly these initial chapters are incremental contributions that we have produced by addressing the challenges and limitations of optic nerve head segmentation, image acquisition and quality assessment, and illumination compensation. In chapter 4 we dive even further in the acquisition procedure, more precisely the auto-focus mechanism for non-mydriatic retinal imaging. We propose a robust method for auto-focusing. In chapters 5 and 6 we discuss the problem of retinal image blurring and propose two approaches for restoring the images. Finally in chapter 7 we discuss the conclusions of the thesis.

Part II

Summary and Analysis of the Results

Chapter 2

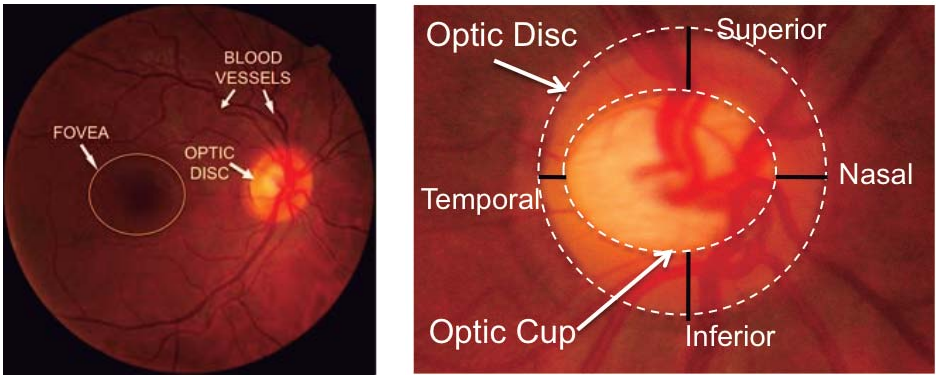
Segmentation of the Optic Disc in Retinal Images

In this chapter we discuss the segmentation of the optic disc (OD), and the method we proposed to carry it out. By way of introduction, we describe several previous works in this topic, particularly the method proposed by Valencia et al. (2006) because it served as the starting point for this thesis. This work was carried out in our research group and the analysis of its limitations and possible extensions was the subject of our early results. A keen analysis of the previous works revealed that the greatest challenge was overcoming the inherent variability of the OD appearance across the general population. This showed that a more general and robust method for the estimation of the OD contour was needed.

2.1 Background

Localization and segmentation of the OD are important tasks in retinal image analysis. Not only is its shape an important indicator of many ophthalmic pathologies (Bock et al., 2010), but its localization is a prerequisite in many algorithms. In great part, this is due to the fact that it is often necessary to differentiate the OD from other retinal features. Moreover, correct segmentation of the OD contour is a non-trivial problem. The natural variation in the characteristics of the OD is a major difficulty for defining the contour. Blood vessels may cross the boundary of the OD obscuring the rim of the disc, with edges of vessels also acting as significant distractors (Winder et al., 2009).

The literature describes a number of algorithms for determining the OD boundary. One technique that deserves especial mention is *active contours* or *snakes*. Kass et al. (1988) proposed the concept of a deformable contour



(a) (b)

Figure 2.1: (a) Retinal image and (b) optic nerve head region.

that changes its shape depending on properties of the image*. However, Mendels et al. (1999) were one of the to first report a two-stage method for the localization of the OD boundary with the use of active contours. The first stage consisted in processing the image using gray-level mathematical morphology to remove the blood vessels. Then a snake was manually placed to detect the OD boundary. The algorithm proposed by Osareh et al. (2002) extended the work of Mendels et al. (1999) with the use of gradient vector flows to localize the OD.

The work of Valencia et al. (2006) is interesting because it included a strategy for segmenting both the OD and the optic cup (OC) (see Fig. 2.1) in order to determine a glaucomatous risk index. In spite of that, their approach is not without limitations. In what follows we briefly review the strategy proposed by Valencia et al. (2006) so that the reader may get an overview of the method. After this, we show our results on the analysis of the method's strengths and limitations.

Developing retinal image analysis techniques to detect diseases like glaucoma and aid in the prevention of blindness in the general population is both a challenge and goal worth pursuing. Glaucoma is one of the most common causes of blindness in the world (Bock et al., 2010). The disease is characterized by the progressive degeneration of optic nerve fibers showing a distinct pathogenetic image of the optic nerve head. Two commonly used parameters for glaucomatous risk assessment are the cup-to-disc ratio (**CDR**) and the **ISNT** rule (Harizman et al., 2006). The CDR is calculated as the area occupied by the OC divided by the OD area, whereas the ISNT is a mnemonic that provides an easy way to remember how the optic nerve is supposed to look in a normal eye. Normally, the neuro-retinal rim is thickest **I**nferiorly and thinnest **T**emporally, thus following the pattern of

*A more detailed description of active contours is given in § 2.3.

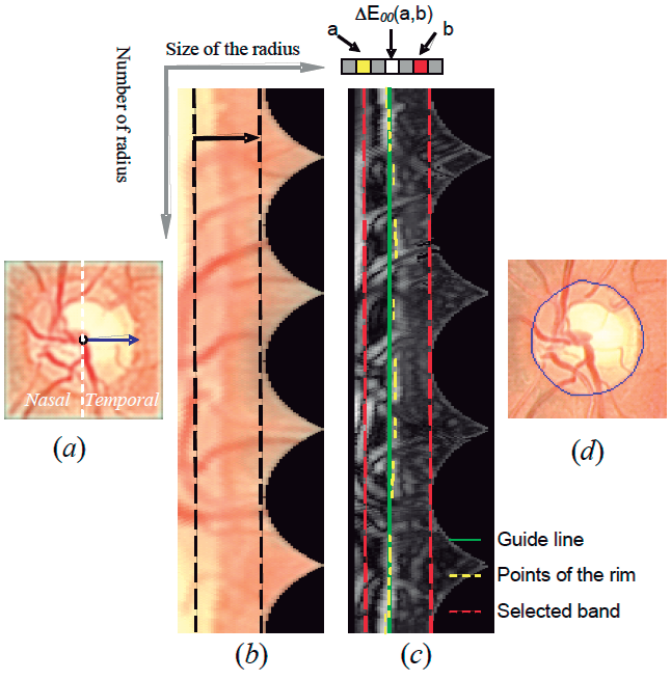


Figure 2.2: Scheme of the algorithm by Valencia et al. (2006) to mark the optic disc boundary.

rim width **I**nferior≥**S**uperior≥**N**asal≥**T**emporal (see Figure 2.1(b)). This rule is commonly used in clinical practice.

2.1.1 Previous work

The algorithm proposed in Refs. (Valencia et al., 2006, Valencia & Millán, 2008) consists in the following. The color images of the eye fundus are pre-processed using a sharpening algorithm (Millán & Valencia, 2006) to smooth noise and sharpen edges by means of the Laplacian of Gaussian operator (LoG). Subsequently, the region of interest (ROI) is manually selected and transformed to polar coordinates for OD boundary extraction (see Figure 2.2). This manual selection is a downside to the algorithm because the user is prompted to assign a manual initial guess for the OD's diameter. In the next stage, ΔE00 color differences (Johnson & Fairchild, 2003) are calculated between neighbor pixels in all radial directions to find the pixels with the highest color difference going from the center to the temporal side (Figure 2.2). These pixels are assumed to belong to the OD boundary. The authors also assumed an initial circular shape of the OD and refined it by a pixel-based criterion. The algorithm searches for local distortions of the OD boundary, and the OD contour is extracted by interpolation and a back-conversion to Cartesian coordinates.

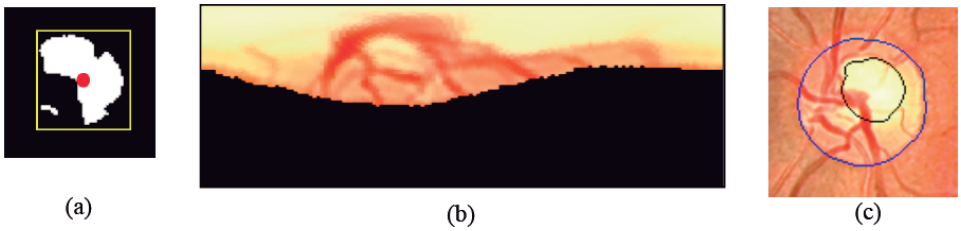


Figure 2.3: (a) Optic cup inside a square, (b) Optic cup in polar coordinates, and (c) Segmented cup (black line) and optic disc (blue line) in the original image. Figure from Valencia et al. (2006).

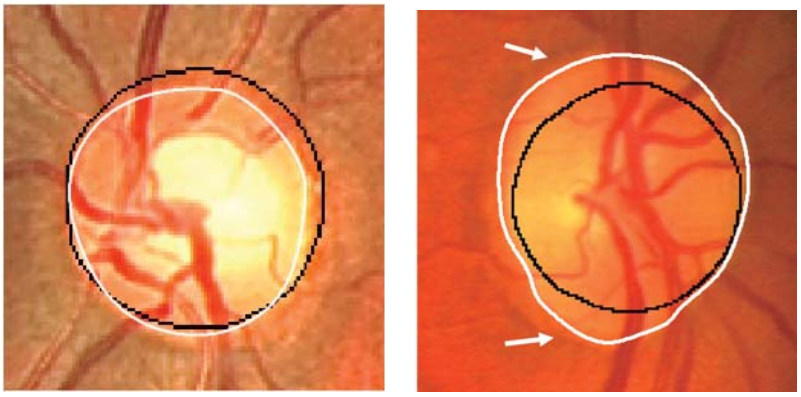


Figure 2.4: Examples for OD segmentation using the algorithm by Valencia et al. (2006). Ground truth in white and contour produced by algorithm in black. (a) nearly circular shaped OD, (b) an elliptical shaped OD

As mentioned previously this approach also deals with OC extraction by means of color seed selection, thresholding, and boundary extraction in the polar coordinate system (see Figure 2.3). From this double segmentation of the OD and the OC, the authors were able to calculate the Cup-to-Disc ratio and determine whether the ISNT rule was fulfilled. In short, they were able to compute quantitatively a glaucomatous risk index.

Figure 2.4(a) shows a satisfactory segmentation for a circular-shaped OD. The resulting segmentation is compared against manually labeled ground truth produced by an expert. It comes as no surprise that the contours match quite well. However, a clear downside of this approach manifests itself when segmenting ODs with not so regular shapes. When the OD tends to a more elliptical shape, which could be a sign of potential risk of glaucoma at early stages, we show that the output may significantly differ from the ground truth. The segmentation of an elliptical-shaped OD is shown in Figure 2.4(b). The arrows indicate regions which cannot be correctly segmented due to the assumption that the OD is approximately circular.

2.2 Compression Effects in Segmentation

REFERENCE TO THE PUBLICATIONS OF THIS THESIS

The content of this section is included in the publications:

A. G. Marrugo and M. S. Millán, "Efectos de compresión en imágenes de la retina para la evaluación del riesgo glaucomatoso", in *IX Reunión Nacional de Óptica*, pp. 140, Orense (Spain) (2009).

A. G. Marrugo and M. S. Millán, "Optic Disc Segmentation in Retinal Images", *Opt. Pura Apl.*, **43**(2), 79–86 (2010).

Nowadays there are many medical instruments that acquire images and, with the purpose of saving memory space, most of them are saved by default under lossy compression standards, such as classic JPEG. Lossy compression involves deliberately discarding information that is not visually or diagnostically important (Clunie, 2000). Until recently, the governing premise for lossy compression was that the average human observer would not notice the difference between the original image and the lossy compressed one. However, for quantitative analysis, measurement of lesion size, computer assisted detection, images often need to be interpreted by non-human observers (computers). Lossy compression may affect such automated or semi-automated methods. So far, standards for a reliable diagnosis using lossy compressed images have been established in radiology and pathology, whereas in ophthalmology this is still an open issue (Conrath et al., 2007, Winder et al., 2009). Therefore, we analyzed the segmentation of the OD in retinal images degraded by lossy compression to determine possible problems and establish a safe ground or a quantifiably lossy threshold.

The lossy compression analysis was carried out using the algorithm described previously (Valencia et al., 2006, Valencia & Millán, 2008). We used two of the most common lossy compression standards, classic JPEG and JPEG-2000, to determine the effect in OD and OC segmentation. JPEG compression is a block-DCT (discrete cosine transform) based technique. It is typically performed on 8×8 blocks in the image and the coefficients in each block are quantized separately. This leads to artificial horizontal and vertical borders between these blocks (Ebrahimi et al., 2004). On the other hand, JPEG-2000 is a wavelet based technique, in which the compression is achieved by quantizing the wavelet coefficients to reduce the number of bits to represent them. In terms of visual artifacts, this produces ringing artifacts manifested as blur and rings near edges in the image due to the attenuation of high frequencies. Figure 2.5 shows a portion of one of the images used in our tests encoded with JPEG and JPEG2000 at a compression ratio of 1:47. Blockiness and blur are visible in the JPEG- and JPEG2000-encoded images, respectively.

A set of 20 color retinal images were compressed with both standards

(a) Original (b) JPEG (c) JPEG 2000

Figure 2.5: (a) The original image, (b) the JPEG- and (c) JPEG-2000-encoded images at a compression ratio of 1:47. The respective artifacts of blockiness and blur are visible in the compressed images.

under ratios of 1:2, 1:8, 1:11, 1:22, 1:31 and 1:47. The low compression ratios are used as a reference and the higher ratios correspond to the ones used by Conrath et al. (2007) in a subjective study for the assessment of diabetic retinopathy. Figure 2.6(a) shows the segmentation of both OD and OC from the original image without compression in TIFF format. An example of the effects of compression in segmentation of OD and OC is shown in Figure 2.6(b)-(g). From these figures we can see that OC segmentation varies considerably under the effects of classic JPEG compression. On the contrary, the OC segmentation under JPEG-2000 is more stable. A commonly used parameter to illustrate the corruption in degraded images is the peak signal to noise ratio (PSNR) defined as

$$\mathrm{PSNR} = 10 \log_{10} \left(\frac{\max_I^2}{\mathrm{MSE}} \right) \quad , \tag{2.1}$$

where $\max_I$ is the maximum possible pixel value of the image, MSE is the mean squared error given by

$$\mathrm{MSE} = \frac{1}{mn} \sum_q \sum_p [I(q,p) - K(q,p)]^2 \quad , \tag{2.2}$$

where I is a noise-free $m \times n$ image and K is its degraded version. Higher PSNR indicates that the image is of better quality (less degraded). Although, there exists other parameters that correlate better with perceptual assessment such as the Structural Similarity Index (SSIM) (Wang et al., 2004), the PSNR gives enough information to set a reference so that our results may be compared against others. The PSNR was calculated for all compression ratios under JPEG and JPEG-2000 standards. In Figure 2.7(a) we show the average PSNR for the ROI. The standard JPEG-2000 slightly outperforms JPEG.

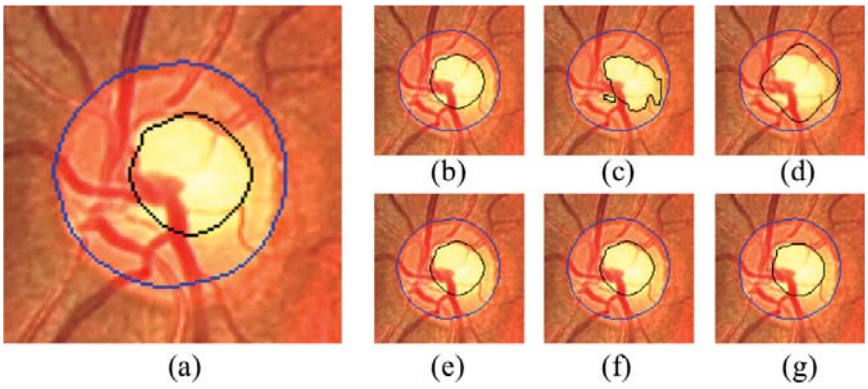
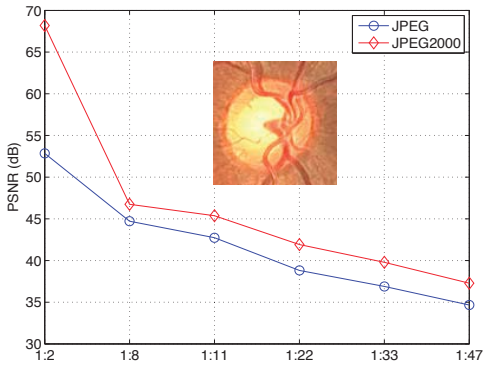


Figure 2.6: (a) OD and OC segmentation in image without compression. Segmentation with JPEG compression ratios of (b) 1:2, (c) 1:8, (d) 1:11 and JPEG-2000 (e) 1:2, (f) 1:8, and (g) 1:11.

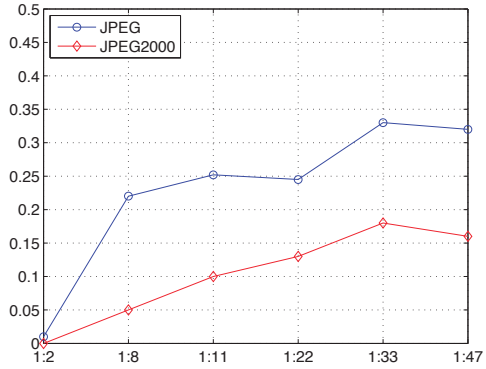
To appropriately assess the effect in the segmentation we derived a measure based on the l_2-norm of the CDR and ISNT parameters. We recall that I, S, N, and T parameters correspond to the widths of the neuro-retinal rim in the inferior, superior, nasal and temporal directions, respectively. In Figure 2.1(b), for instance, these quantities correspond to the black segment lengths measured in pixels. For each image, we take a 1-dimensional vector expressed as $\mathbf{S}^j = \left\{S_i^j\right\}$, $j = A, B$, $i = 1, \ldots, 4$, $S_1^j = \text{CDR}$, $S_2^j = \text{I}$, $S_3^j = \text{S}$, $S_4^j = \text{N}$; where A and B correspond to the original and the compressed images, respectively. In other words, the elements of each vector correspond to the CDR and I, S, N parameters, where the last three components are normalized to T. Therefore, the normalized distance measure is calculated as

$$d = \frac{\|\mathbf{S}^A - \mathbf{S}^B\|}{\|\mathbf{S}^A\|} = \frac{\sqrt{\sum_{i=1}^{n} |S_i^A - S_i^B|^2}}{\sqrt{\sum_{i=1}^{n} |S_i^A|^2}} \quad . \tag{2.3}$$

The average distance measure for all 20 images of the set is shown in Figure 2.7(b). As expected, for very low compression ratios there is a negligible difference, whereas for mid and high ratios the difference does become significant, particularly for JPEG compression. As mentioned previously it is a known fact that JPEG and JPEG-2000 lossy compression standards introduce several degradations or artifacts that increase as the compression ratio increases. The results suggest that the pixel-wise segmentation of the OD and OC is negatively affected by the blocking artifacts from classic JPEG, and more robust to the blurring effect of JPEG-2000. In addition, it is apparent that any algorithm that performs segmentation based on color seed selection is prone to produce an inadequate segmentation under high ratio lossy compression. Moreover, JPEG-2000 compression is clearly more



(a)



(b)

Figure 2.7: (a) PSNR in dB for the different compression ratios under JPEG and JPEG-2000 standards. (b) Normalized distance measure for CDR and ISNT.

reliable than classic JPEG for OD and OC feature extraction. This is most likely due to the the use of the Discrete Wavelet Transform and a more sophisticated entropy encoding scheme which translates into less visible artifacts and nearly no blocking.

2.3 Optic Disc Segmentation by Means of Active Contours

REFERENCE TO THE PUBLICATIONS OF THIS THESIS

The content of this section is included in the publications:

A. G. Marrugo and M. S. Millán, "Optic disc segmentation in retinal images", *Opt. Pura Apl.*, **43**(2), 79–86 (2010).

A. G. Marrugo and M. S. Millán, "Retinal image analysis: preprocessing and feature extraction" *Journal of Physics: Conference Series*, **274**(1), 012039, (2011).

In this section we develop a strategy for OD boundary extraction in ocular fundus images. The pre-processing stage consists in performing color mathematical morphology to remove the blood vessel regions. Subsequently, an active contours approach is used to determine the OD boundary. An active contour is an energy-minimizing spline guided by external constraint forces influenced by image forces that pull it toward features such as lines and edges (Chan & Vese, 2001). Our approach is formulated in the CIELAB (La*b*) (Johnson & Fairchild, 2003) color space (from now on Lab space) to take full advantage of the color features available for the pre-processing and feature extraction stages.

2.3.1 Color Mathematical Morphology

Active contours generally work by locking onto homogeneous regions of a given image. This task is made extremely difficult in the case of OD segmentation because the OD region is fragmented into multiple subregions by blood vessels. Furthermore, the blood vessels enter the OD from different directions with a general tendency to concentrate around the nasal side of the OD region. Mathematical morphology can extract important shape characteristics and also remove irrelevant information. It typically probes an image with a small shape or template known as a structuring element. Using gray-level morphology, the operation can be applied to the intensity or lightness channel. Osareh et al. (2002) showed that in retinal images color morphology outperforms gray- level morphology, which results in more homogeneous regions and better preservation of the OD edges. They used a definition of color morphology within the Lab color space based on a color difference metric. We performed a closing operation, i.e. dilation to first remove the blood vessels and then an erosion to approximately restore the boundaries to their former position.

Color mathematical morphology differs with gray-scale mathematical morphology in that for each point or pixel there exists three color values, in our case (L, a^*, b^*), instead of a single intensity value. The operations are defined in the same way, but a lexicographical order must be defined to be able to perform max and min operations. For the sake of simplicity, let us define the operations in gray-scale. The image is considered a function f that maps a point p in $\mathbb{Z}^2$ with spatial coordinates (i, j) to a real number, that is $f : \mathbb{Z}^2 \to \mathbb{R}$. The morphological operations, erosion ($\ominus$) and dilation ($\oplus$), at the point p by a flat structuring element $B \subseteq \mathbb{Z}^2$ are defined as

$$(f \ominus B)(p) = \min \{f(q) \mid q \in (B)_p\} \quad , \tag{2.4}$$

$$(f \oplus B)(p) = \max \{f(q) \mid q \in (B^s)_p\} \quad , \tag{2.5}$$

where $(B)_p$ is the translation of B by p defined as

$$(B)_p = \{q \in \mathbb{Z}^2 \mid q = b + p \text{ for some } b \in B\} \quad , \tag{2.6}$$

and B^s is the reflection of B

$$B^s = \{q \mid \text{for some } b \in B,\ q = -b\} \ . \tag{2.7}$$

In Fig. 2.8 we show an example to demonstrate how we can use Eqs. (2.4) and (2.5) to perform erosion and dilation on a gray-scale image by a flat structuring element. In that example we are interested in computing $(f \ominus B)$ in point $p = (1, 2)$. The structuring element B is defined as

$$B = \{(0, 0), (0, 1), (-1, 0), (-1, 1)\} \ ,$$

thus, according to Eq. (2.4)

$$\begin{aligned}
(f \ominus B)(p) &= (f \ominus B)(1, 2) \ , \\
&= \min\{f(0 + 1, 0 + 2), f(0 + 1, 1 + 2), f(-1 + 1, 0 + 2), \\
&\quad f(-1 + 1, 1 + 2)\} \ , \\
&= \min\{f(1, 2), f(1, 3), f(0, 2), f(0, 3)\} \ , \\
&= \min\{3, 2, 7, 5\} \ , \\
&= 2 \ .
\end{aligned} \tag{2.8}$$

Osareh et al. (2002) introduced a lexicographical order to color morphology in the Lab space such that basic morphological operations could be performed. This is a problem-oriented formulation based on the knowledge that the OD region contains contrasting pixels: bright, almost saturated regions crossed by dark blood vessel regions. These color differences will reside in well-separated regions of the Lab color space. Given that color differences in the Lab space correspond to the metric distance between them, the basic morphological operations of dilation and erosion are defined using the color difference of all pixels within the structuring element to a certain reference point. The color difference within the Lab color space are obtained using the Euclidean norm, and the reference point is established at the origin $(0, 0, 0)$. The dilation is the furthest point from the origin, and the erosion is the point

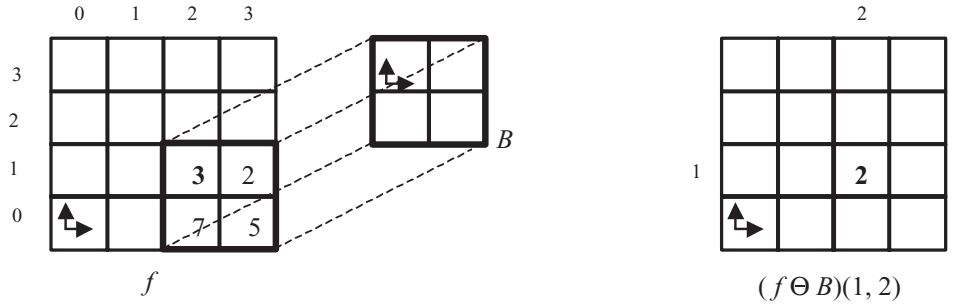


Figure 2.8: Example of a gray-scale mathematical morphology operation, in this case erosion $(f \ominus B)(1, 2)$.

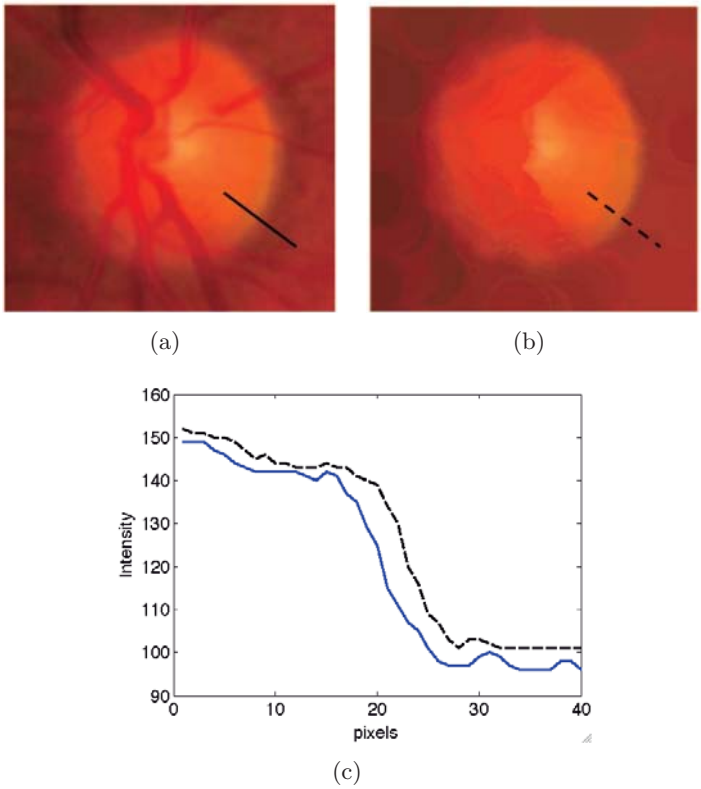


Figure 2.9: Color morphology closing. (a) Original ROI (150×150 pixels) with OD inside. (b) Lab closing with a 25×25 disc-type structuring element. (c) Intensity profile cuts from (a) solid and (b) dashed.

closest to the origin. The closing operation involves a dilation followed by erosion. An example of closing using this formulation with a disc-type structuring element is shown in Figure 2.9(b). It is evident that this approach produces a more homogeneous region while approximately preserving the OD edges Figure 2.9(c). An important aspect of this preprocessing is that any lossy compression artifact that may be present in the image is removed because of the closing operation.

2.3.2 Active Contours

The OD boundary is determined by fitting a geometric active contour model, namely the Chan-Vese (Chan & Vese, 2001) model. In general the active contour consists of a set of points placed near the contour of interest, which are gradually brought closer to the exact shape of the desired region in the image. This is carried out through iterative minimization of an energy function. The Chan-Vese model (Chan & Vese, 2001) establishes the following energy function for an image u_0:

$$F(c_1, c_2, C) = \int\limits_{\text{out(C)}} |u_0(x,y) - c_1|^2 \, \mathrm{d}x\mathrm{d}y + \int\limits_{\text{in(C)}} |u_0(x,y) - c_2|^2 \, \mathrm{d}x\mathrm{d}y + g(C) \ , \tag{2.9}$$

where $C \subset \Omega$ is a piecewise parameterized curve (contour), g is any function evaluated at C, and c_1 and c_2 represent the average intensity value of u_0 inside and outside the curve, respectively. Minimizing the fitting error in Eq. (2.9) the model looks for the best partition of u_0 taking only two values, namely c_1 and c_2, and with one edge C, the boundary between these two regions, given by $\{u_0 \approx c_1\}$ and $\{u_0 \approx c_2\}$. Now, let us define a signed distance function ϕ that is zero exactly at C, that increases in absolute value with respect to the distance from C and that is positive inside and negative outside, as shown in Figure 2.10. By doing so, we have defined implicitly the curve as $C = \{(x,y) \,|\, \phi(x,y) = 0\}$. Therefore, the energy function is expressed as:

$$\begin{aligned} F(c_1, c_2, \phi) &= \int\limits_{\phi>0} |u_0(x,y) - c_1|^2 \, \mathrm{d}x\mathrm{d}y \int\limits_{\phi<0} |u_0(x,y) - c_2|^2 \, \mathrm{d}x\mathrm{d}y + g(\phi) \\ F(c_1, c_2, \phi) &= \int\limits_{\Omega} |u_0(x,y) - c_1|^2 H(\phi) \, \mathrm{d}x\mathrm{d}y \\ &\quad + \int\limits_{\Omega} |u_0(x,y) - c_2|^2 (1 - H(\phi)) \, \mathrm{d}x\mathrm{d}y + g(\phi) \ , \end{aligned} \tag{2.10}$$

where $H(.)$ is the Heaviside function. Keeping c_1 and c_2 fixed, and minimizing F with respect to ϕ we obtain the associated Euler-Lagrange equation for ϕ. Parameterizing the descent direction by an artificial time $t \geq 0$ (or number of iterations), the equation in $\phi(t,x,y)$ (with $\phi(0,x,y) = \phi_0(x,y)$ defining the initial contour) is:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\mathrm{div}\left(\frac{\nabla \phi}{|\nabla \phi|} \right) - (u_0 - c_1)^2 + (u_0 - c_2)^2 \right] \tag{2.11}$$

where $\delta(.)$ is the Dirac function. This partial differential equation is solved numerically using a finite difference scheme. In relation to the problem at hand, we take the initial contour to be a circle big enough to fully contain the OD. From this circle a signed distance map is built for ϕ_0, fulfilling the condition to be positive inside the contour, zero exactly at the boundary, and negative outside. The iterative process consists in calculating the force from the image information, from the curvature penalty, and later evolving the curve (i.e. calculating ϕ_{n+1}).

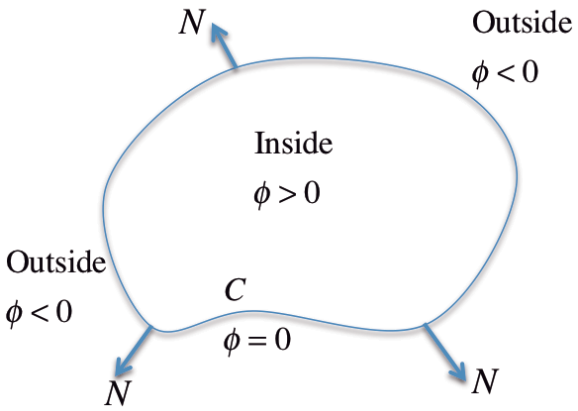


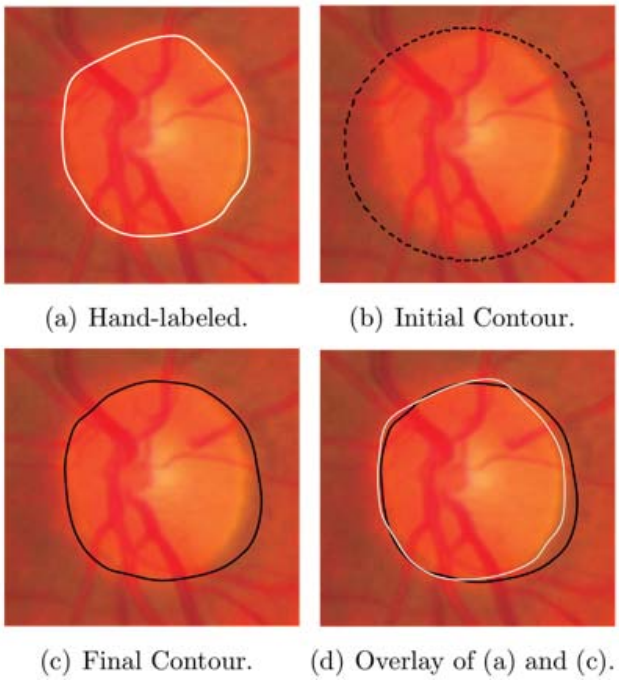
Figure 2.10: Curve C propagating in the normal direction.

2.3.3 Optic Disc Segmentation Results

The ROI is selected manually as a window of 150 × 150 pixels, with the whole OD inside Figure 2.9(a). We applied the Lab closing to all images using a symmetrical 25 × 25 pixels disc-structuring element since the blood vessels were determined not to be wider than 20 pixels. The Lab closing allowed us to remove the blood vessels cleanly and provided the required uniform OD region to initialize the active contour (Figure 2.9(b)). The active contours approach requires an intensity or gray-scale image to perform the optimization procedure. Therefore, instead of solely using the lightness channel L and, more importantly, to be consistent with the color mathematical morphology approach, we decided to use the weighting function based on the Euclidean distance within the Lab space, as described in § 2.3.1. This feature is fundamental to obtain a uniform OD region because our approach is based on the segmentation of pixels with similar color properties.

Following the color morphological pre-processing step, we initialized the contour as a circle with the center at the brightest area and with a diameter equivalent to 80% of the ROI diameter. From these initial conditions the active contour iteratively shrank towards the final boundary. The number of iterations for the final contour convergence was determined empirically and set to 450 for all cases. In Figure 2.11(a)-(c) we show the hand-labeled ground- truth OD, the initial contour, and the final contour respectively.

In Figure 2.11(d) we show the hand-labeled boundary together with the final contour to illustrate the close match achieved. We quantify the accuracy of the boundary localization against the manually labeled ground truth produced by an expert. We use a simple and effective overlap measure of the match between two regions as:



(a) Hand-labeled. (b) Initial Contour.

(c) Final Contour. (d) Overlay of (a) and (c).

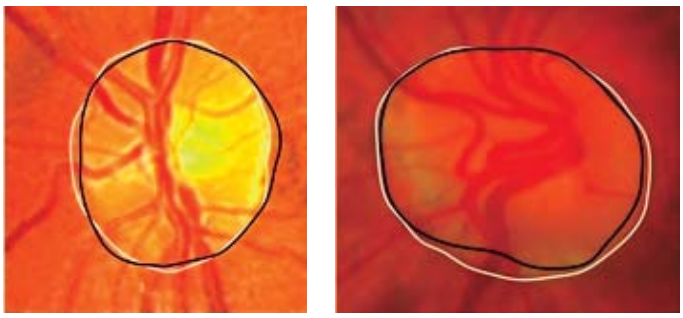
Figure 2.11: Optic disc segmentation results.

$$M = \frac{n(R \cap T)}{n(R \cup T)} \times 100 \ , \tag{2.12}$$

where R and T correspond to the ground-truth and the final OD contour region respectively, and $n(.)$ is the number of pixels in a region. In the optimal case, when both contours perfectly match $M = 100$. The measure M represents the accuracy. When compared with the hand-labeled ground-truth information from the expert, our method was able to localize the OD pixels in all test images with an average accuracy of 85.67 % ($\sigma = 7.82$). Additional tests are shown in Figure 2.12 for some ODs whose shapes differ significantly from a circle. Notice the excellent agreement in Figure 2.12(b) and the improvement achieved in Figure 2.12(c) in comparison with the previous segmentation of Figure 2.4(b).

2.4 Discussion

In this chapter we have discussed two different approaches for OD segmentation. The analysis of the algorithm by Valencia et al. (2006) revealed the need for a more general and robust approach, which would enable the segmentation of OD boundaries that differ considerably from a circular or



(a) (b)

(c)

Figure 2.12: Other optic disc segmentation results. Ground truth in white and algorithm output in black. M values are: (a) 92.61, (b) 90.32, and (c) 88.15.

more regular shape. As regards to compression effects in segmentation of the optic nerve head, we determined that degradation introduced by lossy compression plays an important role and cannot be neglected when processing compressed images. Nonetheless, our results showed that JPEG-2000 compression might provide a safer ground for retinal image segmentation than classical JPEG.

In addition we developed different strategy for OD segmentation based on active contours. The pre-processing stage consisted in performing color mathematical morphology. This provided a vessel-free OD region with uniform color distribution and preservation of the edges. The active contours algorithm for OD segmentation yielded a fair approximation to the actual hand-labeled OD. Our method was able to achieve an average accuracy rate in pixel classification of 85.67 % ($\sigma = 7.82$).

Chapter 3

Acquisition of Retinal Images: Addressing the Limitations

In this chapter we take another look at the procedure for acquiring retinal images and the difficulties that arise from it. Either for medical interpretation or for automated analysis, good quality retinal images are required in order to extract meaningful information. As a result, an initial estimate of image quality is advisable before performing any further processing or analysis on the images. Furthermore, if there are several images of the same patient available it would be desirable to process or analyze the "best" image. Obviously this requires a quality estimation and sorting of images, which we address in this chapter from a *no-reference* image quality point of view.

Most retinal image quality issues are related to problems during the acquisition. More often than not they manifest as improper (non-uniform) illumination or lack of sharp focus (blur). In this chapter we also address the problem of illumination compensation to enhance image visibility. The experiments and results described in this chapter are, for the most part, preliminary work of this thesis. Despite that, they are interesting on their own and have paved the way for further experiments discussed in subsequent chapters.

3.1 Retinal imaging

3.1.1 The retinal image and the fundus camera

Retinal imaging is central to the clinical care and management of patients with retinal diseases. It is widely used for population-based, large scale detection of diabetic retinopathy, glaucoma, age-related macular degeneration,

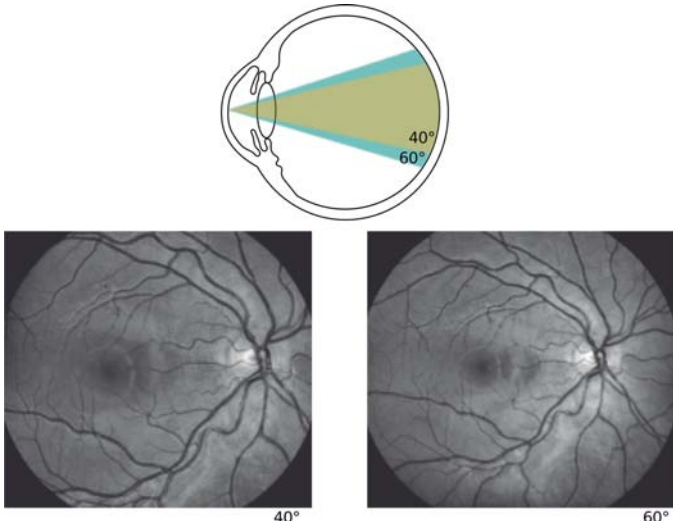


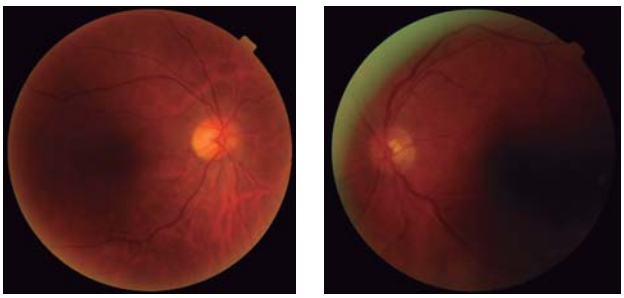
Figure 3.1: Field of view in fundus cameras.

and other eye-related diseases (Abramoff et al., 2010).

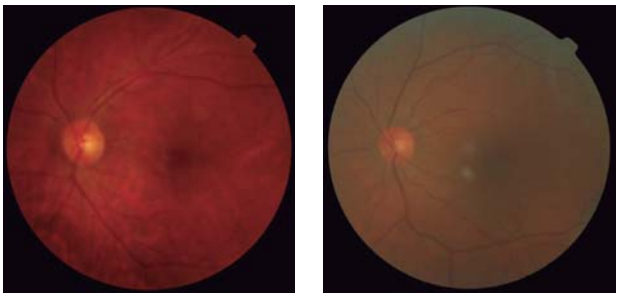
By *retinal imaging or fundus imaging* we refer to the process whereby a 2D representation of the reflected light obtained from the 3D retinal tissues is projected onto the imaging plane (Abramoff et al., 2010). By this definition fundus imaging may refer to a number of different modalities, like fundus photography, scanning laser ophthalmoscopy, fluorescein angiography, among others. The one we are interested in is conventional fundus photography in which the image intensities represent the amount of reflected light in the visible spectral waveband, or the R, G, and B wavebands characteristic of color fundus photography. Therefore, fundus photography or retinal imaging are used interchangeably hereafter. In retinal imaging the retina is photographed directly as the pupil is used as both an entrance and exit for the fundus camera's illuminating and imaging light rays (Saine & Tyler, 2002).

Regarding the imaging device, the *fundus camera* is actually a specialized low power microscope with an attached camera. Its optical design is based on the indirect ophthalmoscope*. Fundus cameras are described by the field of view, more precisely the optical angle of acceptance of the lens (Fig. 3.1). An angle of 30°, considered the normal angle of view, creates an image 2.5 times larger than the real-life object. Wide angle fundus cameras capture images over 45° and provide proportionately less retinal magnification. A narrow angle fundus camera has an angle of view of 20° or less (Saine & Tyler, 2002).

*For further details on the fundus camera see Chapter 1



(a) Uneven illumination (b) Underexposed



(c) Blurred (d) Poor contrast

Figure 3.2: Example of poor quality retinal images.

3.1.2 Errors in fundus photography

Successful fundus photographs are obtained through the mutual interaction and proper alignment of the patient, the camera, and the photographer. The photographer must correctly align and set the camera controls (Saine & Tyler, 2002). The ideal fundus photograph is an accurate visual representation of the retina. Of course, as with any complex process, many things may go wrong, and the end result is an image of poor quality. The most common causes are: (i) Patient movement, like in light-sensitive patients or patients with insufficient fixation capability, which do not cooperate easily. (ii) Insufficient pupil dilation, typically found in patients with glaucoma and diabetes. (iii) The optical quality of the eye, for example elderly patients often have cataracts. These difficulties manifest as image *artifacts* resulting in blurry images, or images with non-uniform illumination and poor contrast as shown in Fig. 3.2.

Whenever possible these limitations should be tackled during the acquisition procedure in order to avoid or, at least, reduce further post-processing of the images. This typically involves having best practice principles. For instance, making sure that the pupils are properly dilated, performing proper eye-camera alignment, carrying out accurate focusing* maneuvering around

*In Chapter 4 we propose a fully automatic focus measure for robust focusing.

local areas of unsharpness such as central cataracts or corneal scars (Saine & Tyler, 2002).

Despite the effort in obtaining the best possible images during acquisition, the resulting images may not always have sufficient quality for human interpretation or computer analysis. In these cases, image processing techniques can help in overcoming the acquisition difficulties by enhancing the images. For example, deconvolving blurry images,* compensating for the uneven illumination distribution and enhancing contrast,† among other image enhancement techniques that leverage the images' clinical use. In addition, automated image quality assessment may be required in order to determine images of sufficient quality.

3.2 Separating the Wheat from the Chaff

REFERENCE TO THE PUBLICATIONS OF THIS THESIS

The content of this section is included in the publications:

A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, and H. C. Abril, "No-reference quality metrics for eye fundus imaging", *Proceedings of the 14th Int. Conf. on Computer analysis of images and patterns*, Lecture Notes in Computer Science, Springer-Verlag, **6854**, 486–493 (2011).

A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, M. Šorel, and F. Šroubek, "Image analysis in modern ophthalmology: from acquisition to computer assisted diagnosis and telemedicine," *Proc. SPIE*, **8436**(1), 84360C, (2012).

3.2.1 On retinal image quality

Retinal image quality is a limiting factor for any type of image analysis technique for the detection of retinopathy (Patton et al., 2006). The imaging procedure is usually carried out in two separate steps: image acquisition and diagnostic interpretation. Image quality is subjectively evaluated by the person capturing the images and they may sometimes mistakenly accept a low quality image (Bartling et al., 2009). A recent study by Abramoff et al. (2008) using an automated system for detection of diabetic retinopathy found that from 10 000 exams 23% had insufficient image quality. Most image restoration algorithms cannot restore an image beyond a certain level of quality degradation. For that reason, accurate quality assessment algorithms, that allow the fundus photographer to avoid poor images, may eliminate the need for correction algorithms (Giancardo et al., 2010). In addition, a quality metric would allow the submission of only the best images if many are available.

*This is discussed in Chapters 5 and 6.

†This is discussed in § 3.3.

The measurement of a precise image quality index is not a straightforward task, because quality is a subjective concept which varies even between experts (Giancardo et al., 2008). Furthermore, because each person has a retina with its own characteristics (vasculature, pigmentation, lesions, etc.) that vary across the population, there is no reference or standard to compare the acquired images to. This lack of reference is why attempting to assess retinal image quality by computer analysis is so difficult. Nonetheless, by constraining the problem, existing image quality assessment methods may be used to gain insight into the complexity of retinal image quality assessment.

3.2.2 No-reference image quality metrics

No-reference assessment of image content is, perhaps, one of the most difficult—yet conceptually simple—problems in the field of image analysis (Wang & Bovik, 2006). It is only until recently that several authors have proposed no-reference metrics in an attempt to shed some light on this uncertain problem.

Image quality assessment through anisotropy

We have considered four metrics to apply them in fundus imaging. The first metric Q_1 was proposed by Gabarda & Cristóbal (2007) and is based on measuring the variance of the expected entropy of a given image upon a set of predefined directions. Before further explaining what the measure does, let us recall several things. First, that entropy histograms provide a measure of the information content of images. Second, because entropy can be applied as a global measure or as a local one, differences in entropy orientations can provide differences in the information content. That is, information can be stored in an anisotropic way. By anisotropy we mean the property of being directionally dependent.

The diversity of textures and edges in images gives rise to the anisotropy. This can be quantified by measuring entropy through the spatial-frequency content of the image in a directional scheme. Thus, differently oriented measures should provide different values of entropy according to the image anisotropy. To get an intuitive idea of a measure of anisotropy consider, for example, an image of a forest that has many vertical components due to the stems of the trees. Here the horizontal component of entropy is unbalanced with the vertical component. The idea that follows is that when an image is degraded (typically blur and/or noise) the anisotropic properties change in such a way that the low-quality images may be differentiated from the high-quality ones.

Directional entropy can be achieved by means of the Rényi entropy (Rényi, 1976), which applied to a discrete space-frequency distribution $P[n, k]$

has the form

$$R_\alpha = \frac{1}{1-\alpha}\log_2\left(\sum_n\sum_k P^\alpha[n,k]\right) \quad , \tag{3.1}$$

where n and k represent the spatial and frequency variables, respectively. In addition, values of $\alpha \geq 2$ are recommended for space-frequency distribution measures (Flandrin et al., 1994). Rényi measures must be normalized in order to preserve the unity energy condition $\sum_n \sum_k P[n,k] = 1$ (Sang & Williams, 1995). Gabarda & Cristóbal (2007) defined several normalization criteria, but determined that the best one was a type of normalization inspired from quantum mechanics in which the space-frequency distribution $P[n,k]$ has a probability density function $\breve{P}[n,k] = P[n,k]P^*[n,k]$ followed by a normalization to unity. From this and Eq. (3.1) with $\alpha = 3$ we obtain

$$\breve{R}_3 = -\frac{1}{2}\log_2\left(\sum_n\sum_k \breve{P}^3[n,k]\right) \quad . \tag{3.2}$$

Gabarda & Cristóbal (2007) used the one-dimensional pseudo-Wigner distribution to obtain the probability density function $\breve{P}[n,k]$ associated with a discrete sequence $z[n]$ of length N. Therefore, the Rényi entropy associated to a position (pixel) n is computed as

$$R[n] = -\frac{1}{2}\log_2\left(\sum_{k=1}^{N}\breve{P}_n^3[k]\right) \quad . \tag{3.3}$$

From Eq. (3.3) a local measure of anisotropy can be obtained by scanning the image with a 1D-window centered at n at different θ_i orientations for $\theta_i \in [\theta_1, \theta_2, \ldots, \theta_K]$. This provides a value of entropy $R[n,\theta_i]$ for each pixel. To define a figure of merit for the image, the expected value of Eq. (3.3) is calculated as

$$\bar{R}[\theta_i] = \sum_n R[n,\theta_i]/M \quad , \tag{3.4}$$

where M is the image size. And finally the standard deviation from the expected entropy for K orientations –the metric itself– is computed as

$$Q_1 = \left(\sum_{i=1}^{K}\left(\mu - \bar{R}[\theta_i]\right)^2/K\right)^{1/2} \quad , \tag{3.5}$$

where μ is the mean of $\bar{R}[\theta_i]$ for all K orientations. Q_1 is a good indicator of anisotropy and the authors were able to show that this measure provides a good estimate for the assessment of fidelity and quality in natural images, because their degradations may be seen as a decrease in their directional

Figure 3.3: Test image set for illustrating the anisotropy measure. Blur decreases from −10 to 0 and noise increases from 0 to 10. The central image is the original source image (from Gabarda & Cristóbal (2007)).

properties. This directional dependency is also true for fundus images, especially due to blurring or uneven illumination. To illustrate the anisotropy measure in Fig. 3.3 we show a set of degraded images. One subset (from −10 to 1) has been progressively blurred by iteratively applying a point-spread function (PSF) to the source image (labeled as "0"). The other (from 1 to 10) has been generated by iteratively adding a constant amount of noise. The noisiest image is labelled as "10". In Fig. 3.4 we show the corresponding Q_1 values for the images of Fig. 3.3. The shape of the plot resembles that of an ideal quality assessment function (Qu et al., 2006) in which a distinct and unique maximum is attained for the best quality.

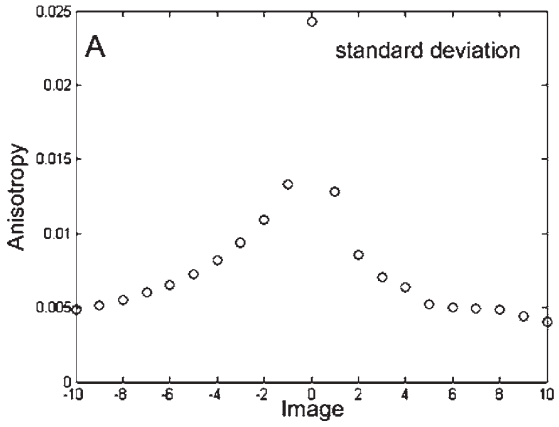


Figure 3.4: Anisotropy measure (Q_1) from the image set in Fig. 3.3.

As described in (Gabarda & Cristóbal, 2007) Q_1 assumes uniform degradation across the whole image in order to work reliably. In retinal imaging uniform degradation is not always the case. Moreover, on some retinal images local degradations or artifacts may even be tolerated because they do not hinder clinical relevancy (Saine & Tyler, 2002). To this end, we make use of a domain knowledge strategy for retinal imaging to adjust the metric so as to meet these local requirements. Our formulation is to multiply every $R[n, \theta_i]$ by a weighting function $w[n] \in [0, 1]$ such that some specific areas

are emphasized more,

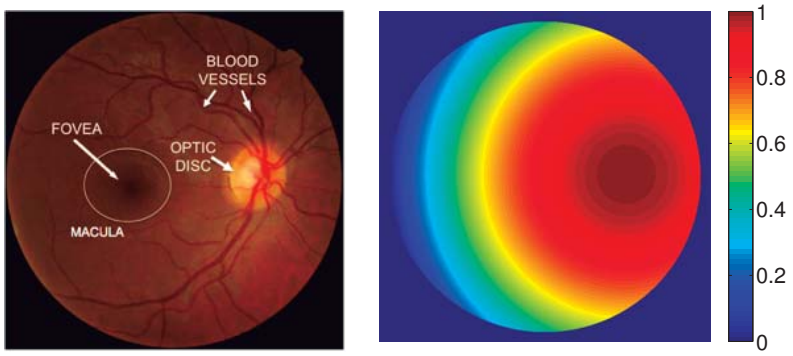
$$\bar{R}[\theta_i] = \sum_n R[n, \theta_i] w[n]/M \ . \tag{3.6}$$

This yields a modified metric Q_1'. The weighting function $w[n]$ can be defined under any arbitrary criterion depending on the desired outcome or the features of interest. In this work we have considered the following. Because two of the most relevant features of a fundus image are the optic disc (OD) and the blood vessels, which for instance are important for assessing a disease like glaucoma, we have designed a weighting function that emphasizes these structures. It is known that in order to assess image sharpness, specialists fixate on the surroundings of the OD to visualize the small blood vessels (Moscaritolo et al., 2009). The weighting function used is an elliptic paraboloid centered at the OD with values ranging from "one" (1) exactly at the position of the OD to approximately "zero" (0) very near the periphery. This function has also been used to model the illumination distribution in fundus images. The approximate position of the OD is determined via template matching (Lowell et al., 2004). The spatial distribution of the weighting function is shown in Fig. 3.5(b).

To illustrate the possible use of Q_1' let us consider the pair of retinal images shown in Fig. 3.6. Both images are from the same retina, but differ in that the image in Fig. 3.6(a) is of good quality and the one in Fig. 3.6(b) has a blue haze artifact that partially obscures the periphery of the retinal image, although the remaining image is of adequate quality. These types of artifacts are common in retinal imaging and this one in particular is caused by improper objective lens-to-cornea distance (Saine & Tyler, 2002). The values Q_1 and Q_1' for both images are shown in Table 3.1. By introducing the weighting function, the relative value of I_2 (because it is normalized to I_1) changes from 0.81 to 0.92. The important aspect to emphasize here is that, while not the ideal image, for most purposes I_2 can indeed be used because the retinal features like the OD and the blood vessels are sharp and properly defined. This is also highlighted in the experiments of § 3.2.4 in Table 3.2 for a set of retinal images with varying degree of quality.

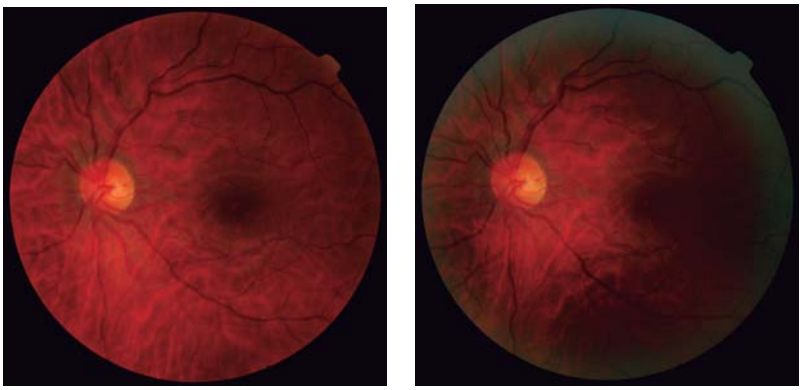
	I_1	I_2	I_2/I_1
Q_1	0.0312	0.0253	0.81
Q_1'	0.0304	0.0281	0.92

Table 3.1: No-reference image quality values Q_1 and Q_1' for the images in Fig. 3.6. The third column represents the metric for I_2 normalized to I_1



(a) (b)

Figure 3.5: (a) Normal fundus image. (b) Weighting function $w[n]$ described by an elliptic paraboloid centered at the OD.



(a) I_1 (b) I_2

Figure 3.6: A pair of retinal images from the same eye for illustrating the effect of a spatial weighting function on the metric.

Local gradients and sharpness metric

The second metric Q_2 was recently proposed by Zhu & Milanfar (2010) and it seeks to provide a quantitative measure of what they call "true image content". Without going into much detail, in the following we briefly explain the basics of Q_2. Let G be the gradient matrix over an $N \times N$ window of an image

$$\mathbf{G} = \begin{bmatrix} \vdots & \vdots \\ p_x(j) & p_x(j) \\ \vdots & \vdots \end{bmatrix} , \tag{3.7}$$

where j is every point inside the window and $[p_x(j), p_x(j)]^T$ denotes the gradient of the image at point (x_j, y_j). The local dominant orientation can be calculated by computing the singular value decomposition (SVD) of G

$$\mathbf{G} = \mathbf{USV^T} = \mathbf{U} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} [\mathbf{v_1} \ \ \mathbf{v_2}]^T \ , \tag{3.8}$$

where $\mathbf{U}$ and $\mathbf{V}$ are both orthonormal matrices. The column vector $\mathbf{v_1}$ represents the dominant orientation of the local gradient field. Correspondingly, the second singular vector $\mathbf{v_2}$ (which is orthogonal to $\mathbf{v_1}$) describes the doninant "edge orientation" of the patch. The singular values $s_1 \geq s_2 \geq 0$ represent the energy in the directions $\mathbf{v_1}$ and $\mathbf{v_2}$, respectively. In Fig. 3.7 an example of the local orientation estimation is shown.

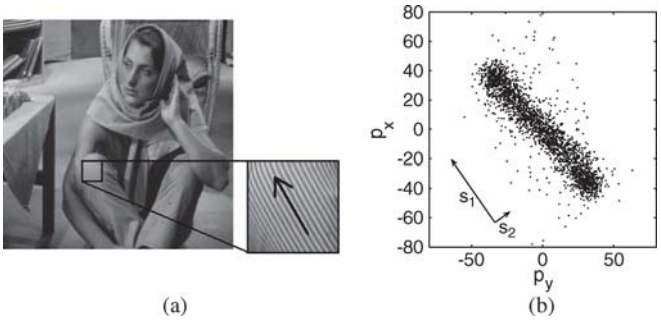


Figure 3.7: Example of local dominant orientation estimation (from Zhu & Milanfar (2010)). (b) Plots the gradient of each pixel within the chosen patch in (a). s_1 and s_2 represent the energy in the dominant orientation and its perpendicular direction, respectively.

From this result Zhu & Milanfar (2010) model different types of image patches like flat, linear, quadratic, and edge patches. They determine what happens to $\mathbf{G}$ and the singular values s_1 and s_2 when the patches are degraded by blur and/or noise. The metric that they derived is the following

$$Q_2 = s_1 \frac{s_1 - s_2}{s_1 + s_2} \ . \tag{3.9}$$

It is correlated with the noise level, sharpness, and intensity contrast manifested in visually salient geometric features such as edges. Its value generally drops if the variance of noise rises, and/or if the image content becomes blurry. To avoid regions without edges this algorithm divides the image into small patches and only processes anisotropic ones (non-homogeneous), thus local information is embedded into the final result.

The just noticeable blur

The third metric Q_3 was proposed by Ferzli & Karam (2009). It is a sharpness metric designed to be able to predict the relative amount of blurriness

in images regardless of their content. Q_3 is conceived based on the notion that the human visual system is able to mask blurriness around an edge up to a certain threshold, called the"just noticeable blur" (JNB). It is an edge-based sharpness metric based on a human visual system model that makes use of probability summation over space. JNB can be defined as the minimum amount of perceived blurriness given a contrast higher than the "Just Noticeable Difference".

As in the previous metric, the perceptual sharpness metric is not applied to the whole image, instead the image is divided into blocks. A flowchart illustrating the algorithm is shown in Fig. 3.8. Each block R_b is processed with a Sobel edge detector and is categorized as a smooth block or an edge block according to a predefined threshold T. For each edge block the edge e_i is located and the corresponding edge width $w_{JNB}(e_i)$ is computed. The perceived blur distortion is calculated as

$$D_{R_b} = \left(\sum_{e_i \in R_b} |w(e_i)/w_{JNB}(e_i)|^{\beta} \right)^{\frac{1}{\beta}} \quad , \tag{3.10}$$

where $w_{JNB}(e_i)$ is the JNB edge width which depends on the local contrast, $w(e_i)$ is the measured width of the edge e_i inside the image block R_b and β is a fitting constant with a median value of 3.6 as defined by Ferzli & Karam (2009). The overall distortion is

$$D = \left(\sum_{R_b} |D_{R_b}|^{\beta} \right)^{\frac{1}{\beta}} \quad , \tag{3.11}$$

and the no-reference objective metric is given by

$$Q_3 = \left(\frac{L}{D} \right) \quad , \tag{3.12}$$

where L is the total number of processed blocks in the image and D is given by Eq. (3.11). In Fig. 3.9 we show the performance of Q_3 against Gaussian blur as reported by Ferzli & Karam (2009). The metric is monotonic with respect to blur and decreases almost linearly.

The image variance

Finally, for the sake of completeness we include the image variance as metric Q_4 defined as

$$Q_4 = \sum_{n} (I[n] - \bar{g})^2 \quad , \tag{3.13}$$

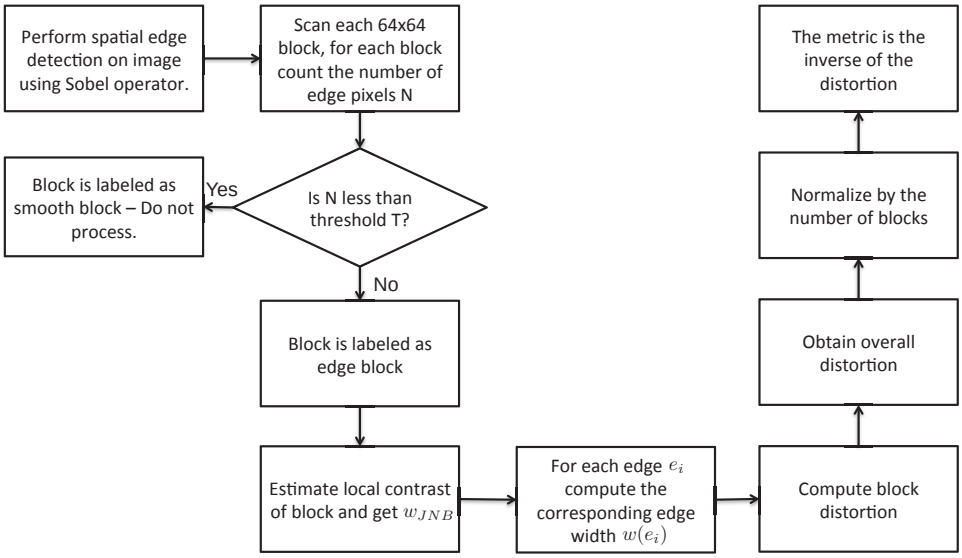


Figure 3.8: Flowchart illustrating the computation of the perceptual-based sharpness metric based on the "Just noticeable blur" (from Ferzli & Karam (2009))

where $I[n]$ indicates the gray level of pixel n, and $\bar{g}$ the gray mean of the image. This measure has been proven to be monotonic and has a straightforward relation with image quality for autoregulative illumination intensity algorithms (Qu et al., 2006).

3.2.3 Constraining the problem

It is often the case that for a given patient several fundus images are acquired. A multilevel quality estimation algorithm at the first few levels has to determine if the images correspond to fundus images, if they are properly illuminated, etc.; in other words, if they meet some minimum quality and content requirements. This is in some way what operators do, they acquire the image and then decides to accept it or not by rapidly visualizing a downscaled version of the image. Once several images of acceptable quality pass this first filter (human or machine), the system would need a final no-reference metric to decide which image to store or to send for further diagnostic interpretation. This metric should in principle yield the sharpest image, with less noise and with the most uniform illumination as possible.

In this work we seek to elucidate the possible use of the no-reference metrics for fundus image quality assessment. Our purpose is to attempt to sort a given set of retinal images acquired from the same eye from the best image down to the worse.

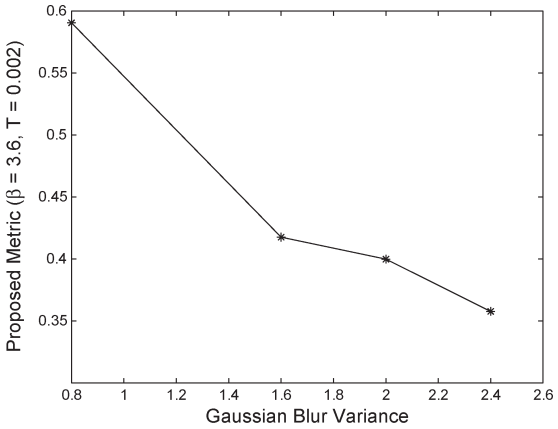


Figure 3.9: Performance of Q_3 versus Gaussian blur when applied to a set of test images of 512×512 pixels and a 7×7 Gaussian filter (from Ferzli & Karam (2009)).

3.2.4 Experiments and results

Experimental details

All images were acquired using a digital fundus camera system (TRC-NW6S, Topcon, Tokyo Japan) with a Fuji FinePix S2 Pro camera, with an image resolution of 1152×768. The images were digitized in color RGB of 24 bit-depth in TIFF format without compression. In all figures the images are shown in color, however all metrics were computed using only the luminance channel (Y) of the YUV color space as usual in image quality assessment. From Fig. 3.5(a) it is evident that the region of interest of the image is that of an approximately circular shaped area that corresponds to the captured object field. The remaining black at pixels at the corners are not of interest, thus all metrics have been modified to solely include pixels within the circular region of interest in the calculation. The neighboring pixels of the sharp black circular contour are also left aside from all calculations.

Experiments

We have analyzed a set of 20 fundus images divided in 5 subsets of 4 images corresponding to the same eye and acquired within the same session. All images within each subset have a varying degree of quality similar to the first subset shown in Fig. 3.10. The relative values from all the metrics applied to this set are shown in Table 3.2. Notice the value Q_1' for image 2. This image is in focus, however it suffers from uneven illumination. Q_1' puts more emphasis on the retinal structures, which are well defined in spite of the illumination, hence the increase with respect to Q_1. Illumination

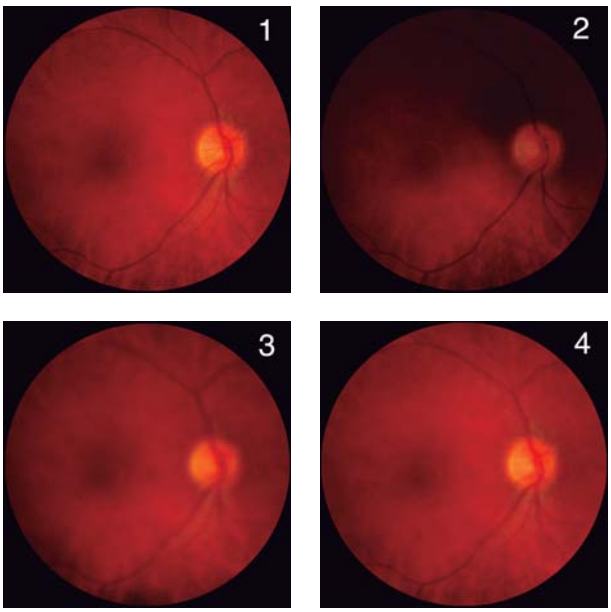


Figure 3.10: Fundus images with varying degree of quality corresponding to the same eye.

problems* are less difficult to compensate as opposed to blurring† (Marrugo et al., 2011a). This is in line with the specialist's evaluation of the images.

To validate the results two optometrists were recruited as readers **A** and **B**. They were familiarized with fundus images and were asked to examine and assess the whole set of images (4 per subject). They evaluated each subset and organized the images from the "best" to the "worse" in terms their subjective perception of sharpness and visibility of retinal structures. The relative scores of the metrics are converted to sorting or permutation indexes so as to compare with the quality sorting carried out by the readers (Table 3.3). Note that in this case only Q_1 and Q_1' agree entirely with the readers. To quantify the agreement we devised a similarity score based on the Spearman's footrule (Fagin et al., 2003). It is basically the l_1-norm of the difference between the reference permutation π_r (from the reader) and the metric Q permutation π_q. Given a set U of m elements (images), a permutation π of this set is defined as a set of indexes mapping to U to produce a particular order of the elements, $\pi : \{1, \cdots, m\} \rightarrow \{1, \cdots, m\}$. The similarity score S of two permutations π_r and π_q is defined as:

$$S = 1 - \frac{\sum_{i=1}^{m} |\pi_r(i) - \pi_q(i)|}{p_{\max}} \ , \tag{3.14}$$

*Illumination compensation is dealt in §3.3.

†Deblurring of retinal images is dealt in Chapters 5 and 6

Image	Q_1	Q_1'	Q_2	Q_3	Q_4
1	1.00	1.00	1.00	0.91	1.00
2	**0.67**	**0.90**	0.40	1.00	0.81
3	0.10	0.12	0.54	0.81	0.85
4	0.38	0.38	0.79	0.70	0.96

Table 3.2: Relative values for all the metrics applied to the set of images in Fig. 3.10.

A	**B**	Q_1	Q_1'	Q_2	Q_3	Q_4
1	**1**	1	1	1	2	1
2	**2**	2	2	4	1	4
4	**4**	4	4	3	3	3
3	**3**	3	3	2	4	2

Table 3.3: Reader A and B vs. metric sorting of images from Fig. 3.10 in accordance to quality. Top to bottom: best to worse.

where $p_{\max}$ is the maximum value of the numerator. It occurs when the permutations are reversed and it can be shown that $p_{\max}$ is equal to $m^2/2$ when m is even and $(m^2-1)/2$ when m is odd. Perfect agreement means $S = 1$, and the opposite $S = 0$. The inter-reader agreement for the whole set of 20 images yielded an S score of 0.90. The S scores for the first 4 image subset and the whole set of images are shown in Table 3.4. The difference in the overall scores for both readers is practically negligible. It is also clear that Q_1 outperforms the other metrics in this experiment with agreement scores of 0.8 and 0.9. The most probable reason is the computation of the metric from normalized space-frequency representation of the image.

3.2.5 Discussion

We have considered four state-of-the-art no-reference image quality metrics and their applicability for eye fundus imaging, particularly the problem of retinal image sorting. To this end, we showed that from the considered metrics, Q_1 and its modified version Q_1' are the most reliable in terms of agreement with expert assessment, evidenced by average similarity scores of 0.8 and 0.9 with readers A and B, respectively. Q_1 performs a measure of anisotropy throughout the whole image. That is, quantifying the fact that structures change in a directional way is good indicator of image sharpness. What this means is that the visibility of retinal structures (caused either by poor illumination or blur) is one of the most important features to take

	Q_1	Q_1'	Q_2	Q_3	Q_4
S_A 1st subset	1.00	1.00	0.50	0.50	0.50
S_B 1st subset	1.00	1.00	0.50	0.50	0.50
S_A all images	0.80	0.80	0.55	0.55	0.40
S_B all images	0.90	0.90	0.60	0.65	0.45

Table 3.4: Evaluation of the no-reference metrics w.r.t. reader grading with the use of the similarity score S in (3.14). The subindex in S indicates reader A or B. The inter-reader agreement for the whole set of 20 images yielded an S score of 0.90.

into account when evaluating retinal image quality. The results lend strong support to the development of a no-reference metric for fundus imaging based on a type of anisotropy measure. This is exactly what we have done in Chapter 4 by further developing these findings into a focus measure for retinal imaging.

3.3 Dealing with Uneven Illumination

REFERENCE TO THE PUBLICATIONS OF THIS THESIS

The content of this section is included in the publications:

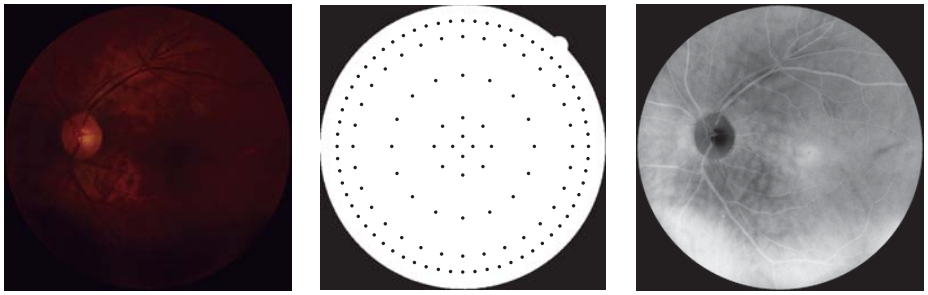
A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, M. Šorel, and F. Šroubek, "Image analysis in modern ophthalmology: from acquisition to computer assisted diagnosis and telemedicine," *Proc. SPIE*, **8436**(1), 84360C, (2012).

A. G. Marrugo,M. S. Millán, G. Cristóbal, S. Gabarda, M. Šorel, and F. Šroubek, "Toward computer-assisted diagnosis and telemedicine in ophthalmology", *SPIE Newsroom*, (doi: 10.1117/2.1201205.004256), (2012).

A. G. Marrugo and M. S. Millán, "Retinal image analysis: preprocessing and feature extraction" *Journal of Physics: Conference Series*, **274**(1), 012039, (2011).

M. S. Millán and **A. G. Marrugo**, "Image Analysis and Optics in Ophthalmology", *Lecture Notes of the Iternational Centre of Byocybernetics Seminar*, Polish Academy of Sciences, Warsaw, October, (2009).

Retinal images are acquired with a digital fundus camera which captures the illumination reflected from the retinal surface. Despite controlled conditions, many retinal images suffer from non-uniform illumination given by several factors: the curved surface of the retina, pupil dilation (highly variable among patients), or presence of diseases, among others. The curved retinal surface and the geometrical configuration of the light source and camera, lead to a poorly illuminated peripheral part of the retina with respect to the central part (Figure 3.11(a)).



(a) (b) (c)

Figure 3.11: (a) Retinal image with uneven illumination and contrast, (b) non-uniform sampling grid, and (c) first principal component of (a) from PCA analysis.

Several techniques have been used to enhance retinal images. Histogram equalization has been shown to be inappropriate for retinal images (Feng et al., 2007). A local normalization of each pixel to zero mean and unit variance aims to compensate lighting variation and enhancing local contrast but also introduces artifacts (Feng et al., 2007). Histogram matching between the red and green planes have been used as a preprocessing step for vessel segmentation (Salem & Nandi, 2007). This improves the contrast of gross dark features like vessels but reduces the contrast of bright objects and tiny dark objects like micro-aneurysms. While most of the aforementioned methods are motivated by automatic analysis, as a preprocessing stage, they are all formulated for a single color plane or for gray-scale images.

Color retinal image enhancement is required for human visual inspection or for the application of vector processing techniques. The work of Foracchia et al. (2005) aimed to introduce a strategy for luminosity and contrast enhancement on each color plane of the RGB color space, independently. This approach tended to produce hue-shifting related artifacts, given by the introduction of new colors to the image. More recently, Joshi & Sivaswamy (2008), proposed a strategy that would reduce the color artifacts by performing the enhancement on single color plane to compensate equally every channel and ultimately perform linear color remapping. However, their method has a serious shortcoming in that the pixels belonging to the OD are not always properly identified. This has a negative impact on the proper estimation of the background illumination distribution. To deal with this we propose to improve the estimation of the luminosity distribution by using principal component analysis (PCA) (Li & Chutatape, 2003) so as to leave out the OD.

3.3.1 Image enhancement on a single color plane

The main idea is that the image can be enhanced by estimating the background luminosity and contrast distribution in order to compensate for uneven illumination. The fact that a retinal image can be expressed in terms of a *background image* or background set $\mathcal{B}$ (the retinal fundus free of any vascular structure) and a *foreground image* or foreground set $1 - \mathcal{B}$ (which contains the vascular structures, optic disc, etc.) was proposed by Foracchia et al. (2005). They also derived the statistical description of a background pixel as a white random field $\mathcal{N}(\mu, \sigma)$ with mean value μ representing the ideally uniform luminosity, and standard deviation σ, representing the natural variability of the fundus pigmentation. They further developed this into the following

$$I(x, y) \sim \mathcal{N}(L(x, y), C(x, y)), \quad (x, y) \in \mathcal{B} \ , \tag{3.15}$$

where I is the original image, C and L are the contrast and luminosity drifts, respectively. L and C can also be understood in terms of gain and offset. It follows from (Foracchia et al., 2005) that estimates $\hat{L}$ and $\hat{C}$ can be computed by measuring the mean and standard deviation of the background image. With the proper estimation of the background luminosity we can subtract it from the original image to make the illumination uniform. Thus, the enhanced image $U(x, y)$ is expressed as:

$$U(x, y) = \frac{I(x, y) - L(x, y)}{C(x, y)}, \tag{3.16}$$

The sampling approach of Foracchia et al. (2005) divides the whole image into a square sampling grid, whereas in Joshi & Sivaswamy (2008) they use a more intuitive sampling scheme based on the knowledge of the illumination distribution that leads to less computational burden. Therefore, we decided to use a similar type of non-uniform sampling grid shown in figure 3.11(b). The sampling tries to compensate for the decreasing illumination outwards, therefore it is coarse in the central region and dense in the periphery.

This enhancement is oriented toward compensating the green channel of the RGB retinal image because it is the component with highest contrast. The algorithm for the enhancement is shown in Algorithm 1. In the first stage (Stage 0), the image is separated into a set of background and foreground pixels. The second stage (Stage 1) consists in computing estimates $\hat{C}$ and $\hat{L}$ of the C and L components from the background image.

This strategy is motivated by the fact that the retinal structures can bias the luminosity component. For instance, the OD is a naturally high luminosity zone, while the blood vessels typically exhibit low luminosity. The sampling scheme is as follows: for each sampling point p on the grid we take a window of size $w_0 \times w_0$ large enough to include retinal structures and the background. We compute the local mean $\mu_0[p]$ and the standard

Algorithm 1: Algorithm for enhancing the illumination of retinal images.

input : Original image $I(x, y)$
output: Enhanced image $U(x, y)$

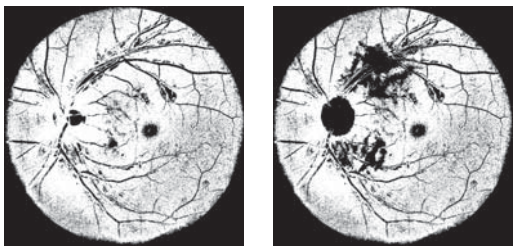
`// Stage 0: identify background pixels`
foreach *sampling point p in grid* **do**
  measure $\sigma_0[p]$, $\mu_0[p]$ in windows of $w_0 \times w_0$ from $I(x, y)$;
end
interpolate $\sigma_0[p]$, $\mu_0[p]$ for all (x, y) to obtain $\sigma_0(x, y)$, $\mu_0(x, y)$;
classify background ($\mathcal{B}$) pixels with Eq. (3.17) ;
improve background ($\mathcal{B}$) classification with PCA and thresholding ;

`// Stage 1: estimate L and C from background`
foreach *sampling point p in grid* **do**
  measure $\sigma_1[p]$, $\mu_1[p]$ in windows of $w_1 \times w_1$ from $I(x, y),\ x, y \in \mathcal{B}$;
end
interpolate $\sigma_1[p]$, $\mu_1[p]$ for all (x, y) to obtain $\sigma_1(x, y)$, $\mu_1(x, y)$;
compute $U(x, y)$ from Eq. (3.16) and $\hat{L}(x, y) = \mu_1(x, y)$,
$\hat{C}(x, y) = \sigma_1(x, y)$;

deviation $\sigma_0[p]$ for each p point. We perform bi-cubic interpolation to obtain $\mu_0(x, y)$ and $\sigma_0(x, y)$ for all (x, y) points of the retinal image. To identify background pixels the criteria defined by Foracchia et al. (2005) states that a pixel is considered to belong to the background if its Mahalanobis distance from $\mu_0(x, y)$, defined as

$$D(x, y) = \left| \frac{I(x, y) - \mu_0(x, y)}{\sigma_0(x, y)} \right|, \tag{3.17}$$

is lower that a certain threshold t, which for this work is taken as 1. This threshold is somewhat critical, because, as pointed out before, any retinal structure that does not belong to the background, especially the OD, can bias the background components. Therefore, to ensure that the OD region is not taken into account in this estimation we developed a strategy using PCA. We have not used the template matching algorithm mentioned in § 3.2.2 by Lowell et al. (2004), because it only provides an approximate location of the OD, whereas for this we need a rough estimate of the OD area. Along the same line a precise segmentation as described in § 2.3 is not needed as well. Li & Chutatape (2003) used a PCA based model to approximately localize the OD region in an intensity retinal image. More recently, Fadzil et al. (2008) developed a model for retinal pigment identification using independent component analysis (ICA) on the RGB retinal image, although no experiments were carried out including the OD. The main drawback of ICA is that it does not prioritize the output components, whereas in PCA



(a) (b)

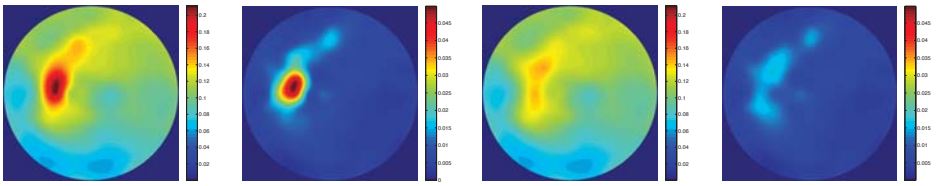
Figure 3.12: Background pixel classification from Eq. 3.17 using (a) the strategy in Foracchia et al. (2005) and (b) with additional PCA analysis. Notice that the OD region has been left out in order not to bias the estimation of the luminosity component.

this is not an issue. As a result we used PCA on the three RGB channels to identify the OD region. The first principal component from the image in Fig. 3.11(a) is shown in Fig. 3.11(c). It can clearly be seen that the OD region has different properties than the surrounding retinal regions. Using the first principal component and a simple thresholding approach, like that of Otsu (1979), the OD region can be entirely left out from the background image as shown in Figure 3.12(b).

In the following stage (Stage 1) we estimate the $\hat{L}$ and $\hat{C}$ components from the background image to perform the enhancement. It involves repeating the sampling scheme but this time including just background pixels ($\mathcal{B}$) and with a smaller window of size $w_1 \times w_1$ to increase the precision in the estimation of $\mu_1[p]$ and $\sigma_1[p]$. Bi-cubic interpolation is carried out to obtain $\mu_1(x, y)$ and $\sigma_1(x, y)$ for all (x, y). From Foracchia et al. (2005) $\hat{L}$ and $\hat{C}$ can be approximated as $\mu_1(x, y)$ and $\sigma_1(x, y)$. The enhanced image is obtained by applying Eq.(3.16). In our experiments we set $w_0 = 125$ and $w_1 = 51$. The estimated $\hat{C}$ and $\hat{L}$ components are shown in Figure 3.13. Notice how the OD region has little influence on the components in Figures 3.13(c) and (d) that were computed from the background pixels after the PCA and thresholding operation. To illustrate the impact on the single channel enhancement by applying Eq. (3.16), both images are shown in Figure 3.14. Note that in Fig. 3.14(b) the illumination in the surrounding area of the OD has not been significantly modified when compared to Fig. 3.14(a).

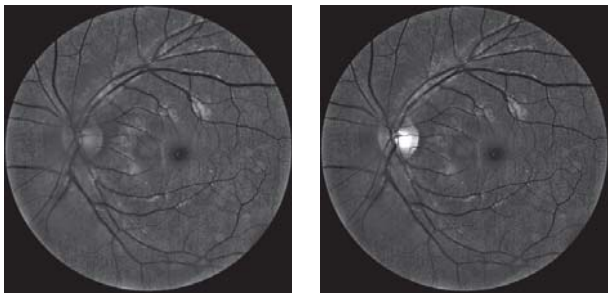
3.3.2 Color remapping

After the single channel enhancement we perform the following color remapping: given a color image with color components (r, g, b), the single plane enhancement is applied to the g plane and g_{enh} is obtained. Next, the en-



(a) (b) (c) (d)

Figure 3.13: Estimated $\hat{L}$ and $\hat{C}$ components using background pixels ($\mathcal{B}$) (a)-(b) from Figure 3.12(a) and (c)-(d) from Figure 3.12(b). For the purpose of comparison (a) and (c), as well as (b) and (d) are in the same scale. As expected, notice how the OD region has little influence on the components in (c)-(d).



(a) (b)

Figure 3.14: Image enhancement on single channel from (a) the strategy by Joshi & Sivaswamy (2008) and (b) with additional PCA analysis. In (b) the illumination in the surrounding area of the OD has not been modified significantly compared to that in (a).

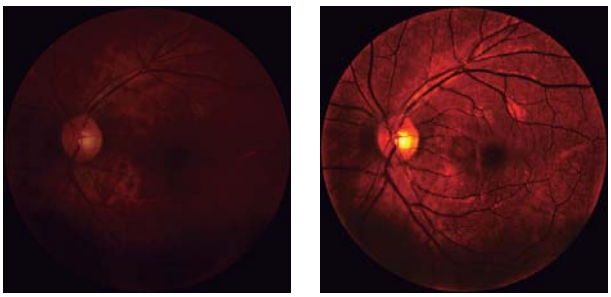
hanced color image $(\hat{r}, \hat{g}, \hat{b})$ is computed on pixel basis as:

$$\hat{r} = \frac{g_{enh}}{v} \cdot r, \ \hat{g} = \frac{g_{enh}}{v} \cdot g, \ \hat{b} = \frac{g_{enh}}{v} \cdot b, \tag{3.18}$$

where v is a scalar defined as $v = max[r_{max}, g_{max}, b_{max}]$ to play a normalization role in the enhancement. Thus, the ratio of the original r, g, and b components is maintained. Figure 3.15(b) shows the enhanced color retinal image. Notice that it has good luminosity and different retinal structures are contrasted well against the background.

3.3.3 Discussion

We have developed a strategy for retinal image enhancement. We showed that the problem of non-uniform illumination and poor contrast in retinal images may be addressed via an image enhancement technique based on the knowledge of luminosity distribution in the retina. If not taken into consideration, the retinal structures like the OD have a negative impact on the



(a) (b)

Figure 3.15: (a) Original color retinal image with uneven illumination and (b) resulting enhanced color retinal image.

estimation of the background luminosity. With the use of additional PCA analysis we have been able to leave out the OD region so as to estimate proper luminosity components for the illumination compensation. The resulting enhanced image shows remarkable gain in contrast related to retinal structures against the background. The background exhibits a much more uniform illumination distribution, in spite of a minor decrease in intensity.

Chapter 4

Robust Automated Focusing in Retinal Imaging

In the previous chapter we discussed the problems associated with the acquisition of retinal images, for example blurry retinal images because of lack of focus. In this chapter we discuss the problem of focusing as the search for the sharpest image within a sequence of images. To carry out such a task we propose a robust focus measure based on a calculation of image anisotropy. This was implemented in a non-mydriatic retinal imaging system with real subjects.

REFERENCE TO THE PUBLICATIONS OF THIS THESIS

The content of this chapter is included in the publications:

A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, M. Šorel, and F. Šroubek, "Image analysis in modern ophthalmology: from acquisition to computer assisted diagnosis and telemedicine," *Proc. SPIE*, **8436**(1), 84360C, (2012).

A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, and H. C. Abril, "Anisotropy-based robust focus measure for non-mydriatic retinal imaging," *J. Biomed. Opt.*, **17**(7), 076021, (2012).

A. G. Marrugo, M. S. Millán, and H. C. Abril, "Implementation of an Image Based Focusing Algorithm for Retinal Imaging," presented at the *X Reunión Nacional de Óptica*, Zaragoza, 40–43, (2012)

4.1 Non-Mydriatic Retinal Imaging

Fundus cameras can be mydriatic or non-mydriatic. Mydriatic fundus cameras require pharmacological dilation, while non-mydriatic cameras use a near infrared (NIR) viewing system to exploit the patient's natural dilation in a dark room (Bennett & Barry, 2009). Infrared light is used to preview the retina on a video monitor. Once the monitor's image is focused and

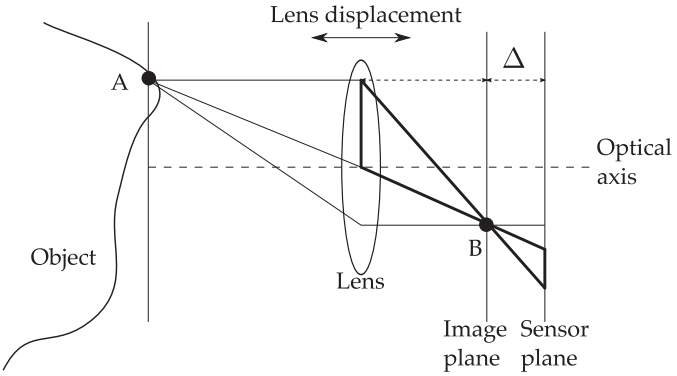


Figure 4.1: A simplified diagram of the focusing mechanism. A sharp image is obtained when a retinal layer is the object (A) and the sensor coincides with the image plane (or plane of sharp focus $\Delta \approx 0$). The focus control compensates refractive errors by displacing the lens.

aligned, a flash of visible light from a Xenon arc lamp is fired and the image is captured.

Non-mydriatic fundus cameras perform the focusing mechanism by displacing a compensation lens. This compensation lens is basically an aspheric objective lens design that, when combined with the optics of the eye, matches the image plane to the eye fundus. The focus control of the fundus camera is used to compensate for refractive errors in the subject's eye. Figure 4.1 shows a simplified diagram of the focusing mechanism. Until recently (Moscaritolo et al., 2009), these cameras were entirely operated manually with the focusing mechanism assisted by a split line visual aid. Manual focusing is error prone especially in the presence of inexperienced photographers and may lead to images that require additional restoration or enhancement (Marrugo et al., 2011a). The autofocus feature offered in new retinal cameras is a significant advance that ultimately leads to a more robust imaging system, especially for medical screening purposes. However, it still relies on the split line mechanism. Alternatively, we have proposed a passive focus measure (FM) completely based on image analysis, which we describe in the following.

4.1.1 Focusing

In a single lens optical imaging system operating within the paraxial regime the process of focusing consists in adjusting the relative position of the object, the lens, the image sensor, or a certain combination of the three to obtain a focused image (Fig. 4.1). Let $f(x, y)$ be the focused image of a planar object and $g_i(x, y)$ a sequence of images recorded for a sequence of camera parameter settings. The eye fundus is actually a curved surface, however in our case $f(x, y)$ corresponds to a small region of the fundus so

that it can be considered as an isoplanatic patch(Bedggood et al., 2008). We consider the variation of only one camera parameter at a time –either the lens position or the focal length. The acquired set of images can be expressed by convolution

$$g_i(x, y) = (f * h_i)(x, y),\ i = 1, \ldots, m\ , \tag{4.1}$$

where $h_i(x, y)$ is the point spread function (PSF) of the blur in the ith observation. In a practical imaging system the image magnification and mean image brightness change while focusing even if nothing has changed in the scene. Normalization with respect to these two parameters can be carried out. However, illumination normalization is more easily performed. Image magnification may be neglected because in most practical applications the magnification is less than 3 percent (Subbarao et al., 1993). Ideally, the best possible case occurs when $h_i(x, y) = \delta(x, y)$, therefore $g_i(x, y) = f(x, y)$. In practice all $h_i(x, y)$ have an unknown low-pass filter effect.

A FM may be understood as a functional defined on the image space which reflects the amount of blurring introduced by $h_i(x, y)$. Let S be the FM with which we look for the "best" (or sharpest) image by maximizing/minimizing $S(g_i)$ over $i = 1, \ldots, m$. A reasonable FM should be monotonic with respect to blur and robust to noise. Groen et al. (1985) used eight different criteria for the evaluation of focus functions. Ideally the focus function should be unimodal, but in practice it can present various local maxima which can affect the convergence of the autofocus procedure. Moreover, the focus curve should be ideally sharp at the top and long tailed, which can accelerate the convergence of the screening procedure.

4.2 The Focus Measure in Related Works

Various FMs have been reported in the literature (Subbarao et al., 1993, Lee et al., 2008, Kautsky et al., 2002, Aslantas & Kurban, 2009, Moscaritolo et al., 2009). They mainly consist of a focus measuring operator that estimates the sharpness of the image. The image that yields a maximum FM is considered as the focused one. Almost all FMs depend directly on the amount of high frequency information in the image. The high frequency components correspond to edge information. On the other hand, their accuracy can deviate depending on the content of the processed images. Since well focused images have sharper edges, they are expected to have higher frequency content than blurred ones (Aslantas & Kurban, 2009). The common FMs are based on norm of gradient or second derivative of the image, gray level variance and energy of Laplacian. Surprisingly, little is known about the performance of these methods for fundus imaging and the literature on this subject is scarce.

To the best of our knowledge, before our work was published (Marrugo et al., 2012b), only two other works concerning auto-focusing in retinal imaging had been published (Liatsis & Kantartzis, 2005, Moscaritolo et al., 2009). In these other works, the authors used conventional mydriatic imaging in the visible spectrum, which is not our case. In the work by Liatsis & Kantartzis (2005), the authors do not propose a FM, instead they use several preprocessing operations to improve the performance of traditional FMs for segmentation purposes. On the other hand, in the work by Moscaritolo et al. (2009), they propose a filtering technique to assess the sharpness of optic nerve head images, however no comparison with other methods was carried out. In this section we briefly summarize five notable approaches–including (Moscaritolo et al., 2009)–for later comparison with our proposed method.

The first FM S_1 was proposed in (Moscaritolo et al., 2009). It may be defined mathematically as

$$S_1 = \mathrm{Var}\left(z_{\mathrm{med}}\left|g_i - z_{\mathrm{lp}}(g_i)\right|\right) \quad , \tag{4.2}$$

where z_{lp} is a low-pass filtering of $g_i(x,y)$, z_{med} is a nonlinear median filter of the absolute value $|.|$ of the difference for removing noise, and Var(.) is the variance. Another important measure is the l_2-norm of image gradient, also called the energy of gradient, is defined as

$$S_2 = \sum_x \sum_y \left(\frac{\partial g_i(x,y)}{\partial x}\right)^2 + \left(\frac{\partial g_i(x,y)}{\partial y}\right)^2 \quad . \tag{4.3}$$

The third measure is the energy of Laplacian. It can analyze high frequencies associated with image edges and is calculated as

$$S_3 = \sum_x \sum_y \left(\nabla^2 g_i(x,y)\right)^2 \quad . \tag{4.4}$$

Nayar & Nakagawa (1994) proposed a noise-insensitive FM based on the summed modified Laplacian operators. When two second partial derivatives with respect to horizontal and vertical directions have different sign, one offsets the other and the evaluated focus value is incorrect. The method is a modification to obtain the absolute value of each second partial derivative as

$$S_4 = \sum_x \sum_y \left(\left|\partial^2 \frac{g_i(x,y)}{\partial x^2}\right| + \left|\partial^2 \frac{g_i(x,y)}{\partial y^2}\right|\right) \quad . \tag{4.5}$$

The Frequency-Selective Weighted Median (FSWM) Filter (Choi et al., 1999) is a high-pass nonlinear filter based on the difference of medians. It is well known as a nonlinear edge detector that removes impulsive noise effectively. The FSWM uses several nonlinear subfilters having a weight according to the frequency acting like a bandpass filter as

$$z_F(x) = \sum_j^P \beta_j \hat{z}_j(x) \quad , \tag{4.6}$$

where $z_F(x)$ is the FSWM filter, P is the number of subfilters, $\beta_j \in R$, and $\hat{z}_j(x)$ is the weighted median filter. The FM is produced by summing FSWM results, Fx and Fy , applied to an image along the horizontal and vertical directions as

$$S_5 = \sum_x \sum_y \left(F_x^2 + F_y^2\right) \quad . \tag{4.7}$$

Subbarao & Tyan (1998) analyzed the robustness of three FMs: the image variance (not included here), S_2, and S_3. They recommended to use S_3 because of its tolerance to additive noise. However, the differences among individual measures were not significant. There are many other FMs like the wavelet based FM proposed in Ref. (Kautsky et al., 2002), or the mid-frequency discrete cosine FM in Ref. (Lee et al., 2008) but were not included in our study either because of their lack of robustness to noise or for their complex implementation. For a review and evaluation of FMs in natural images the reader is referred to (Aslantas & Kurban, 2009, Subbarao et al., 1993).

4.3 The Focus Measure in Our Proposal

In this section we describe the proposed FM, the theoretical basis supporting it, the optimization procedure and the hardware implementation with the experimental setup.

4.3.1 Representing the blur

Discrete cosine transform

The discrete cosine transform (DCT) is an invertible, linear transformation $\mathcal{T} : \mathbb{R}^N \to \mathbb{R}^N$. An image is transformed to its spectral representations by projection onto a set of of orthogonal 2-D basis functions. The amplitude of these projections are called the DCT coefficients. Let $g(x, y)$, for $x = 0, 1, 2, \ldots, M-1$ and $y = 0, 1, 2, \ldots, N-1$, denote an $M \times N$ image and its DCT denoted by $\mathcal{T}[g(x, y)] : G(u, v)$, given by the equation

$$G(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x,y)\alpha(u)\alpha(v) \cos\left[\frac{(2x+1)u\pi}{2M}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad , \tag{4.8}$$

where

$$\alpha(\xi; \mathrm{A}) = \begin{cases} \sqrt{\frac{1}{A}} & \xi = 0, \\ \sqrt{\frac{2}{A}} & \text{otherwise} \ , \end{cases} \tag{4.9}$$

where $A = \{M, N\}$ depending on variables u and v, respectively. Low-order basis functions represent low spatial frequencies, while those of higher

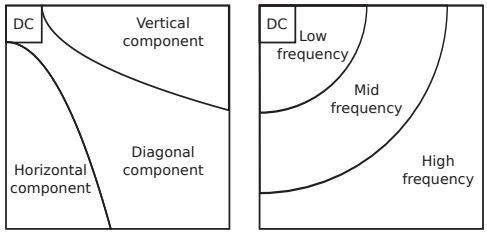


Figure 4.2: Relationship between DCT coefficients and frequency components of an image.

orders represent high spatial frequencies (Fig. 4.2). Therefore, low-order coefficients depict slow spatial variations in image intensity, while those of higher orders depict rapid variations.

The DCT is closely related to the discrete Fourier transform (DFT), a standard tool in signal processing and has been reported as a suitable transform for spectral-based focusing algorithms (Ng Kuang Chern et al., 2001). However, the DCT has a greater energy compaction property than the DFT, i.e. most of the image information tends to be concentrated in a few low frequency DCT coefficients. This is also the reason why the JPEG compression standard is based on the DCT. In addition, many efficient schemes for the computation of DCT exist (Wallace, 1992) and hardware implementations are commonly available (Ramirez et al., 2000).

Normalized DCT

The normalized DCT (Kristan et al., 2006) of an image is defined as

$$\tilde{G}(u,v) = \tilde{\mathcal{T}}[g](u,v) = \frac{|\mathcal{T}[g](u,v)|}{\sum_{(u,v)}|\mathcal{T}[g](u,v)|} \ , \tag{4.10}$$

This normalization is important because it leads to invariance to changes in the contrast of the image. This can be shown with the following: let $g'(x,y) = cg(x,y)$, where c is a non-zero scaling factor. Given that the DCT is linear, the normalized DCT of g' is

$$\tilde{\mathcal{T}}[g'](u,v) = \frac{c|\mathcal{T}[g](u,v)|}{c\sum_{(u,v)}|\mathcal{T}[g](u,v)|} = \tilde{\mathcal{T}}[g](u,v) \ , \tag{4.11}$$

which implies that the normalized DCT is contrast invariant and any measure based on this transform as well.

For illustrating the nature of blurring and the behavior of the DCT we take the red channel from a sharp RGB fundus image (because it resembles more to the NIR image) and simulate the imaging system as a linear shift-invariant system to acquire a sequence of images by varying the lens

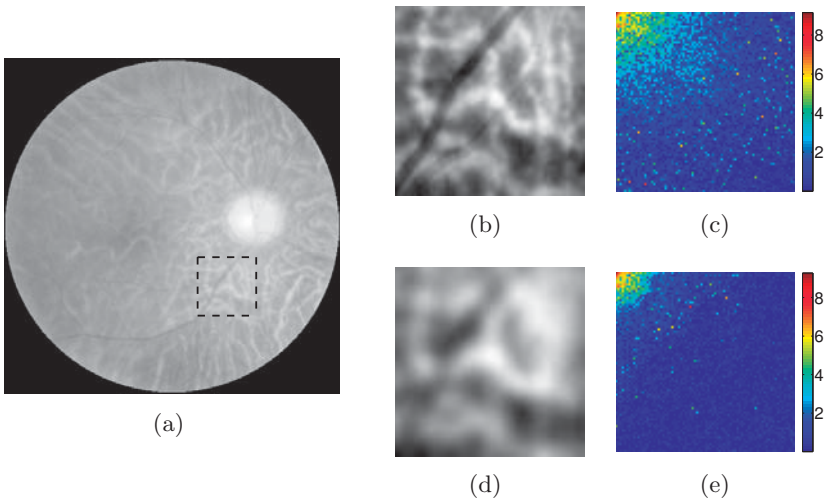


Figure 4.3: (a) Original sharp fundus image (R channel from RGB fundus image). (b) ROI from sharp image and (c) its DCT spectrum. (d) ROI from blurred image and (e) its DCT spectrum. For visualization purposes both spectra are shown in log scale. Coefficients with the highest values are shown in red and those with the lowest values are shown in blue. The blurred image spectrum is dominated by low order coefficients.

position. This was carried out by means of Fresnel propagation. In Fig. 4.3 we show the original sharp image, image patches of both the sharp and blurred images, and their DCT spectra (in the same log scale). Notice how the spectrum changes, there is less high and mid frequency content in the blurred image spectrum. In addition, in the original spectrum there are some favored orientations in the mid and low frequency coefficients, while in the blurred spectrum they seem to become more uniformly distributed. Another important feature is that in the blurred spectrum the coefficients related to high frequency have decreased significantly, and, as described in Section 4.2, many FMs are actually based on the notion of emphasizing high frequencies. While this may be true in theory, in practice there will always be noise contributing to the high frequency content due to different acquisition conditions. Furthermore, given that the focusing mechanism involves acquiring a sequence of images, there will be spatial and temporal variations of noise.

4.3.2 A Measure of Anisotropy

As we have seen in the previous example, the overall nature of blurring can be described as a low-pass filtering that tends to break down the characteristic anisotropy of the original image. The FM proposed here aims to quantify this anisotropic dependence based on the normalized DCT of the

image.

To define our measure we shall introduce the notation. From Eq. (4.10) $\tilde{G}(u,v)$ is the normalized DCT of $g(x,y)$ of size $N \times N$, and λ_j, for $j = 1,2,3$, is a vector along one of the three main orientations of the spectrum depicted in Fig. 4.4. We will restrict our study to angular partitions of the spectrum roughly equivalent to vertical, diagonal and horizontal components of the image space. Our measure of anisotropy mainly consists in calculating a difference of weighted coefficients along these orientations. Let $\tilde{\mathrm{G}}_{\mathrm{j}} = \{\tilde{G}(u,v) : \theta = \arctan\left(\frac{v}{u}\right), \theta_j \leq \theta < \theta_{j+1}, j = 1,2,3\}$ be the set of DCT coefficients located between θ_j and θ_{j+1} angles, for $\theta_j \in \{0°, 30°, 60°, 90°\}$. The function $\psi_{\lambda_j}(.)$ takes as input $\tilde{\mathrm{G}}_{\mathrm{j}}$, performs orthogonal projection of all its elements along vector λ_j and averages the elements that after projection fall on the same discrete (u,v) coordinates. With $\psi_{\lambda_j}(.)$ we seek to compact the information around the three main orientations in a one dimensional vector of N elements. To illustrate, lets compute $\psi_{\lambda_1}(\tilde{\mathrm{G}}_1) = \left[\psi^1_{\lambda_1}, \psi^2_{\lambda_1}, \ldots, \psi^N_{\lambda_1}\right]^T$, where $\tilde{\mathrm{G}}_1$ is the set of coefficients located between $\theta_1 = 0°$ and $\theta_2 = 30°$. In Fig. 4.4(b) we show the projection of the coefficient with coordinates $(4,2)$ along λ_1. After projection this coefficient has coordinates $(4,1)$. Therefore, the element $\psi^4_{\lambda_1} = \mathrm{mean}[\tilde{G}(4,1), \tilde{G}(4,2)]$. Consequently, we can stack all ψ_{λ_j} to form the following matrix,

$$\mathbf{\Psi} = \begin{bmatrix} \psi^1_{\lambda_1} & \psi^1_{\lambda_2} & \psi^1_{\lambda_3} \\ \psi^2_{\lambda_1} & \psi^2_{\lambda_2} & \psi^2_{\lambda_3} \\ \vdots & \vdots & \vdots \\ \psi^N_{\lambda_1} & \psi^N_{\lambda_2} & \psi^N_{\lambda_3} \end{bmatrix} .$$

Note that the first element of each vector corresponds to the dc coefficient. This coefficient does not convey any directional information of the image, however we decided to keep it in the matrix for the sake of completeness. To obtain a measure of anisotropy –the FM itself– from $\mathbf{\Psi}$ we compute the variance of the weighted sum of the columns, computed as the matrix product $\mathbf{w\Psi}$,

$$S_a(g) = \mathrm{Var}(\mathbf{w\Psi}) = \mathrm{E}\left[(\mathbf{w\Psi} - \mu)^2\right] \quad , \tag{4.12}$$

where $\mathbf{w} = [w_1, w_2, \ldots, w_N]$, E is the expected value and μ is the mean of the matrix product $\mathbf{w\Psi}$. The vector $\mathbf{w}$ can be regarded as a weighting procedure and with it we aim to achieve robustness to noise and illumination variation.

DCT coefficient weighting

The first issue to address is the selection of a suitable $\mathbf{w}$. In DCT-based pattern recognition, robustness is achieved by means of coefficient truncation (Lian & Er, 2010). It is known that low frequencies are related to illumination variation and smooth regions, and high frequencies represent

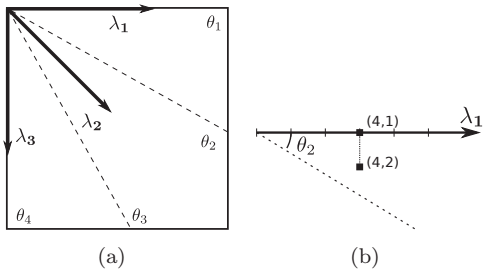


Figure 4.4: (a) Vectors along the main directions of the DCT and (b) projection of a coefficient along λ_1.

noise as well as rapid variations (like edges and details) of the image. The middle frequency coefficients contain useful information of basic structure, therefore these are suitable candidates for recognition (Chen et al., 2006). Consequently, a trade-off between low frequency and high frequency truncation should be achieved to obtain a robust FM that is monotonic with respect to blur, unimodal, and at the same time robust to noise and illumination variations.

We decided to find a $\mathbf{w}$ that meets our requirements based on a training set of m images. This can be formulated as an optimization problem. The goal would be to find the vector $\mathbf{w} = [w_1, w_2, \ldots, w_N]$ that simultaneously optimizes K objective values $\{J_1(\mathbf{w}), J_2(\mathbf{w}), \ldots, J_K(\mathbf{w})\}$. Every objective value $J_k(\mathbf{w})$ is formulated so that the FM S_a decreases with respect to blur, $S_a(g_i^k) > S_a(g_{i+1}^k)\ \forall i = 1, \ldots, m$. There are K subsets of $g_i(x, y)$ all generated in the same way as described in Eq. 4.1, but they differ in that every k stands for a different kind of noise degradation, except for $k = 1$ the noise free case. In other words, we want to find a $\mathbf{w}$ that guarantees monotonicity of S_a with respect to blur under different types of noise. The objective values are implicitly defined in terms of permutations of the ordered set $H = \{S_a(g_1), S_a(g_2), \ldots, S_a(g_m)\}$. Thus, the reference permutation is $\pi_r = \{1, 2, \ldots, m\}$, and any other arbitrary permutation of H violates the decreasing property of S_a with respect to blur. As a result, our goal is to find a $\mathbf{w}$ that produces permutations π_k for all K types of noise equal to that of π_r. The objective value is defined as the l_1-norm of the difference between π_r and π_k,

$$J_k(\mathbf{w}) : \sum_{j}^{m} |\pi_r(j) - \pi_k(j)| \quad . \tag{4.13}$$

It is zero for two identical permutations, and approaches zero as π_k approaches π_r. This is the same for all $J_k(\mathbf{w})$, hence our single aggregate objective function (Suppapitnarm et al., 2000) is the weighted linear sum of all $J_k(\mathbf{w})$, where all weights are equal to 1.

The solution to this problem is not a straightforward task as the search

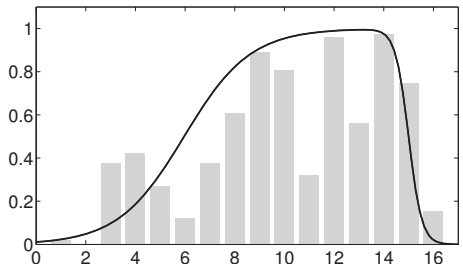


Figure 4.5: DCT coefficient weights obtained from the optimization procedure. The distribution resembles a bandpass filter.

space is multivariate and a unique global optimum cannot be guaranteed to exist. Therefore, we solved it using a probabilistic meta-heuristic approach called simulated annealing (Granville et al., 1994). It provides an acceptably good solution in a fixed amount of time. Each step of the algorithm replaces the current solution by a random nearby solution, chosen with a probability that depends both on the difference between the corresponding function values and also on a global parameter T (called the temperature), that is gradually decreased during the process. The dependency is such that the current solution changes almost randomly when T is large, but increasingly downhill as T goes to zero (For further details see Ref. (Suppapitnarm et al., 2000)).

4.3.3 Implementation

A diagram that summarizes the algorithm is shown in Figure 4.6. In order to reduce the computation time, the FM is applied to a small region of the image that contains retinal structures. This window is transformed to the DCT domain and a weighted directional sampling procedure is carried out, as previously described in § 4.3.2. The output is what we call a measure of anisotropy or a directional variance from the vectors. To achieve real time computation we decided to implement our measure by dividing the focusing window into sub-windows, to which the same procedure is applied. The measure is computed in the following manner:

1. The focusing window is divided into non-overlapping sub-images of size 16×16. This is chosen so that the most basic structures of the image fit in the sub-windows.

2. Each sub-window image is transformed with the normalized DCT and the FM S_a is computed.

3. An overall FM $\bar{S_a}$ is computed by taking the mean of all S_a values from the sub-windows.

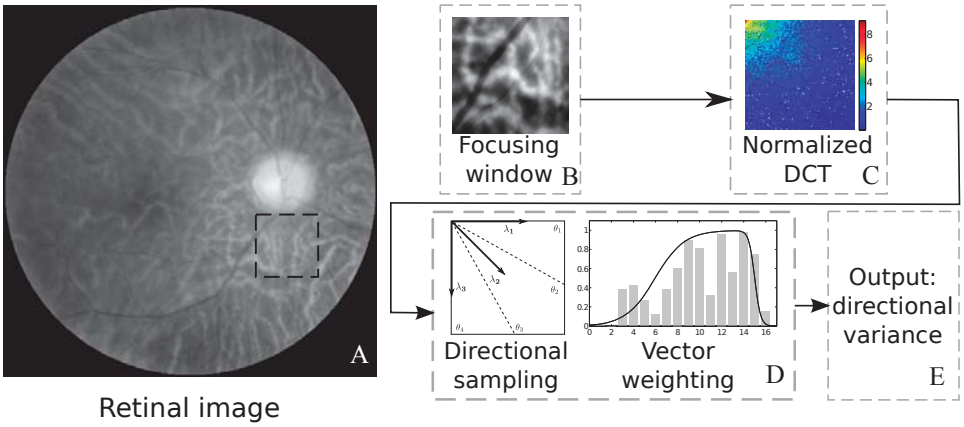


Figure 4.6: Block diagram illustrating the focus measure algorithm.

According to this implementation, the parameter $\mathbf{w}$ consists of 16 elements. The considered noise degradations for the procedure described in § 4.3.2 are: Gaussian noise ($\sigma^2 = 0.001$), Speckle noise ($\sigma^2 = 0.001$) and Impulsive noise ($d = 0.01$). The resulting $\mathbf{w}$ is shown in Fig. 4.5. As expected the first two coefficients are practically zero. Observe the distribution of $\mathbf{w}$ instead of the individual values per coefficient. This means that a strong emphasis should be put to mid-frequency coefficients. It is perhaps not surprising that the distribution resembles a bandpass filter. This finding is consistent with the work of Subbarao et al. (1993) where they showed that bandpass filtering causes the FMs to have sharp peaks while retaining monotonicity and unimodality. Interestingly, these weights also resemble the band pass response of the contrast sensitivity function of the human visual system. In the DCT domain different approaches have been considered for computing visually optimized coefficients for a given image (Watson, 1994).

A major feature of our approach is the fast computation of the FM. The average execution time per frame, in MATLAB implementation in a PC with a 2.66 GHz Intel Core 2 Duo processor, is 40 milliseconds. In most cases this is sufficient, however if needed, implementation in a low level programming language could significantly reduce the execution time. In addition, because we divide the focusing window into sub-windows, our implementation could be further improved by taking advantage of large parallel architectures such as in GPU (Graphics Processor Unit) computing.

Experimental setup

The experimental set-up consisted mainly of an AF mechanism attached to a digital fundus camera system (TRC-NW6S, Topcon, Tokyo Japan) showed in Figure 4.7. The AF apparatus consisted of in-house assembled stepper motor mechanism for the displacement of the compensation lens controlled

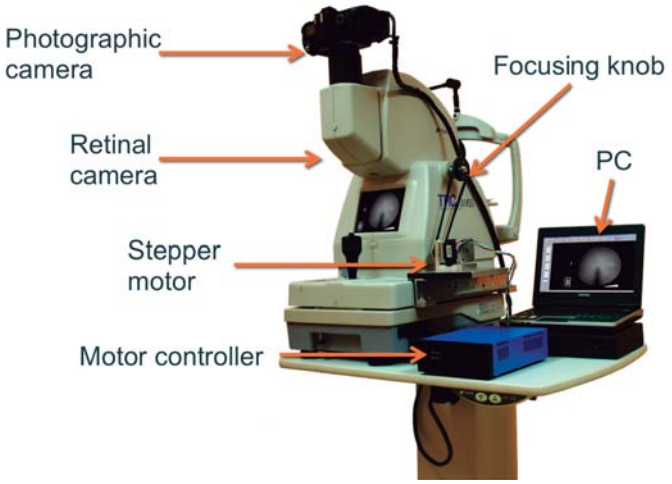


Figure 4.7: Experimental setup.

Figure 4.8: The autofocus system in operation while examining a subject.

via RS232 with a PC. This mechanism was coupled to the fundus camera. The image acquisition and processing, along with the motor control, were carried out in MATLAB. The images were acquired from the video output of the infra-red focusing system with a resolution of 640×480. The fundus camera focusing system enables a compensation range of $-13D : 12D$ (D stands for diopters) in normal operation. For strong myopia or hyperopia two additional compensation lenses are available to compensate the ranges $-12D : -33D$ and $+9D : +40D$, respectively. The image sequences analyzed in this chapter were acquired for the normal operation range. In Figure 4.8 we show the assembled autofocus system in operation while examining a subject.

4.4 Results

4.4.1 Simulated images and robustness assessment

To evaluate the robustness of our proposed FM S_a we have simulated the focusing procedure. We generate a sequence $g_i(x, y)$ for $i = 1, \ldots, m$ from the red channel of a sharp RGB fundus image and propagate it at different distances through a linear imaging system of fixed focal length by means of Fresnel propagation. This is equivalent to displacing the lens or the sensor to look for the optimal focus position. From this noise free sequence we generate 6 additional sequences by corrupting it with 2 levels of 3 different types of noise: Gaussian, speckle, and impulse noise. We carried out this procedure for 20 retinal images for a total of 140 focusing sequences. Ideally, a noise robust FM should produce the same (or similar) focusing curve for both the noise free and the corrupted sequences. To quantify the similarity between two focusing curves S_r and S_c we used the zero-lag normalized cross correlation defined as

$$R(S_r, S_c) = \frac{\sum_i S_r(i) \cdot S_c(i)}{\sqrt{\sum_i {S_r}^2(i) \cdot \sum_i {S_c}^2(i)}} \ , \tag{4.14}$$

where r stands for the reference curve computed from the noise free sequence, and c the curve computed from the noise corrupted sequence. The output is 1 in the case of perfect correlation and 0 for no correlation at all. The reason for the zero-lag calculation, as opposed to the regular cross correlation by sliding dot product, is that we need the maxima of the curves to coincide at the horizontal position in addition to the matching of the curves solely by shape.

All FMs were computed using a focusing window of 128×128 pixels located over retinal structures. In Fig. 4.9 we show an example to illustrate the robustness assessment of the FMs. The FM curves represent the normalized measure value over the search space for different lens positions. The highest value should be obtained when the lens is on the optimal focus position identified by the dashed vertical line, which was verified via the split-line focusing mechanism. As the lens gets farther from the optimal position the measure value should decrease with the distance. It comes as no surprise that all measures perform sufficiently well in the noise free sequence shown in Fig. 4.9(a), where all curves follow a typical bell shape with a unique maximum. However, in the curves shown in Figs. 4.9(b)-(d) where the focusing sequence is corrupted by different types of noise, the proposed FM S_a clearly outperforms the other measures in terms of monotonicity and unimodality. Notice that under Gaussian and speckle noise (Figs. 4.9(b)-(c)) the S_a curves are nearly identical to the noise free S_a curve in Fig. 4.9(a). Without jumping to conclusions this result is interesting because it graphically shows the

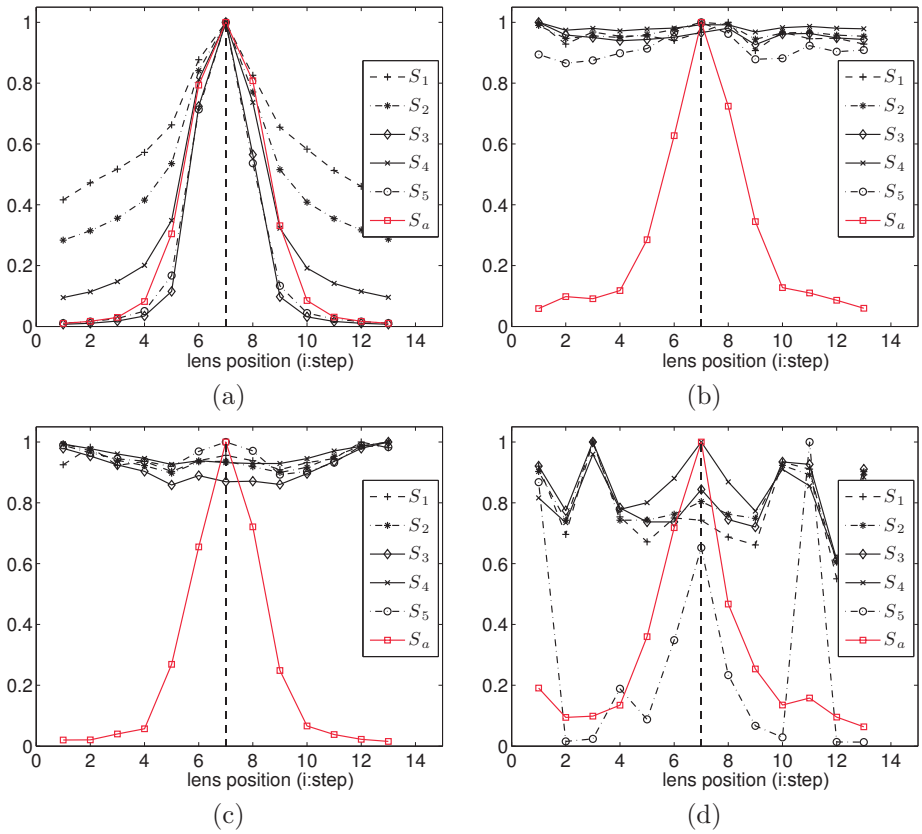


Figure 4.9: Focus measures curves for the simulated images, the dashed vertical line indicates the correct in-focus position. (a) Noise free images; images corrupted with: (b) Gaussian noise ($\sigma^2 = 0.001$), (c) Speckle noise ($\sigma^2 = 0.001$) and (d) Impulsive noise ($d = 0.01$).

robustness of the proposed FM. The results for all 140 sequences are summarized in Table 4.1. Each value represents the average cross-correlation obtained for all 20 sequences corrupted with a specified type and level of noise for a particular FM. The overall average for each FM is shown in the last column. These results provide further evidence that the proposed FM S_a has a considerable robustness to noise with an overall performance value of 0.929 and an exceptional 0.996 for the sequence corrupted with Gaussian noise with $\sigma^2 = 0.001$. The second and third best FMs were S_4 and S_1, with overall values of 0.781 and 0.502 respectively. In comparison with S_a these values represent a moderate to mild noise robustness. In the following section we use these two FMs to compare with S_a in real images.

	Gaussian (0.001*)	**Gaussian** (0.005*)	**Speckle** (0.001*)	**Speckle** (0.005*)	**Impulse** (0.01**)	**Impulse** (0.05**)	**Overall Average**
S_1	0.554	0.486	0.635	0.422	0.477	0.438	**0.502**
S_2	0.524	0.499	0.468	0.408	0.476	0.462	0.473
S_3	0.449	0.444	0.370	0.359	0.420	0.417	0.410
S_4	0.784	0.782	0.750	0.746	0.836	0.791	**0.781**
S_5	0.495	0.380	0.495	0.304	0.795	0.362	0.472
S_a	0.996	0.939	0.997	0.992	0.979	0.667	**0.929**

Table 4.1: Average normalized cross correlation results for noise robustness assessment of focus measures from 140 sequences generated from 20 retinal images corrupted with different types and levels of noise. (* : Standard deviation σ^2, ** : Noise ratio d)

4.4.2 Real images

In this subsection we show the results obtained from real NIR focusing eye fundus images. The images have a relatively low SNR which justifies the need for a robust FM.

It is a well known fact that as a person ages the crystalline lens of the eye gradually gets opacified obstructing the passage of light. This is called a cataract. A complete loss of transparency is only observed in advanced stages in untreated patients. In early stages of cataracts retinal examination is considered practicable–however, it is not without difficulty. For this reason we decided to test our focusing method on healthy young subjects and elderly subjects with first signs of cataracts not only to demonstrate its applicability on real images, but to assess its limitations as well. In this work we show results from five representative subjects with ages 27, 40, 68, 70 and 81 years old for a total number of ten eye fundi.

First we show the effects of placing the focusing window on different regions of the retinal image. A retinal image has distinct sharp structures such as the blood vessels and the optic disc, as opposed to the relatively uniform background. No FM is reliable without placing the focusing window on top of structures with edges, a fact easily appreciable from the three focusing curves shown in Fig. 4.10, which were computed from the right eye fundus of the 27-year-old subject. The S_a curves computed from the regions (b) and (c) are clearly reliable in terms of monotonicity and unimodality, and coincide on the optimal focus position. Conversely, the S_1 and S_4 curves when compared against S_a failed to produce a profile in which the maximum is easy to identify. In that regard, the S_a curves display a steeper peak at the optimal focus position, evidence of the measure's robustness to noise. In contrast, all measures computed from region (d) are unusable because they are mainly given by noise.

To illustrate the link between the focusing curves and the image qual-

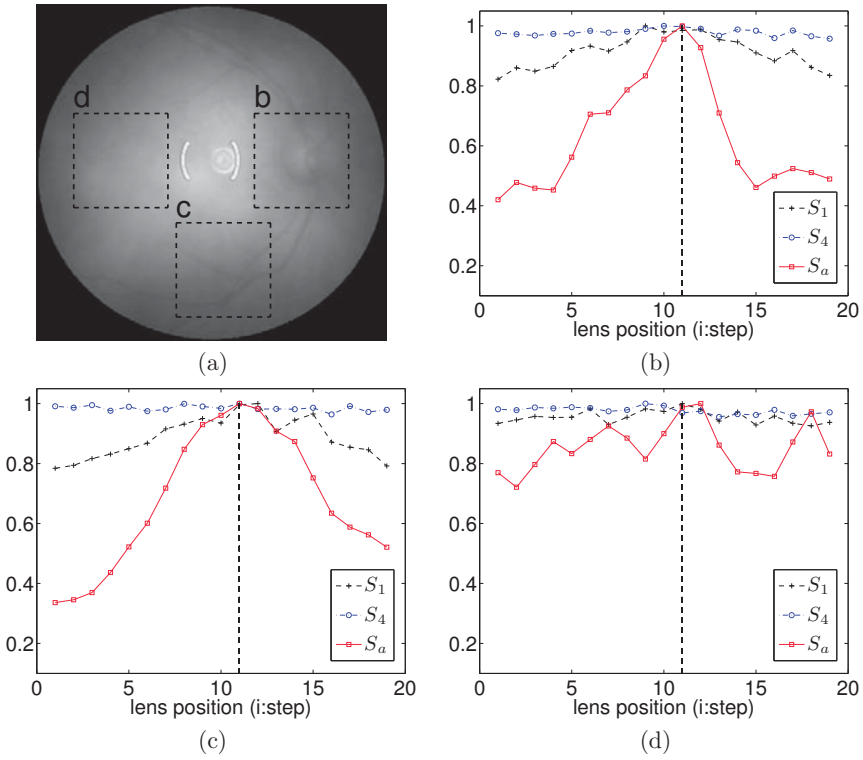


Figure 4.10: Focus measure curves obtained by placing the focusing window over different regions of the retinal image (a): (b) and (c) are located over prominent retinal structures, whereas (d) is located over a relatively uniform region. The dashed vertical line indicates the correct focused position in (b)-(d).

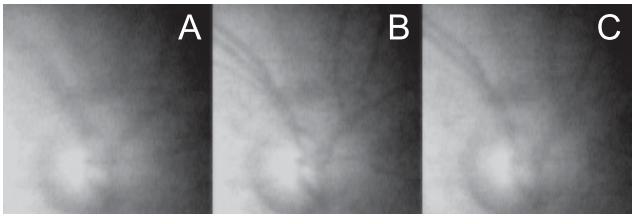


Figure 4.11: Image detail from Fig. 4.10 for different focusing positions: (a) 6 (b) 11 (optimal focus), and (c) 15. The positions are in reference to Figures 4.10(b)-(c).

ity, in Fig. 4.11 we show three image details depicting the optic disc region for three different focusing positions. The image detail in Fig. 4.11(b) corresponds to the focused image (optimal focus position 11 in the S_a curve Fig. 4.10). Notice how this image is properly focused, it has sharp details like the blood vessels. The other two images are blurred demonstrating the consistency of the S_a curves with image quality or sharpness. The result that emerges from this example is that to effectively locate the best focused image, homogeneous regions should be avoided. An adaptive technique, based e.g. on an edge detector, could prove useful for detecting such prominent structures and therefore candidate regions for applying automatically the focusing technique. The focusing curves shown hereafter, however, were all computed from a focusing window located manually over retinal structures.

To further analyze the performance of the FM in Fig. 4.12 we show the focusing curves obtained from four of the five subjects; the ages are shown in the figure caption. In general, from the comparison against S_1 and S_4 it can clearly be stated that the proposed FM S_a outperforms them in the considered cases. From the four cases shown only in one (Fig. 4.12(c)) the S_a measure peak did not coincide precisely with the optimal focus position. However, the error is no more than a single position. The FMs curves of S_1 and S_4 are generally flatter than those of S_a which in a focus search strategy is not wanted because of the difficulty to properly distinguish the optimum position in a coarse or initial search. From the curves in Fig. 4.12 we can also note that there appears to be little difference between the curves from young and elderly subjects. In Fig. 4.13 we show the focusing curves obtained from the 81-year-old subject for both eye fundi. This case is interesting on its own because in the right eye (Fig. 4.13(a)) the crystalline lens has been extracted and replaced with an intraocular lens, whereas the left eye (Fig. 4.13(b)) is in an early stage of cataract. While both focusing curves are able to successfully identify the optimal focus position, the curve in Fig. 4.13(b) is certainly flatter throughout most of the search space. This is most likely due to the difference in visibility and clarity from both eyes.

A close examination of the results reveal that the shape of the focusing curve is not exclusively given by the degree of defocus, but by the subject's state of the eye and the analyzed region of the fundus as well. This is

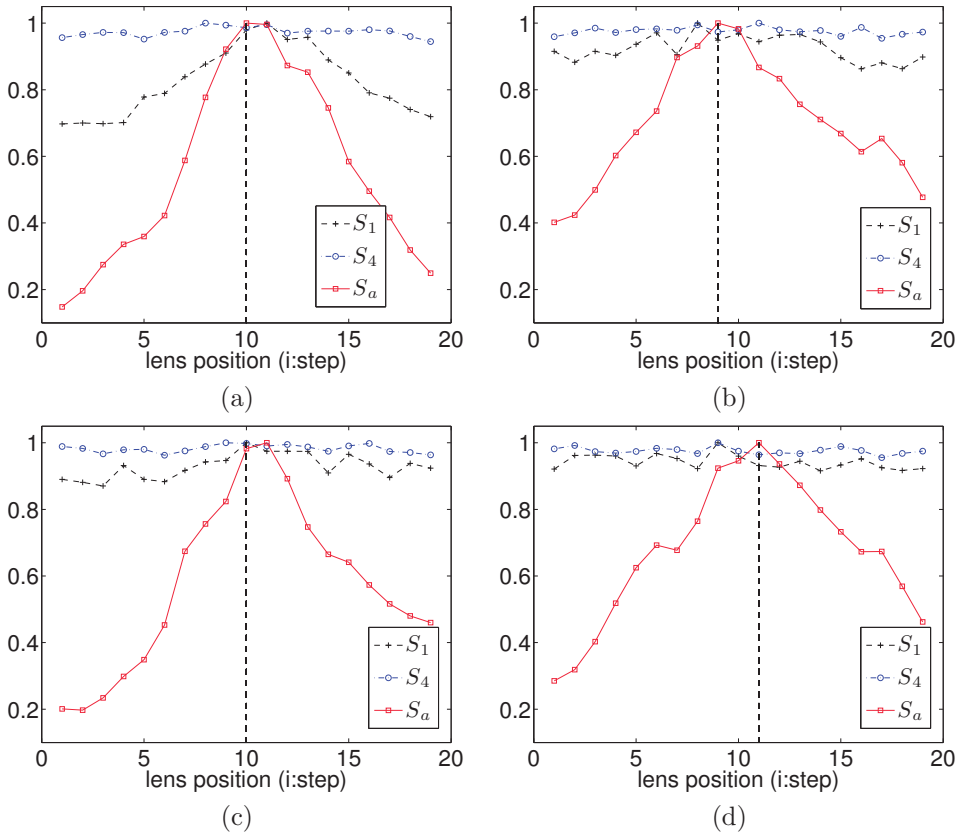


Figure 4.12: Focusing curves obtained from four subjects with ages (a) 27, (b) 40, (c) 68, and (d) 70 years old. The dashed vertical line indicates the correct focused position.

important because it conditions the strategy for searching the optimal focus position (Marrugo et al., 2012a). Finally, even though the results seem to indicate that the FM could be successfully applied to both young and elderly subjects, further research on a higher number and variety of subjects is necessary. Additionally, we report here that we encountered some difficulty in the procedure with the elderly subjects related to sustaining fixation during the acquisition procedure. From an initial number of six subjects one was excluded from all calculations due to this condition. Patient inability to successfully establish fixation is a true challenge in fundus photography and dealing with it is out of the scope of this work.

4.5 Discussion

In this chapter we have introduced a new focus measure for non-mydriatic retinal imaging. It is based on a measure of anisotropy, mainly the weighted

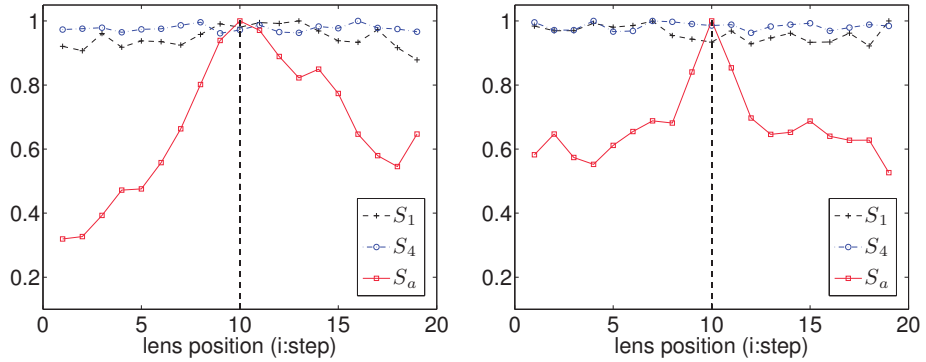


Figure 4.13: Focusing curves obtained from the 81-year-old subject for each eye fundus. In the (a) right eye the crystalline lens has been extracted and replaced with an intraocular lens, the (b) left eye is in an early stage of cataract. The dashed vertical line indicates the correct focused position.

directional variance of the normalized discrete cosine transform. The weights were calculated by means of an optimization procedure to maximize the noise robustness of the focus measure. Not only were the resulting weights in agreement with previous works (Subbarao et al., 1993), but they also provide a key insight into the design of noise invariant focus measures. Both by simulation and real fundus imaging we demonstrated the robustness and the accuracy of the novel focus measure, clearly outperforming the other considered measures. The findings presented here may have a number of implications for the design and operation of autofocusing in modern retinal cameras. Finally, in this study we included several young and elderly subjects to assess the limitations of the proposed focus measure. Even though we found no significant differences between the focusing curves, there was some difficulty in the acquisition of images from the elderly mainly given by inability to sustain fixation. As with all such studies, there are limitations that offer opportunities for further research. Adapting our method to these variations within the patient population is a goal worth pursuing.

Chapter 5

Deblurring Retinal Images and Longitudinal Change Detection

5.1 Introduction

Blur is one of the main image quality degradations in eye fundus imaging, which along with other factors such as non-uniform illumination or scattering hinder the clinical use of the images. In this chapter we focus on the enhancement of blurred retinal images. A *successful* retinal image restoration with no knowledge of the *point-spread-function* (PSF) that blurred the image is rather recent in the literature (Chenegros et al., 2007). Besides, results on real images have been rarely produced. This deblurring problem is referred to as *blind deconvolution* (Kundur & Hatzinakos, 1996, Campisi & Egiazarian, 2007), in which the goal is to recover a sharp version of the input blurry image when the blur kernel is unknown. Mathematically, we wish to decompose a blurred image $z(x, y)$ as

$$z(x, y) = u(x, y) * h(x, y) = \int u(s, t)h(x - s, y - t)\,\mathrm{d}s\mathrm{d}t \ , \qquad (5.1)$$

where $u(x, y)$ is a visually plausible sharp image, $h(x, y)$ is a non-negative linear shift blur kernel (PSF), and $*$ is the two-dimensional linear convolution operator. This problem is severely ill-posed* and there is an infinite set of pairs (u, h) explaining any observed z. For example, one undesirable solution that perfectly satisfies Eq. (5.1) is the no blur explanation: h is a delta (identity) kernel and $u = z$. The ill-posed nature of the problem implies that additional assumptions on u and h must be introduced.

*Ill-posed problems, *Encyclopedia of Mathematics*, URL: `http://www.encyclopediaofmath.org/index.php/Ill-posed_problems`

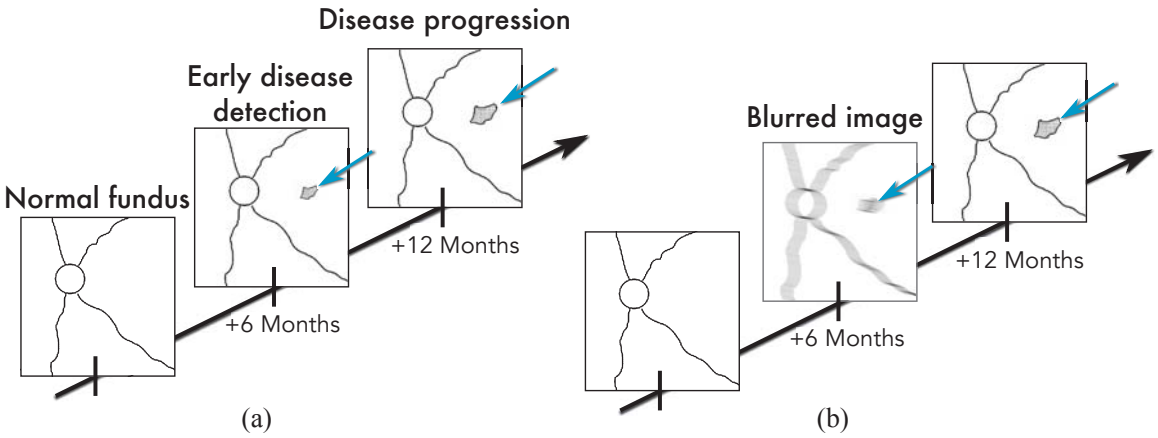


Figure 5.1: (a) Patients receive regular examination either for early disease detection or disease-progression assessment. (b) Low-quality image occurrence is not uncommon. Archival blurred images are of little clinical use unless they are enhanced.

It has been shown that a way to mitigate this ill-posedness is to have several images of the same scene blurred differently (Sroubek & Flusser, 2003). This better-poses the problem in that there is redundant information and makes it more robust to noise (Sroubek & Milanfar, 2012). Taking advantage of this approach, in a recent publication (Marrugo et al., 2011a) we proposed a novel strategy for retinal image deblurring where we considered a general image degradation scenario: a sequence of blurred retinal images acquired with different time lapses (ranging from minutes to months), which meant that disease progression was also a key factor to consider.

The reason for considering images acquired with such long time-lapses comes from the fact that a correct assessment of a patient's state evolution requires sharp images from all moments in time (see Fig. 5.1). In other words, because single-image deblurring is—to date—a difficult problem to tackle successfully, we set out to deblur retinal images from a patient's archive so that we could readily have several images of the same scene (eye fundus) for implementing a multi-image deblurring strategy. However, equally important to the image enhancement, we have developed a longitudinal change-detection algorithm because disease progression could manifest as structural changes in the images. This has enabled the identification and localization of changes of pathological origin in the retina.

5.1.1 Motivation and Background

Retinal imaging is acknowledged to be an important tool for the detection and the progression-assessment of diseases affecting the eye such as diabetic retinopathy, glaucoma, and age-related macular degeneration (Kanski,

2005). The digital format provides a permanent record of the appearance of the retina at any point in time (Winder et al., 2009).

The imaging procedure is usually carried in two separate steps: image acquisition and diagnostic interpretation. Image quality is subjectively evaluated by the person capturing the images and they can sometimes mistakenly accept a low quality image (Bartling et al., 2009). Low-quality image occurrence rate has been reported at 3.7-19.7% in clinical studies (Agrawal & McKibbin, 2003, Boucher et al., 2003, Herbert et al., 2003), which is not a minor fact. A recent study by Abramoff et al. (2008) using an automated system for detection of diabetic retinopathy found that from 10 000 exams 23% had insufficient image quality.

Major source of retinal image quality degradation are aberrations of the human eye, imperfections in the fundus camera optics, and improper camera adjustment, flash lighting or focusing during the exam (Larichev et al., 2001). Moreover, regardless of how well controlled the aforementioned parameters are, in practice it may not always be possible to obtain good enough image quality as a result of additional factors such as lens opacities in the examined eye, scattering, insufficient pupil dilation or patient difficulty in steady fixating a target in the camera; such as in patients suffering from amblyopia (Bartling et al., 2009). Out of all possible retinal image degradations, some can be properly compensated via enhancement or restoration techniques; e.g., low-contrast, non-uniform illumination, noise, and blur (Winder et al., 2009). However, this compensation is also dependent on the extent of the degradation. Regarding retinal image blurring its main causes are: relative camera-eye motion, inherent optical aberrations in the eye, and improper focusing.

In the past decade many wavefront technologies (that originated from astronomy) such as Adaptive Optics (AO) (Christou et al., 2004) and Deconvolution from Wavefront Sensing (DWFS) (Primot et al., 1990) gave rise to the correction of monochromatic aberrations of the eye and also created new opportunities to image the retina at unprecedented spatial resolution. However, AO-corrected and DWFS-based fundus imagers usually aim at resolving details at the level of individual photoreceptors, thus have a Field of View (FOV) of a couple degrees and a high resolution in the order of 1 or 2 microns (Catlin & Dainty, 2002). Despite the fact that greater FOVs can be achieved ($\sim 5^\circ$) (Yang et al., 2008, Burns et al., 2007), with additional hardware constraints, diffraction limited imaging is not guaranteed due to an increase in aberrations (Bedggood et al., 2008). Nevertheless it is still a considerably narrow FOV and a major disadvantage with clinical subjects because of the need to examine larger areas of the retina. Alternatively, regular non-AO-corrected fundus imagers used for routine check-ups have a large FOV (typically 30°) at the expense of lower spatial resolution, but still sufficient for practical detection and progression-assessment of observable clinical signs such as, microaneurysms, dot and blot hemorrhages,

exudates, among others. Consequently, large FOV fundus imagers are the major imaging modality available to patients visiting an eye-care clinic. All images analyzed in this chapter were acquired with a conventional large FOV fundus imager.

5.1.2 Contribution

REFERENCE TO THE PUBLICATIONS OF THIS THESIS

The content of this chapter is included in the publications:

A. G. Marrugo, Michal Šorel, Filip Šroubek, and María S Millán, "Retinal image restoration by means of blind deconvolution", in *Journal of Biomedical Optics*, **16**(11):116016, (2011).

A. G. Marrugo, Filip Šroubek, Michal Šorel, and María S Millán. "Multichannel blind deconvolution in eye fundus imaging", In *ISABEL '11-Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*, **7**:1–5. NY, USA, (2011).

A. G Marrugo, María S Millán, Gabriel Cristóbal, Salvador Gabarda, Michal Šrorel, and Filip Šroubek. Image analysis in modern ophthalmology: from acquisition to computer assisted diagnosis and telemedicine **Invited Paper**. In *SPIE Photonics Europe, Proceedings SPIE*, **8436**:84360C, (2012).

A. G Marrugo, María S Millán, Gabriel Cristóbal, Salvador Gabarda, Michal Šorel, and Filip Šroubek, "Toward computer-assisted diagnosis and telemedicine in ophthalmology", *SPIE Newsroom*, (doi: 10.1117/2.1201205.004256), (2012).

In this chapter our novel contributions to the retinal image processing task are twofold. First, we propose a degradation model for time-series retinal images, which captures the underlying distortions resulting from instrument limitations and changes between patient visits; we are also able to identify and highlight such changes. Second, we propose a restoration strategy based on blind deconvolution that is able to obtain image enhancement and resolution improvement using inexpensive digital methods applied to images acquired with a conventional fundus camera.

5.2 The Blind Deconvolution Problem

The goal of blind deconvolution is to recover the original (or unblurred) scene from a single or a set of blurred images in the presence of a poorly determined or unknown PSF. The main assumption is that the blur can be described by a convolution of a sharp image with the unknown PSF (Eq. (5.1)). Restoration by deconvolution improves contrast and resolution of digital images, which means that it is easier to resolve and distinguish features in the restored image. To avoid confusion with super-resolution, we explain what we mean by *resolution improvement*. Digital deconvolution can be described as any scheme that sharpens up the PSF while the spatial frequency bandwidth remains unchanged. This means that the spatial frequency response

and the two-point resolution is improved but the cut-off frequency is unchanged (Sheppard, 2007). In the super-resolution context the goal is to increase the cut-off frequency.*

5.2.1 The Multichannel Approach

Blind deconvolution algorithms can be single input (single-image blind deconvolution SBD) or multiple input (multi-channel blind deconvolution MBD). SBD is truly a complicated problem in that it is an under-determined inverse problem as there are more unknowns (image and blur) than equations. For a long time, the problem seemed too difficult to solve for general blur kernels. Past algorithms usually worked only for special cases, such as astronomical images with uniform (black) background, and their performance depended on initial estimates of PSFs (Sroubek & Milanfar, 2012). Despite the fact that SBD is one of the most ill-posed problems, there are several reliable SBD algorithms (Levin et al., 2011), although most of them require that the blurred image be governed by relatively strong edges or that the blur is exclusively caused by motion. In either case, insofar as retinal imaging is concerned, these assumptions do not hold and these methods would likely fail (we show this in § 5.5.1).

As regards MBD its robustness lies in the requirement of multiple images of the same scene blurred in a slightly different way, which adds information redundancy and better poses the problem. In many practical applications it is certainly difficult to obtain multiple images with these requirements. However, retinal images are special in that the acquisition conditions change relatively little between patient visits (even with a time span of months), and most importantly the retina or the retinal features like blood vessels—the scene itself—is *highly stable* in time. Most common diseases do not change the distribution of the blood vessels in a way that its topology is affected. From this consideration we have hypothesized that a pair of fundus images of the same retina, acquired at different moments in time, contain enough common information for their restoration via existing multi-channel deconvolution techniques.

A block diagram illustrating our proposed method (Marrugo et al., 2011a) is shown in Fig. 5.2. The input are two color retinal images acquired with a conventional fundus camera within a time lapse that can span from several minutes to months given by routine patient check-ups, as illustrated in Fig. 5.1. The images correspond to the same retina, but can differ to some extent with respect to illumination distribution, object field and perspective, blur, and local structural changes of possible pathological origin. These differences cannot solely be accounted for by the convolutional model described in Section 5.3. For that reason the images have to be preprocessed

*For an extensive discussion on the topic see (Sheppard, 2007).

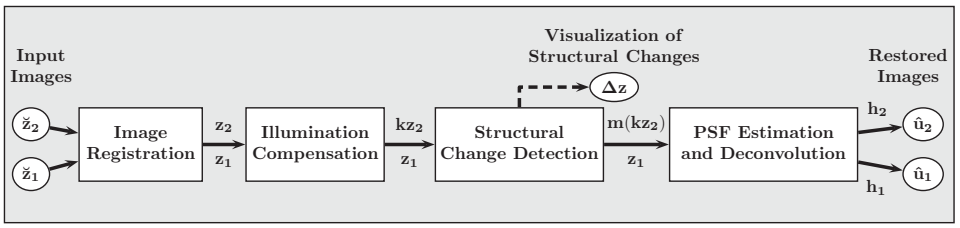


Figure 5.2: Block diagram illustrating the proposed method. $\check{z}_i$ are the unregistered degraded input images and $\hat{u}_i$ are their restored versions. The other variables are intermediate outputs of every stage, their meaning is given in the text.

before the blind deconvolution stage can take place. This consists in image registration to properly align the images, compensation of both inter-image illumination variation and structural changes. In fact, this preprocessing work is a great opportunity to meet one of the main concerns of ophthalmologists when they visually compare fundus images of the same retina over time. It enables the identification of *true* structural or morphological changes pertaining to possible pathological damage and consequently disregarding other changes merely caused by variation of illumination or blur. Once the input images have been preprocessed they are ready for blind deconvolution. Our blind deconvolution strategy is a two-stage one. The first stage consists in the estimation of the PSFs following a multi-channel scheme and the second stage is the image deconvolution, where every image is restored with its corresponding PSF independently.

The multi-channel scheme is based on the work by Sroubek & Flusser (2005), which has proved to work well in practice with sufficient experimental data. It is an alternating minimization scheme based on a maximum *a posteriori* (MAP) estimation, with *a priori* distribution of blurs derived from the multichannel framework, and *a priori* distribution of the ideal sharp image defined by regularization with the total variation of the image (Rudin et al., 1992). MAP is formulated as an optimization problem, where regularization terms are directly related to priors. Regularization involves the introduction of additional information in order to solve an ill-posed problem in the form of a penalty or restriction in the minimization routine (See Section 5.4.4). This provides good quality of restoration—significantly better than for example Lucy-Richardson algorithm (Richardson, 1972) still widely used in biomedical applications. We have modified the algorithm by Sroubek & Flusser (2005) to leave out regions where the eye fundus has structurally changed (it only takes into account one image in these regions) with the use of a *masking operator*, similarly to the solution proposed by Šroubek et al. (2008) within the super-resolution context. This has allowed the restoration of both degraded input images.

5.3 Mathematical Model of Image Degradation

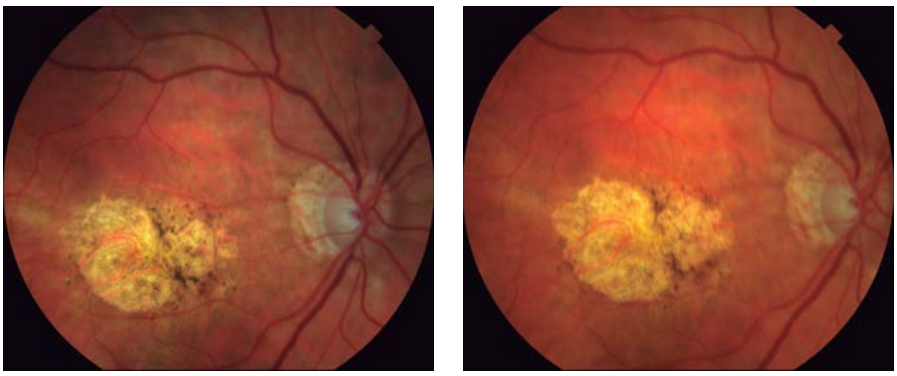
Let $\breve{z}_1$ and $\breve{z}_2$ be two unregistered degraded input images, as depicted in Fig. 5.2. After registration we obtain two degraded registered images z_1 and z_2, which we model as originating from an ideal sharp image u. Mathematically, the degradation model is defined as

$$\begin{aligned} z_1 &= u * h_1 + n_1 \\ z_2 &= \left(uk^{-1}\right) * h_2 + n_2 \ , \end{aligned} \tag{5.2}$$

where $*$ is the standard convolution, h_i are called convolution kernels or PSFs, and k is a function accounting for relative local illumination change between images z_1 and z_2 (see Eq. (5.3)). For pixels where no illumination changes occur, $k \approx 1$. The noise n_i is assumed Gaussian additive with zero mean in both images. From Eq. (5.2), the PSFs and k comprise all radiometric degradations except structural changes in the eye, which are taken into account by a masking operator defined in Section 5.4.3. Despite the fact that we consider the PSFs to vary between the two image acquisitions, we assume them to be spatially invariant within each image. Since the FOV is of 30° or less, this assumption can be accepted in first approach. Spatially variant blur is considered in Chapter 6. The ideal sharp image u is actually unknown and its estimation is the purpose of this chapter. Thus to avoid confusion the estimated (restored) image is denoted by $\hat{u}$. In Section 5.5.1 we test the performance of our method with synthetically degraded images, which means that we know the original u. This is important, because we can compare the restored image $\hat{u}$ with the original one u and, therefore assess the quality of the method.

5.4 The Deblurring Method

In this section we describe every stage of the proposed method. To illustrate each stage we use the images shown in Fig. 5.3. They were acquired using a non-mydriatic digital fundus camera system with conventional xenon flash lighting source (in the visible spectrum). The fundus images were acquired from a patient that suffered from age-related macular degeneration and were captured within a seven-month time lapse. They are color RGB 24 bit-depth fundus images of size 1500×1200 digitized in TIFF format. This is a general example were both images do not correspond exactly to the same object field, the illumination distribution across both images is not exactly the same, and there are some structural differences between them given by the pathological development in the macula (centered yellowish region).



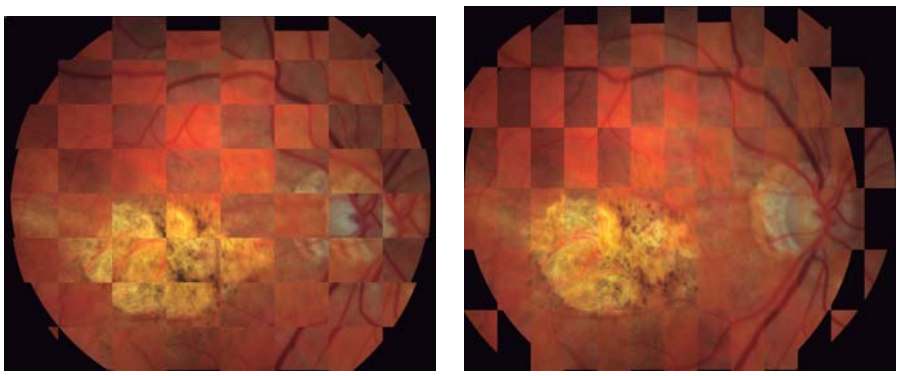
(a) (b)

Figure 5.3: Color fundus images of a human eye affected by age-related macular degeneration. Images (a)-(b) were captured within a seven-month time lapse, (a) was captured before (b).

5.4.1 Image Registration

Image registration is a procedure that consists of spatial alignment of two or more images. General and application-specific image registration, such as in retinal imaging, has been investigated from the beginning of image processing research. The interested reader is referred to the image registration review by Zitova & Flusser (2003) and the recent work by Lee et al. (2010) for objective validation of several retinal image registration algorithms. Image registration techniques are usually divided into two groups: intensity-based and feature-based methods. Intensity based methods have the drawback of poor performance under varying illumination conditions. Feature based methods are robust to such effects but rely on accurate and repeatable extraction of the features. The retinal vasculature is known to provide a stable set of features for registration in the conditions of interest.

For registering the images we use the robust dual-bootstrap iterative closest point algorithm; we describe the main features here, for a full description of the method the reader is referred to the work by Stewart et al. (2003). The vasculature from each image is automatically traced starting from initial seed points extracted from a 1D edge detection and later recursively tracking the vessels using directional templates. The vessel branching and crossover points are used as landmarks to register the images to subpixel accuracy. The registration algorithm starts from initial low-order estimates that are accurate only in small image regions called bootstrap regions. The transformation is then refined using constraints in the region, and the bootstrap region is expanded iteratively. The algorithm stops when the bootstrap region expands to cover the overlap between the images, and uses a 12-dimensional quadratic mapping. This transformation model includes



(a) (b)

Figure 5.4: Registration of images from Fig. 5.3 in checkerboard representation. (a) Before and (b) after registration.

rotation, scale, translation, a shearing term and a quadratic term that describes the spherical shape of the retina. We refer the interested reader to Ref. (Can et al., 2002) for details on the model derivation. This registration algorithm is very robust to local changes and low overlap between images as demonstrated by its high success rate on test images with at least one common landmark point and overlaps even as low as 35% (Stewart et al., 2003). Even though the reported accuracy by Stewart et al. (2003) is of up to subpixel accuracy, in our case of degraded images this can be slightly worse without compromising the outcome. Minor local misregistration errors may occur when landmark points do not match precisely, but they are not be taken into account in the restoration because they are masked out before the PSF estimation and image deconvolution stages (see Section 5.4.3).

To confirm the registration outcome the pair of images before and after registration are shown in Fig. 5.4 in checkerboard representation, where the images are merged together in a chess-like pattern where each square alternates information from one image to the other. Notice how after registration the images have been correctly aligned, specially the blood vessel distribution (see the continuous transitions of vessels between neighbor squares).

5.4.2 Compensation of uneven illumination

Despite controlled conditions in retinal image acquisition such as optical stops to prevent glares and provide a diffuse illumination, there are many patient-dependent aspects that are difficult to control and mainly affect the illumination component with gradual non-uniform spatial variations. Some of the contributing factors are:

- The curved surface of the retina. As a consequence, all regions cannot be illuminated uniformly.

- Imaging requires either a naturally or an artificially dilated pupil. The degree of dilation is highly variable across patients.
- Unexpected movements of the patient's eye.
- The presence of diseases.

The non-uniform illumination across the image results in shading artifacts and vignetting. This effect hinders both quantitative image analysis and the reliable operation of subsequent global operators.

In our model, described by Eq. (5.2), the relative changes in intensity between the two fundus images cannot be described exclusively by convolution with different PSFs and must be compensated by k. A number of general-purpose techniques have been investigated to attenuate the variation of illumination. However, most techniques are oriented towards single-image compensation (Winder et al., 2009), for instance using the red channel to estimate background illumination (Muramatsu et al., 2010). Therefore, no consistency between two images is guaranteed. For our case this uneven illumination can be compensated by properly adjusting the intensity values on one image to approximately match that of the other while satisfying a predetermined illumination model. This can be carried out if the blurring is not too large and the illumination changes smoothly, which is usually the case for fundus images. This assumption can be expressed mathematically as

$$(k^{-1} \cdot u) * h \approx k^{-1} \cdot (u * h) \ .$$

The illumination of the fundus is formed by a slowly varying light field over a smooth surface, thus it can be modeled by a low-order parametric surface. Narasimha-Iyer et al. (2006) used a 4th-order polynomial to effectively model the light pattern formed by an illumination source passing through the attenuating ocular media. Here we use a similar approach, but fitting the surface with respect to both images. The parametric surface fitting equation can then be formulated as

$$\arg\min_{k} \| z_1(x,y) - k(x,y) \cdot z_2(x,y) \| \ , \tag{5.3}$$

where $k(x,y) = \alpha_{15}y^4 + \alpha_{14}y^3x + \cdots + \alpha_2 y + \alpha_1$, and z_1, z_2 are the registered fundus images. We minimize Eq. (5.3) in the least squares sense to estimate the 15 parameters. This procedure can be both carried out using the luminance channel or the green channel as is commonplace in retinal image processing (Foracchia et al., 2005). Here we have used the green channel. Owing to the fact that the illumination can be compensated globally by the polynomial function k, it is important to realize that the structural changes remain unaffected. The interpretation of k from Eq. (5.3) is straightforward. If the registered images z_1 and z_2 had neither illumination changes

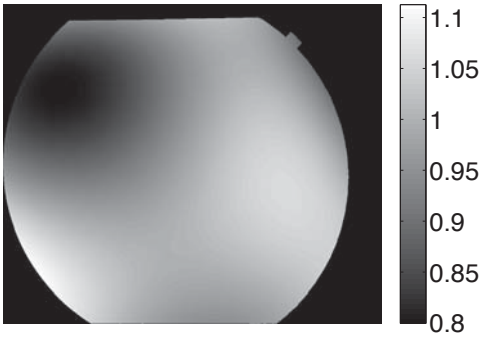


Figure 5.5: Illumination compensation function $k(x, y)$. This is applied only within the region of interest (the central circle).

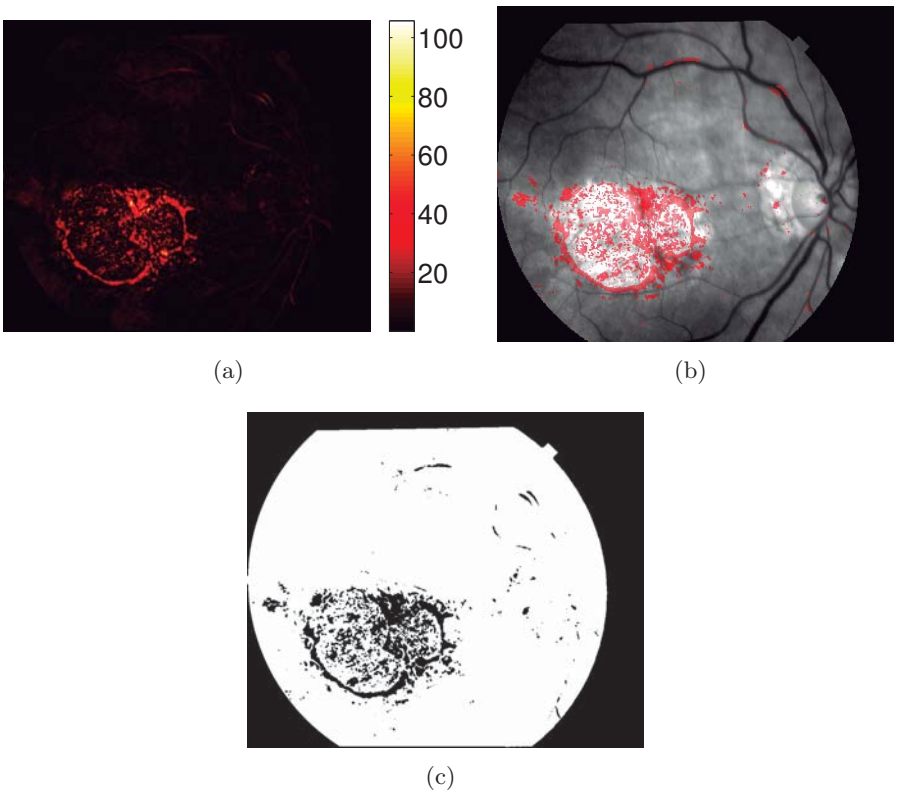
nor structural changes, then $k \approx 1$ throughout the common object field. In Fig. 5.5 we show the resulting $k(x, y)$ for the images in Fig. 5.3. The different shades of gray indicate the average contrast and intensity difference between the two images. From the image it can be seen that most areas have similar intensity values except for the upper left part (dark region).

5.4.3 Segmentation of Areas with Structural Changes

Up to this point we have addressed two conditions required to make the images comply with the convolutional model: the registration or spatial alignment of the images, and the compensation for illumination variation. As we discussed previously, because we consider images acquired between patient visits, we must also take into account the presence of possible longitudinal changes in the given retina over time due to treatments and/or disease progression (Fig. 5.1). The detection of such changes is no small feat. Traditionally, the clinical procedures for inspecting the images have been mostly manual. Until recently (Narasimha-Iyer et al., 2006) there was no fully automated and robust way to carry it out. The work of Narasimha-Iyer et al. (2006) has prompted further research (Narasimha-Iyer et al., 2007, Xiao et al., 2012) as we have done here (Marrugo et al., 2011b;a) by adapting the general methodology to suit our purposes for restoring retinal images.

Longitudinal Change Detection

The longitudinal changes in the retina are typically structural changes, and as such they need to be segmented and masked out. Image change analysis is of interest in various fields and many algorithms have been developed for such a task (Aach & Kaup, 1995, Chang et al., 2005). A complete overview of change detection methods can be found in the survey by Radke et al. (2005). An initial step in order to identify these changes comes from computing the difference from the two registered images including the illumination com-



(a) (b)

(c)

Figure 5.6: Intermediate outputs from the algorithm: (a) Image difference $\Delta z(x, y)$ in absolute value, (b) Image difference in pseudocolor on top of gray-scale fundus image, and (c) mask m for avoiding areas with structural changes in subsequent PSF estimation.

pensation as

$$\Delta z(x, y) = z_1(x, y) - k(x, y) \cdot z_2(x, y) \ . \tag{5.4}$$

The difference image $\Delta z(x, y)$ from the two input images (Fig. 5.3) is shown in absolute value in Fig. 5.6(a). To better understand this result, in Fig. 5.6(b) we show one of the retinal images in gray-scale where the pixels related to structural changes are highlighted in pseudo-color. This image constitutes an important output of our algorithm. The structural changes can now be visualized and detected from the difference image $\Delta z(x, y)$ by taking a statistical significance test, in the same fashion as described by Narasimha-Iyer et al. (2006). First, structural changes are often associated with a group of pixels, thus the change decision at a given pixel j should be based on a small block of pixels in the neighborhood of j denoted as w_j. Second, in the absence of any change, the difference can be assumed to be due to noise alone. Therefore, the decision as to whether or not a change has occurred corresponds to choosing one of two competing hypothesis: the *null*

hypothesis $\mathcal{H}_0$ or the *alternative hypothesis* $\mathcal{H}_1$, corresponding to *no-change* and *change* decisions, respectively. Assuming a Gaussian distribution for the difference values, the changes can be identified by comparing the normalized sum square of the differences within the neighborhood w_j to a predetermined threshold τ as described by Aach & Kaup (1995). The test is carried out as

$$\Omega_j = \frac{1}{\sigma_n^2} \sum_{(x,y)\in w_j} \Delta z(x,y)^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau \ , \tag{5.5}$$

where σ_n is the noise standard deviation of the difference in the no-change regions. The threshold τ is derived from the fact that Ω_j follows a χ^2 distribution with N degrees of freedom, where N is the number of pixels in the window w_j. It can be obtained for a particular false positive rate α from the χ^2 tables. The choice of an appropriate α is both guided by mathematical considerations, e.g. the conventional 5% level for statistical significance (Stigler, 2008), and the consequences that false alarms and misses might have. In this case, the effect of false alarms is unimportant because there would still be a large number of remaining pixels from where to compute the PSFs. On the other hand, misses do have a considerable impact in view of the fact that these pixels do not fulfill the convolutional model. As a result, α values below 0.05 might yield a more accurate change detection at the expense of possible undesirable misses. For all experiments we use a 3×3 window ($N = 9$) and set $\alpha = 0.05$. The parameter σ_n was estimated by manually picking out no-change regions from a training set of images, computing Eq. (5.4) and the standard deviation inside these regions. Using Eq. (5.5) at each pixel, we can determine a change mask between the images or conversely a no-change mask. Given that for the MBD procedure we are interested in estimating the PSF from the no-change regions the masking function m is obtained directly from the no-change mask of the significance test. The mask for the input images (Fig. 5.3) is shown in Fig 5.6(c). Notice that the pathological region is the main cause of structural changes.

5.4.4 PSF Estimation

In this subsection, we describe the basic principles of the blind deconvolution method used for the estimation of the PSFs. For this purpose we have chosen one of the best working MBD methods (Sroubek & Flusser, 2005). The algorithm can be viewed as a Bayesian MAP estimation of the most probable sharp image and blur kernels. For our purposes, we modified the original method so that it ignores regions affected by structural changes, which improves stability and precision of the computation. Without this modification, represented by the mask m in Eq. (5.6), the algorithm does not work reliably. The algorithm can be described as a minimization of the

functional

$$\arg\min_{u,h_1,h_2} \left(\frac{1}{2}||u * h_1 - z_1||^2 + \frac{1}{2}||m(u * h_2 - kz_2)||^2 + \lambda_u \int |\nabla u| \, \mathrm{d}x \, \mathrm{d}y \right. \\ \left. + \lambda_h ||m(z_1 * h_2 - kz_2 * h_1)||^2 \right) \ , \quad h_1, h_2 \geq 0 \ , \tag{5.6}$$

with respect to the latent image u and blur kernels h_1 and h_2. The first and second terms measure the difference between the input blurred images and the searched image u blurred by kernels h_1 and h_2. The size of this difference is measured by L_2 norm $||.||$ and should be small for the correct solution; ideally, it should correspond to the noise variance in the given image. Function k compensates for uneven illumination as described in Section 5.4.2. The value of the masking function m is 1 in the valid points (white in Fig. 5.6(d)) and 0 in the pixels where the eye fundus has structurally changed. Any of the first two terms could be masked, but not both at the same time. This is because the latent image u cannot have pixels with no value at all, hence these pixels must take values from any of the two images. In this case z_2 is masked, as a result these pixels take values from the first term. The two remaining terms are regularization terms with positive weighting constants λ_u and λ_h. The third term is nothing else than the total variation of image u. It improves stability of the minimization and from the statistical viewpoint incorporates prior knowledge about the solution. The last term is a condition linking the PSFs h_1 and h_2 of both images, which also improves the numerical stability of the minimization.

The functional is alternately minimized in the subspaces corresponding to the image and the PSFs. The advantage of this scheme lies in its simplicity, this alternating minimization approach is actually a variation of the steepest-descent algorithm. The minimization in the PSF subspace is equivalent to the solution of a system of linear equations in the least square sense with the non-negativity constraint, in our implementation solved by the MATLAB *fmincon* function (MATLAB, 2010). The non-blind deconvolution realized by the minimization in the image subspace, is solved by half-quadratic iterative scheme (Chambolle & Lions, 1997), replacing the total variation by $\int \sqrt{|\nabla u|^2 + \epsilon^2}$, where ϵ is an auxiliary variable in the range $0 < \epsilon << 1$. It is a small relaxation parameter which makes total variation differentiable around zero. A typical value for ϵ is 10^{-1}.

The main difference with respect to the original method (Sroubek & Flusser, 2005) is the introduction of the masking function m, which is computed in the beginning of the algorithm as described in Section 5.4.3. During the minimization, the multiplication by m is included in the operator corresponding to the convolution with u (in the PSF minimization step) and in the operator corresponding to the convolution with h_2 (in the image minimization step). Because of the simplicity of this masking operation, the

speed is practically the same as the speed of the original algorithm. In addition, even though we work with a complicated set of pixels, we can use the standard operation of convolution, which can eventually be speeded up using the Fast Fourier transform (FFT).

5.4.5 Image Restoration

Once the PSFs have been properly estimated the restoration step is basically a non-blind deconvolution operation. Nevertheless, note that from Eq. (5.6) the restored version of z_1 ($\hat{u}_1$) is obtained because z_2 is masked. Conversely, $\hat{u}_2$ could be obtained by minimizing Eq. (5.6) again with fixed PSFs and masking z_1. However, this procedure has the disadvantage that both images are restored only within the common object field. Therefore, the proper solution is to restore each image z_i via single-channel (non-blind) deconvolution with their corresponding PSF h_i (estimated from the previous step) by the minimization of the functional

$$\arg\min_{u_i} \left(||u_i * h_i - z_i||^2 + \lambda_u \int |\nabla u_i| \, \mathrm{d}x \, \mathrm{d}y \right) \quad . \tag{5.7}$$

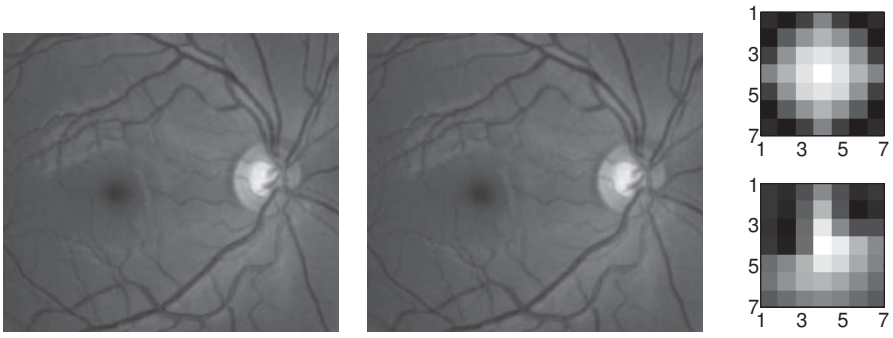
Assuming that the PSFs are spatially invariant in the object field, this approach has a further advantage in that the PSF estimation can be computed from a relatively small area of the common object field, provided that there are retinal structures within, thus greatly reducing the computational cost of the combined PSF estimation plus image deconvolution.

Finally, it should also be noted that the whole process of PSF estimation plus deconvolution can be computed for every channel of the RGB color fundus image. However, in spite of the increase in computational burden, tests showed no real advantage to estimate the PSF for each channel. Moreover, we found the green channel to be the most suitable channel for PSF estimation because it provides the best contrast. Whereas the blue channel encompasses the wavelengths most scattered and absorbed by the optical media of the eye, hence the image has very low energy and relatively high level of noise. As a result, the RGB deconvolved fundus image was computed by deconvolving every R, G, and B channel from the green channel PSF.

5.5 Experiments and Results

5.5.1 Synthetic Images

In this subsection, we use synthetically degraded retinal images to test the performance of the proposed method. We use blurred signal-to-noise ratio (BSNR) to measure the noise contained in the degraded image, and improvement in signal-to-noise ratio (ISNR) to measure the quality of restored



(a) (b) (c)

Figure 5.7: (a)-(b) Degraded images (BSNR = 40 dB). (c)-(d) PSFs.

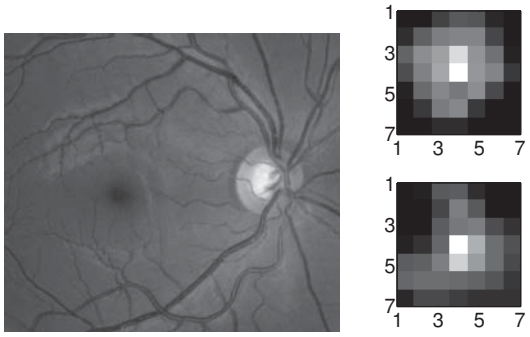
images. They are defined as follows:

$$\begin{aligned} \text{BSNR} &= 20\log_{10}\left(\frac{\|z\|}{\|n\|}\right) \\ \text{ISNR} &= 20\log_{10}\left(\frac{\|u-z\|}{\|u-\hat{u}\|}\right) \end{aligned}$$

where u, z, $\hat{u}$, and n are the original image, degraded image, restored image, and the noise vector, respectively. For ISNR higher means better restoration, whereas for BSNR lower means noisier degraded image. These metrics are mainly used to provide an objective standard for comparison with other techniques and they can only be used for simulated cases.

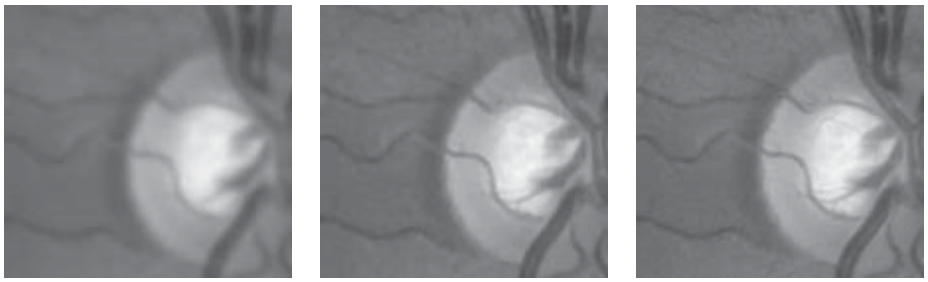
The first example is shown in Fig. 5.7, where the degraded images are synthesized from a sharp real image and the kernels shown in Fig. 5.7(c) plus Gaussian noise with zero mean and variance $\sigma^2 = 10^{-6}$ (BSNR = 40dB). The recovered image and PSFs are shown in Fig. 5.8. The restoration provides an ISNR = 4.45dB. In this case for synthetically degraded images the masking operation of Section 5.4.3 was not applied. Visual inspection of the details shown in Fig. 5.9 clearly reveal the accuracy of the method. Under these circumstances the algorithm is able to produce a significant restoration of fine details like small blood vessels around the optic disc.

To further test our approach under a more realistic degradation we produced an initial geometrical distortion via a quadratic model (Can et al., 2002, Lee et al., 2010) as the one used for registration (Fig. 5.10). After the geometric distortion the degradation (blur plus noise) is produced on both images (BSNR = 40dB). They are then registered and the restored image is recovered via MBD. The restored image and the estimated PSFs are shown in Fig. 5.11. The ISNR is slightly less (4.11dB) than in the previous case, but still sufficient to produce a significant restoration. To corroborate our assumption that MBD methods seem better suited for this type of images,



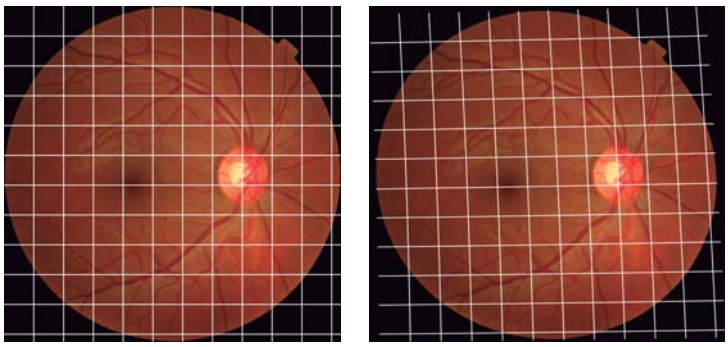
(a) (b)

Figure 5.8: (a) Restored image (ISNR = 4.45dB). (b) Estimated PSFs.



(a) (b) (c)

Figure 5.9: Details from (a) Degraded image, (b) Restored image, and (c) Original image.



(a) (b)

Figure 5.10: (a) Original image and (b) geometrically distorted image.

we tried to restore the image with a recent SBD method proposed by Xu & Jia (2010). The result is shown in Fig. 5.11(c) and visually reveals that it does not follow the true nature of the blurring with artifacts around the blood vessels, thus being prone to produce a poor restoration evidenced by an ISNR $= -0.72$dB.

Concerning parameter setting, in Fig. 5.12 we show the sensitivity of the two parameters λ_u and λ_h for the minimization of Eq. (5.6) in ISNR of the restored images. In Fig. 5.12(a) we fix the value of λ_h to 10 and check the ISNR of the restored images for different initial values of $\lambda_u = \left\{10^0, 10^1, 10^2, 10^3, 10^4, 10^5\right\}$. The best restoration is obtained with $\lambda_u = 10^3$, thus in Fig. 5.12(b) we carried out the same procedure by fixing the value of λ_u to 10^3 and checking the ISNR of the restored images for different values of $\lambda_h = \{1, 10, 20, 30, 40, 50\}$. The best restoration was obtained with an initial value of $\lambda_h = 30$. For this type of images, when scaled to the interval $\langle 0, 1\rangle$, we find $20 < \lambda_h < 40$ to be a suitable range to produce an optimal restoration.

5.5.2 Real Images

The experiments shown in this subsection aim to demonstrate the applicability of the proposed method for retinal image deblurring in real scenarios. Three different cases are shown in Fig. 5.13 including the retinal images that were used to illustrate the method (Fig. 5.3). The estimated PSFs are shown at the bottom of the restored images. All images contain some pathological damage and have been acquired within considerable lapses of time (several months). In all examples the resolution improvement can be visually assessed by the clear distinction of details such as small blood vessels or the increase in sharpness of edges specially in the pathological areas. We emphasize the fact that these images correspond to real routine patient follow-up and were not intentionally degraded. From a clinical viewpoint,

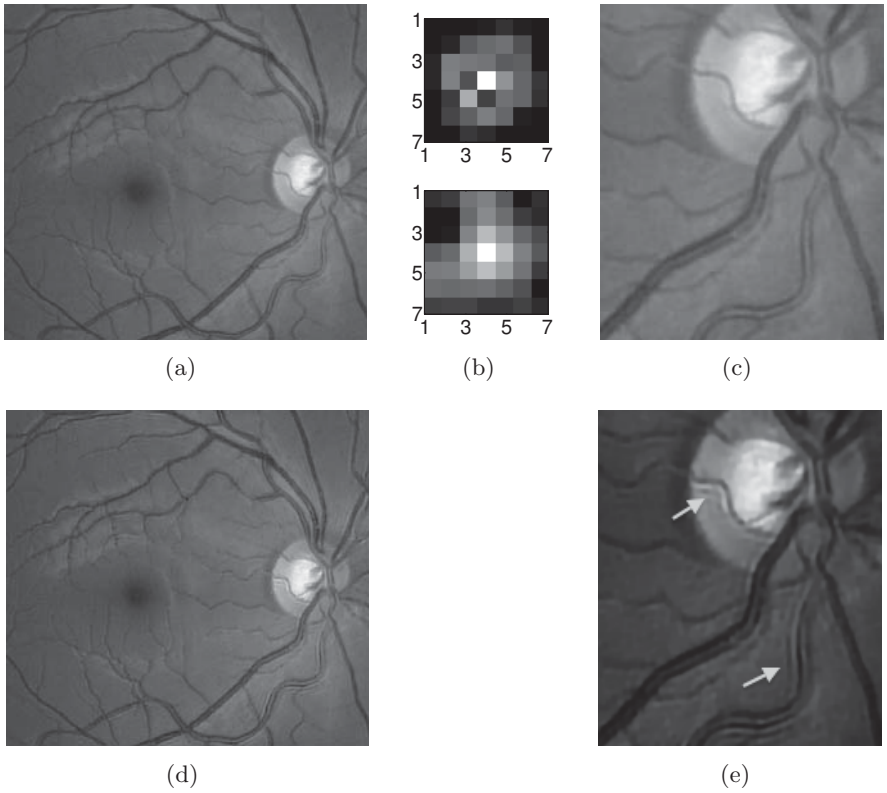


Figure 5.11: Image restoration from degraded and geometrically distorted images. (a) Restored image by the proposed method (ISNR = 4.11dB), (b) estimated PSFs and (c) image detail. (c) Restored image with the method of Xu & Jia (2010) (ISNR = −0.72dB) and (d) image detail.

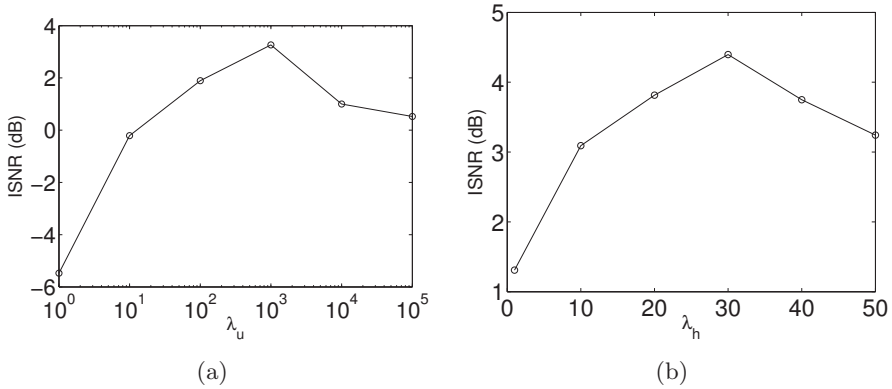


Figure 5.12: Test on parameter setting (BSNR = 40dB). Average ISNR with respect to different initial values of (a) λ_u and (b) λ_h.

the enhancement can be used for a more precise assessment of a patient's state. Likewise, the images are more suitable for subsequent processing such as for the detection of retinal pathology (Muramatsu et al., 2010, Xu & Luo, 2010).

In Fig. 5.14 the same images are shown but in grey-scale to highlight the areas of structural change in pseudocolor. As we have mentioned earlier, this is an important result for its potential impact in the medical practice. Subtle changes can be identified by this approach such as the ones in Fig. 5.14(b) and the hemorrhage in the region of the optic disc in Fig. 5.14(c). Another technique to rapidly identify changes from the two images is by alternating both restored images in a video sequence.* Videos 1 and 2 (Fig. 5.13) correspond to the first two real cases.

*The videos are available in `http://www.goapi.upc.edu/usr/andre/retvideo.html`

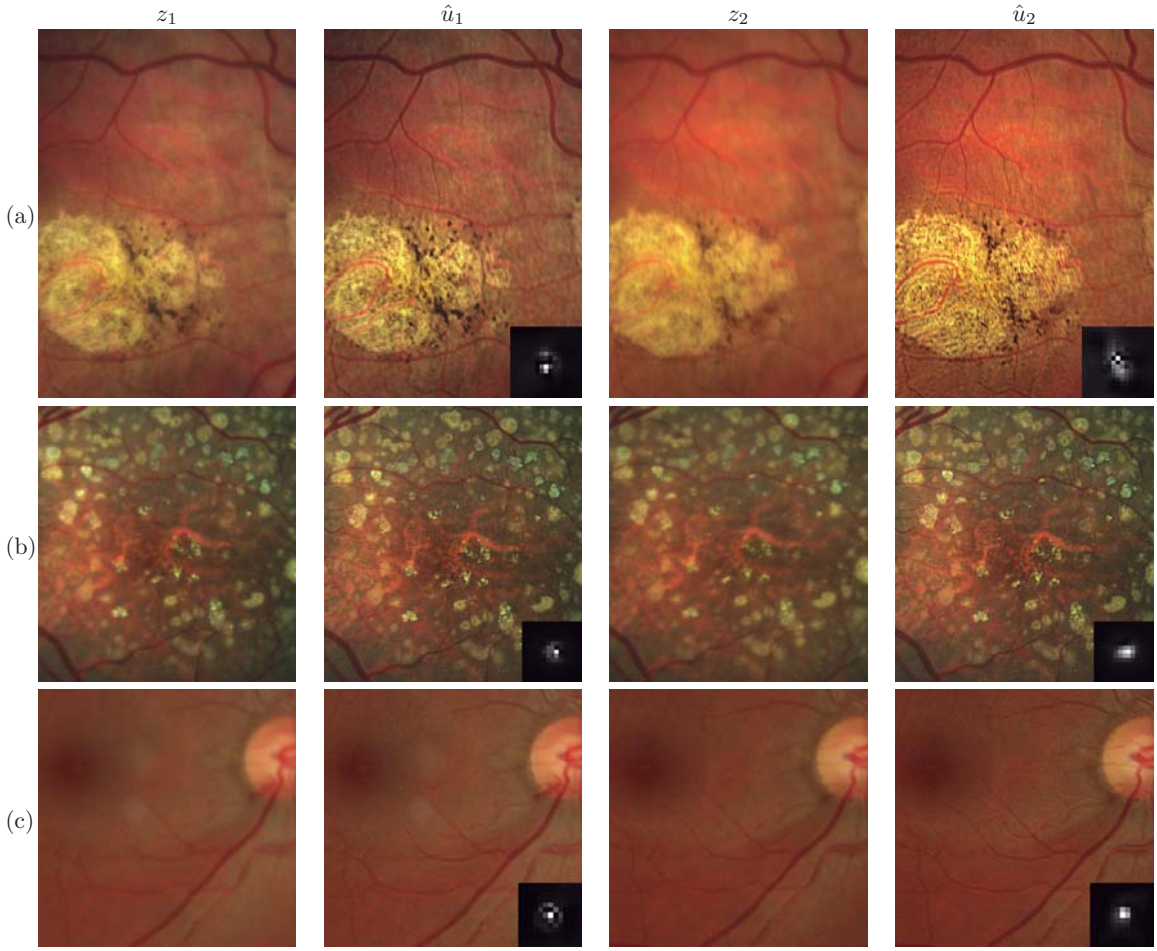
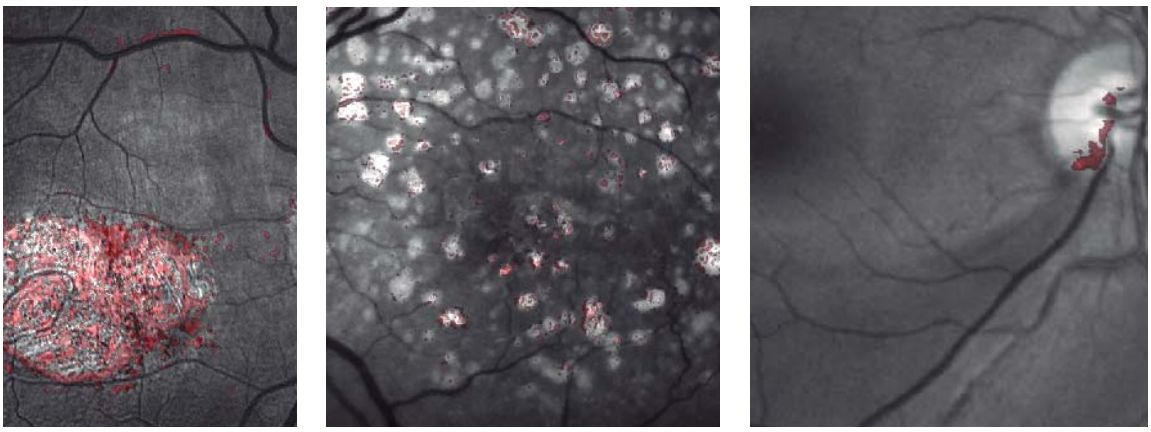


Figure 5.13: Original and restored color retinal images; (a), (b) and (c) indicate three separate cases arranged from left to right following our notation for degraded (z_i) and restored ($\hat{u}_i$) images. The images are cropped to represent the region of interest given by the pathological area. The estimated PSF is shown at the bottom of the restored image. Video files are also included for change detection in cases (a) and (b) `http://www.goapi.upc.edu/usr/andre/retvideo.html` (Video 1, Quicktime, 0.5MB; Video 2, Quicktime, 0.4MB).



(a) (b) (c)

Figure 5.14: (*This figure is meant to be viewed in color*) Visualization of structural changes in pseudo-color for the images of Fig. 5.13.

5.6 Discussion

In this chapter we have investigated a new approach for retinal image restoration based on multichannel blind deconvolution. In addition, we have developed a strategy for identifying and highlighting areas of structural change with possible relation to pathological damage. We have verified that fundus images of the same retina over time contain enough common information to be restored with the proposed method. The method consists of a series of preprocessing steps to adjust the images so they comply with the convolutional model, followed by the final stages of PSF estimation and deconvolution. At this level of approach, spatially invariant PSFs are assumed. The synthetically degraded images have enabled us to test the performance of the proposed approach and also to compare with a state-of-the-art single-channel blind deconvolution method. The results have showed a remarkable enhancement evidenced by the increased visibility of details such as small blood vessels or pathological areas.

Chapter 6

Deblurring Retinal Images with Space-Variant Blur

6.1 Introduction

In the previous chapter we introduced the problem of retinal image blurring and proposed a deblurring strategy (see § 5.4) to overcome such degradation. However, that strategy is limited to images blurred uniformly; in other words, we assumed the blur to be space-invariant. The space-invariant assumption is commonplace in most of the restoration methods reported in the literature, (Levin et al., 2011) but in reality it is a known fact that blur changes throughout the image (Bedggood et al., 2008). In Fig. 6.3(a) we show an attempt at restoring an image degraded with spatially variant blur with the space-invariant approach. Note how the restoration fails. In this chapter we consider the blur to be both unknown and space-variant (SV). This in itself is a novel approach in retinal imaging; relevant to such extent that many common eye related conditions, such as astigmatism, keratoconus, corneal refractive surgery, or even tear break-up, may contribute significantly to a decline in image quality typically in the form of a SV degradation (Tutt et al., 2000, Xu et al., 2011). An example of such a condition is shown in Fig. 6.11(a). The image corresponds to an eye from a patient with corneal abnormalities that lead to a loss in visual acuity and a quality degradation of the retinal image (Fig. 6.11(b)).

Restoration of images with SV blur from optical aberrations has been reported in the literature (Costello & Mikhael, 2003), although the main limitation is that the blurred image is often restored in regions or small sections, which are then stitched together. This inevitably leads to *ringing* artifacts associated with frequency-domain filtering like in Wiener filtering. Another clear disadvantage is a significant complexity for accurately estimating the SV PSF, for instance Bardsley et al. (2006) use a phase-diversity based scheme to obtain the PSF associated with an image patch. This type of ap-

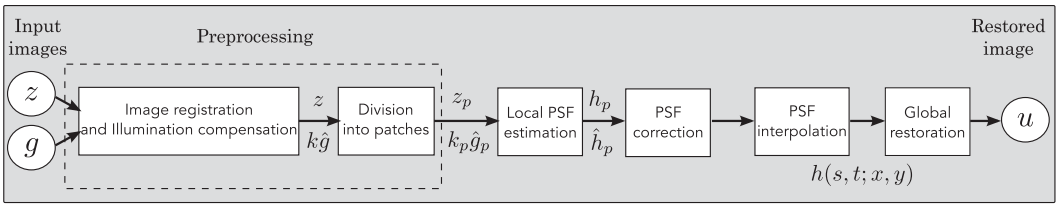


Figure 6.1: Block diagram illustrating the proposed method. z is the degraded image, g is an auxiliary image of the same eye fundus used for the PSF estimation, and u is the restored image. The other variables are intermediate outputs of every stage; their meaning is given in the text.

proach is common in atmospheric optics where the conditions and set-up of the imaging apparatus (typically a telescope) are well known and calibrated. Unfortunately, this is not immediately applicable to retinal imaging, at least non-adaptive optics retinal imaging.

If we assume that the SV PSF is locally space-invariant, its estimation can be carried out on a patch basis. However, the estimation may fail in patches with no structural information. Tallón et al. (2012) developed a strategy for detecting these *non-valid* patches in a SV deconvolution and denoising algorithm from a pair of images acquired with different exposures: a sharp noisy image with a short exposure and a blurry image with a long exposure. Because they had two distinct input images they were able to: (i) Identify patches where the blur estimates were poor based on a comparison (via a thresholding operation) of the deconvolved patches with the sharp noisy patches. (ii) In those patches, instead of correcting the local PSFs and deconvolving the patches again, they performed denoising in the noisy sharp image patch. The end result is a patchwork-image of deconvolved patches stitched together with denoised patches. Their method is mainly oriented at motion blur, this is the reason for a dual exposure strategy. This is not readily implementable in the retinal imaging scenario where the SV blur is generally caused by factors like aberrations.

Finally, there have been several works (Harmeling et al., 2010, Whyte et al., 2010, Gupta et al., 2010) that try to solve the SV blind deconvolution problem from a single image. The common ground in these works is that the authors assume that the blur is only due to camera motion. They do this in order to reduce the space in which to search for SV blurs. Despite their approach being more general, the strong assumption of camera motion, as in (Tallón et al., 2012), is simply too restrictive to be applied in the retinal imaging scenario.

6.1.1 Contribution

In this chapter we propose a method for removing blur from retinal images. We consider images degraded with SV blur, which may be due to factors

like aberrations in the eye or relative camera-eye motion. Because restoring a single blurred image is an ill-posed problem, we make use of two blurred retinal images from the same eye fundus to accurately estimate the SV PSF. Before the PSF estimation and restoration stages take place, we preprocess the images to accurately register them and compensate for illumination variations not caused by blur, but by the lighting system of the fundus camera. The principles of these stages are shared with the deblurring method already described in § 5.4 of the previous chapter. This is depicted in the block diagram shown in Fig. 6.1. The individual stages of the method are explained in § 6.3.

We assume that in small regions of the image the SV blur can be approximated by a spatially invariant PSF. In other words, that in a small region the wavefront aberrations remain relatively constant; the so-called *isoplanatic* patch (Bedggood et al., 2008, Goodman, 1968). An important aspect of our approach is that instead of deblurring each patch with its corresponding space-invariant PSF—and later stitching together the results—we sew the individual PSFs by interpolation, and restore the image globally. The estimation of local space-invariant PSFs, however, may fail in patches with no structural information. Unlike other methods, we incorporate prior knowledge of the blur that originates through the optics of the eye to address this limitation. To this end we propose a strategy based on eye-domain knowledge for identifying non-valid local PSFs and replacing them with appropriate ones. Even though methods for processing retinal images in a space-dependent way (like locally adaptive filtering techniques (Salem & Nandi, 2007, Marrugo & Millán, 2011)) have been proposed in the literature, to the best of our knowledge this is the first time a method for SV deblurring of retinal images is proposed.

6.2 Space-Variant Model of Blur

In the previous chapter we modeled the blurring of a retinal image by convolution with a unique global PSF, assuming space-invariant blurring by a model

$$z = u * h + n \tag{6.1}$$

where z and u are the blurred and sharp (original) images, respectively, h is a convolution kernel and n white Gaussian noise $N(0, \sigma^2)$. This approximation is valid as long as the PSF changes little throughout the field of view (FOV). In other words, that the blurring is homogenous. In reality we know that the PSF is indeed spatially variant (Bedggood et al., 2008), to such extent that in some cases the space-invariant approach completely fails, bringing forth the need for a SV approach. To address this limitation we now model

the blurred retinal image z by a general linear operator

$$z = Hu + n \ , \tag{6.2}$$

The operator H can be written in the following form

$$z(x,y) = [Hu](x,y) = \int u(s,t)h(s,t,x-s,y-t)\ dsdt \ , \tag{6.3}$$

where h the SV PSF. The operator H is a generalization of standard convolution where h is now a function of four variables. We can think of this operation as a convolution with a PSF $h(s,t,x,y)$ that is now dependent on the position (x,y) in the image. Standard convolution is a special case of Eq. (6.3), where $h(s,t,x,y) = h(s,t)$ for an arbitrary position (x,y). Note that the PSF h is a general construct that can represent other complex image degradations which depend on spatial coordinates, such as motion blur, optical aberrations, lens distortions and out-of-focus blur.

6.2.1 Representation of space-variant PSF

An obvious problem of spatially varying blur is that the PSF is now a function of four variables. Except trivial cases, it is hard to express it by an explicit formula. Even if the PSF is known, we must solve the problem of a computationally-efficient representation.

In practice we work with a discrete representation, where the same notation can be used but with the following differences: the PSF h is defined on a discrete set of coordinates, the integral sign in Eq. (6.3) becomes a sum, operator H corresponds to a sparse matrix and u to a vector obtained by stacking the columns of the image into one long vector. For example in the case of standard convolution, H is a block-Toeplitz matrix with Toeplitz blocks and each column of H corresponds to the same kernel $h(s,t)$ (Golub & Van Loan, 1996). In the SV case that we address here, as each column of H corresponds to a different position (x,y), it may contain a different kernel $h(s,t,x,y)$.

In retinal imaging, all typical causes of blur change in a continuous gradual way, which is why we assume the blur to be locally constant. Therefore, we can make the approximation that locally the PSFs are space-invariant. By taking advantage of this property we do not have to estimate local PSFs for every pixel. Instead, we divide the image into rectangular windows and estimate only a small set of local PSFs (see Fig. 6.5) following the method described in (Marrugo et al., 2011a) and outlined in Section 6.3. The estimated PSFs are assigned to the centers of the windows from where they were computed. In the rest of the image, the PSF h is approximated by bilinear interpolation from the four adjacent local PSFs. This procedure is explained in further detail in the following section.

6.3 Description of the Method

In this section we describe the different stages of the proposed restoration method depicted in Fig. 6.1. This work follows from Chapter 5 and addresses a more general problem: restoration of retinal images in the presence of a SV PSF. Blind deconvolution of blur from a single image is a difficult problem, but it can be made easier if multiple images are available. In § 5.5.1 we showed that single image blind deconvolution for blurred retinal images does not provide a suitable restoration. Moreover, in images with SV blur the restoration is even worse. Alternatively, by taking two images of the same retina we can use a multi-channel blind deconvolution strategy that is mathematically better-posed (Sroubek & Flusser, 2005). In fact, in order to properly estimate the SV PSF we use the degraded original image z (Fig. 6.2(a)) and a second auxiliary image g of the same retina, shown in Fig. 6.3(b). Unlike the previous chapter, the images used here were acquired in the same session. It is important to note that because the eye is a dynamical system the two images are not blurred in exactly the same way, which is a requirement in the multi-channel approach. The second image is used only for the purpose of PSF estimation. We make clear that the method is proposed so that there is no preference over which of the two images is restored; meaning that in ideal conditions restoring one or the other would produce the similar results. A practical approach, nevertheless, would be to restore the image that is less degraded, thereby obtaining the best restoration possible. Once the PSF is properly estimated, we restore the image as described in the last part of this section.

To illustrate the method we used the image shown in Fig. 6.2(a) as the degraded original image to be restored. This retinal image was acquired from a person with strong astigmatism in the eye, which introduces a SV blur. The strong aberrations hinder the imaging procedure making it impossible to achieve a perfectly sharp image throughout the whole FOV. Notice how the upper regions are much more blurred than anywhere else in the image.

6.3.1 Preprocessing

Because we use a multi-channel scheme for the estimation of the local PSFs, the images are preprocessed so that they meet the requirements imposed by the space-invariant convolutional model given by Eq. (6.5). In the same way as in § 5.4 this consists in registering the images and adjusting their illumination distribution. By accounting for these differences the remaining radiometric differences between the images are assumed to be caused by blur and noise.

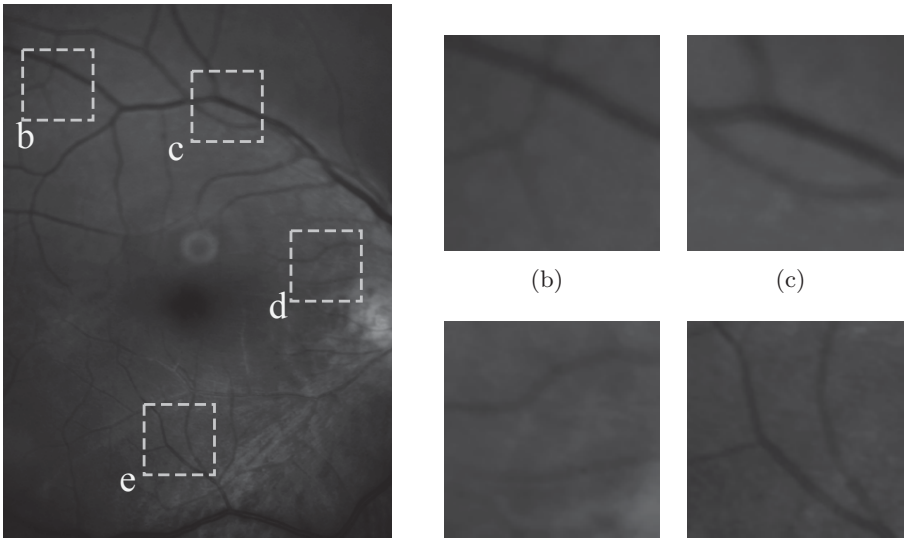


Figure 6.2: (a) Retinal image degraded with space-variant blur given by strong astigmatism. (b), (c), (d), and (e) zoomed regions to show the space-variant nature of the blur.

Image registration

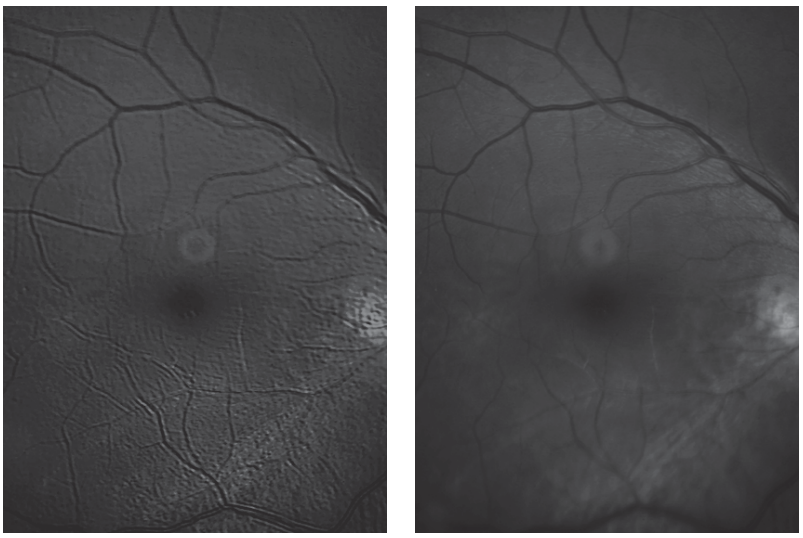
For registering the images we have used the robust dual-bootstrap iterative closest point algorithm (Stewart et al., 2003), which we already described in § 5.4.1. The transformation model includes rotation, scale, translation, a shearing term and a quadratic term that accounts for the spherical shape of the retina (Can et al., 2002). With this procedure we transform the auxiliary image g (Fig. 6.3(b) in our example) to the coordinates of z. The transformed image is denoted by $\hat{g}$.

Illumination compensation

Nonuniform illumination is particularly a problem in retinal imaging due to a combination of several factors. These include: the difficulty of achieving uniform illumination even with a fully dilated pupil, instrument limitations such as the ring-shaped illumination pattern, and curved surface of the retina. Because we use two images to estimate the blurring kernels we must compensate for this spatially nonuniform illumination.

In the same fashion as in § 5.4.2 we compensate the illumination differences between the two images z and $\hat{g}$ by properly adjusting the intensity values of one image to approximately match those of the other, while satisfying a predetermined illumination model (Marrugo et al., 2011a).

We use a low-order parametric surface fitting equation to model the



(a) (b)

Figure 6.3: (a) The space-invariant restoration of Fig. 6.2(a) (Marrugo et al. (2011a)). Notice the artifacts. (b) Second (auxiliary) image g for the PSF estimation.

illumination $k(x, y)$ of the fundus as a slowly varying light field over a smooth surface as (Narasimha-Iyer et al., 2006)

$$\arg\min_{k} \|z - k \cdot \hat{g}\|^2 \ , \tag{6.4}$$

where $k(x, y) = \alpha_{15}y^4 + \alpha_{14}y^3x + \cdots + \alpha_2 y + \alpha_1$, and $\hat{g}$, z are the retinal images. We minimize Eq. (6.4) in the least squares sense to estimate the 15 parameters. The interpretation of k is straightforward. If the images $\hat{g}$ and z had neither illumination changes nor structural changes, then $k \approx 1$ throughout the common object field. In Fig. 6.4 we show the resulting $k(x, y)$ for the pair of images shown in Fig. 6.2(a) and Fig. 6.3(b). The different shades of gray indicate the average contrast and intensity difference between the two images. This is applied to image $\hat{g}$ to match the illumination distribution of the background of image z (Fig. 6.2(a)).

6.3.2 Estimation of local PSFs

In section 6.2 we described the model for a spatially varying blur in which we assume the PSF h to vary gradually, which means that within small regions the blur can be locally approximated by convolution with a space-invariant PSF. For this reason we approximate the global function h from Eq. (6.3) by interpolating local PSFs estimated on a set of discrete positions. The main

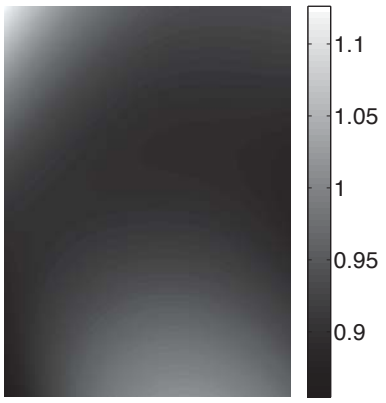


Figure 6.4: Illumination compensation function $k(x, y)$.

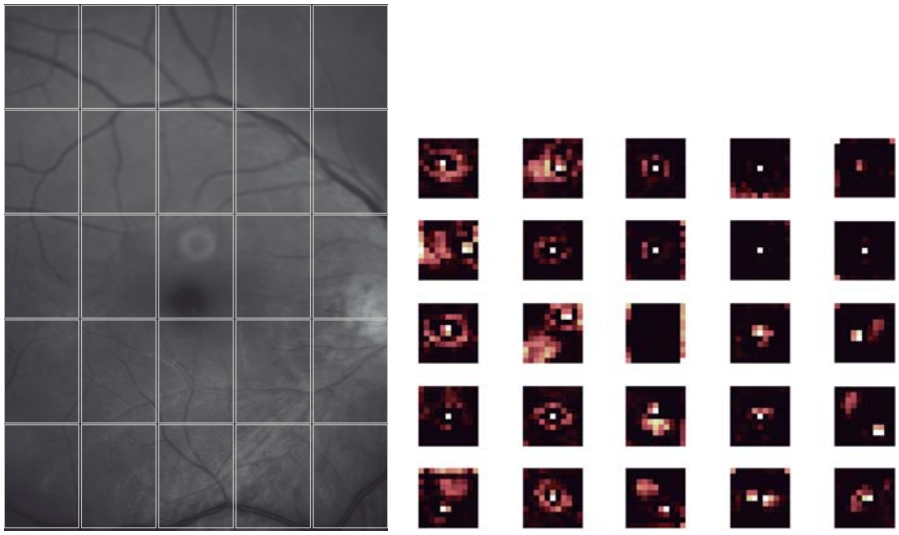
advantage of this approach is that the global PSF needs not be computed on a per-pixel basis which is inherently time-consuming.

The procedure for estimating the local PSFs is the following. We divide the images z and $\hat{g}$ with a grid of $m \times m$ patches. In each patch p we assume a convolutional blurring model, like in § 5.2, where an ideal sharp patch u_p originates from two degraded patches z_p and $\hat{g}_p$ (for $p = 1 \dots m \times m$). The local blurring model is

$$\begin{aligned} z_p &= h_p * u_p + n \\ \hat{g}_p &= \hat{h}_p * (k_p^{-1} \cdot u_p) + \hat{n} \ , \end{aligned} \tag{6.5}$$

where $*$ is the standard convolution, h_p and $\hat{h}_p$ are the convolution kernels or local PSFs, and k_p is the function accounting for relative local illumination change between patches z_p and $\hat{g}_p$. The noise (n and $\hat{n}$) is assumed to be Gaussian additive with zero mean.

From this model we can estimate the local PSFs with an alternating minimization procedure as described in § 5.4.4, but applied locally. We do so on a grid of 5–by–5 image patches to compute 25 PSFs, as shown in Fig. 6.5. The grid size was chosen so that the image patches were big enough to include retinal structures. This makes it easier to successfully recover a PSF. In our case, with retinal images of size approximately of 1280×1280 pixels (the region of interest is slightly smaller) local PSFs h_p of size 11×11 pixels proved sufficient, with image patches of approximately $\frac{1}{5}$ of the size of region of interest. Every local PSF is computed on each patch



(a) (b)

Figure 6.5: (a) Fig. 6.2(a) with a 5 by 5 grid that defines the image patches. (b) Estimated local PSFs.

p by minimizing the functional

$$\mathbf{E}(u_p, h_p, \hat{h}_p) = \arg \min_{u_p, h_p, \hat{h}_p} \left(\frac{1}{2}\|u_p * h_p - z_p\|^2 + \frac{1}{2}\|u_p * \hat{h}_p - k_p \hat{g}_p)\|^2 + \right.$$
$$\left. + \lambda_1 \int |\nabla u_p| \, \mathrm{d}x \, \mathrm{d}y + \lambda_2 \|z_p * \hat{h}_p - k_p \hat{g}_p * h_p)\|^2 \right),$$
$$h_p, \hat{h}_p(s,t) \geq 0, \qquad (6.6)$$

with respect to the ideal sharp patch u_p and the blur kernels h_p and $\hat{h}_p$. The blur kernel $h_p(s,t)$ is an estimate of $h(s,t,x_0,y_0)$, where (x_0,y_0) is the center of the current window z_p, and $\|.\|$ is the L_2 norm. The first and second terms of Eq. (6.6) measure the difference between the input blurred patches (z_p and $\hat{g}_p$) and the sharp patch u_p blurred by kernels h_p and $\hat{h}_p$. This difference should be small for the correct solution. Ideally, it should correspond to the noise variance in the image. The two remaining terms of Eq. (6.6) are regularization terms with positive weighting constants λ_1 and λ_2, which we have set following the fine-tuning procedure described in (Marrugo et al., 2011a). The third term is the total variation of u_p. It improves stability of the minimization and from a statistical perspective it incorporates prior knowledge about the solution. The last term is a condition linking the convolution kernels which also improves the numerical stability of the minimization. The functional is alternately minimized in the subspaces corresponding to the image and the PSFs. Note that Eq. (6.6) is the local version of the expression derived in Eq. (5.6) of previous chapter for the

computation of the space-invariant PSF.

Although u_p is a restored patch, however it is discarded. This is because our method does not work by performing local deconvolutions and sewing restored patches together, which in practice would produce artifacts on the seams. Instead we perform a global restoration method explained in § 6.3.5.

6.3.3 Identifying and correcting non-valid PSFs

Strategy based on eye domain knowledge

The local PSF estimation procedure not always succeeds. For instance, if we use the estimated local PSFs shown in Fig. 6.5 for deconvolution, the reconstruction develops ringing artifacts (see Fig. 6.9(b)). Consequently, such non-valid PSFs must be identified, removed and replaced. In our case we replace them by an average of adjacent valid kernels. The main reason why the kernel estimation fails is due to the existence of textureless or nearly homogenous regions bereft of structures with edges (e.g. blood vessels) to provide sufficient information (Tallón et al., 2012). To identify them we devised an eye-domain knowledge strategy. The incorporation of proper *a priori* assumptions and domain knowledge about the blur into the method provides an effective mechanism for a successful identification of non-valid PSFs. First, because the patient's eye is part of the imaging system, it is logical to consider that the PSF's shape is partly determined by the eye's PSF. Moreover, experimental measurements of the eye's PSF (Navarro, 2009) have shown it to be characterized by a ring or star shape. Great deviations from this pattern are unlikely and have no physical basis supporting them; despite the fact that it could well satisfy a numerical solution. Along the same line of thought, Meitav & Ribak (2012) proposed a method for enhancing the contrast of high-resolution retinal images. For the reconstruction process they avoided distant lobes of the estimated PSF and used only the PSF area that would be under the central lobe and the first ring of the Airy pattern.

The second aspect to consider is the following. Because the images have been registered, the peak of maximum intensity of the PSF should tend to be close to the geometrical center of the PSF space. Or at least, because we assume the PSF to change smoothly, great shifts between adjacent PSFs are not expected. From this analysis we defined the following two criteria:

Criterion 1. Characterizing the energy distribution along the local PSF space. PSFs that are too wide, or have a high energy content concentrated far from the center, are most likely regions where the PSF estimation failed. We characterize the distribution by adding the PSF values along concentric squares and normalizing the resulting vector. Valid kernels should not have most of the energy concentrated in the last position.

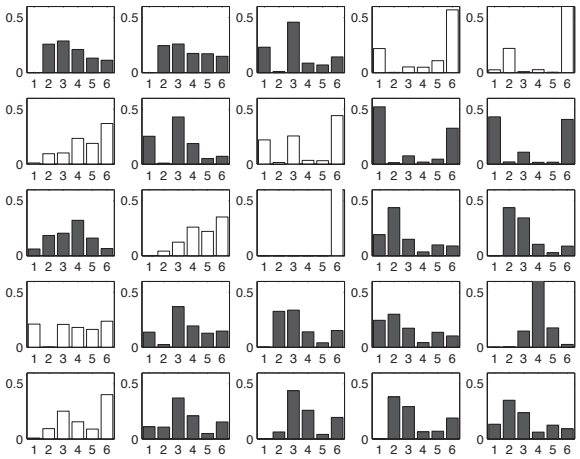


Figure 6.6: Characterization of estimated local PSFs by energy distribution. The ones plotted with white bars correspond to local PSFs with most of the energy concentrated around the boundaries.

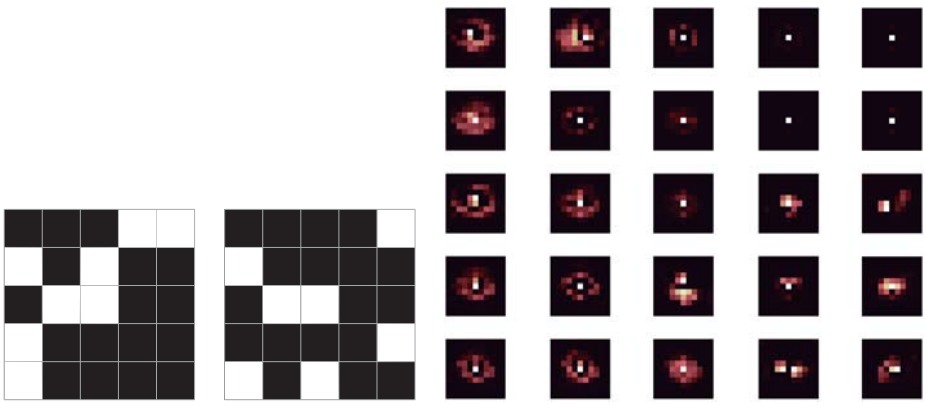
Criterion 2. Measuring the distance from the maximum peak of the local PSF to the geometrical center. Any PSF with its maximum at a distance greater than three pixels from the center, in a PSF size of 11×11 pixels, is discarded.

In Fig. 6.6 we show the PSF energy distribution and the identification of regions where the PSF estimation failed (regions with energy distribution plotted with white bars) based on Criterion 1. The regions identified by both criteria are represented by white squares in Figures 6.7(a) and (b). The overall criterion, for identifying the non-valid local PSFs, is the combination of both criteria (equivalent to a logical OR operation). The procedure for correcting the non-valid local PSFs consists in replacing them with the average of adjacent valid kernels. The resulting set of valid local PSFs is shown in Fig. 6.7(c).

6.3.4 PSF interpolation

The computation of the SV PSF h is carried out by interpolating the local PSFs estimated on the regular grid of positions. The PSF values at intermediate positions are computed by bilinear interpolation of four adjacent known PSFs, (Nagy & O'Leary, 1998) as shown in Fig. 6.8. Indexing any four adjacent grid points as $p = 1 \ldots 4$ (starting from the top-left corner and continuing clockwise), the SV PSF in the position between them is defined as

$$h(s,t;x,y) = \sum_{p=1}^{4} \alpha_p(x,y) h_p(s,t) \ , \tag{6.7}$$



(a) (b) (c)

Figure 6.7: Identification and replacement of non-valid local PSFs. The white squares correspond to non-valid local PSFs identified by: (a) Criterion 1, and (b) Criterion 2. (c) New corrected set of 5×5 local PSFs (compare with Fig. 6.5(b)).

where α_p are the coefficients of bilinear interpolation. Let us denote x_1 and x_2 minimum and maximum x-coordinates of the sub-window, respectively. Analogously, y_1 and y_2 in the y-coordinates. Using auxiliary quantities

$$t_x = \frac{x - x_1}{x_2 - x_1} \ , \quad t_y = \frac{y - y_1}{y_2 - y_1} \ , \tag{6.8}$$

the bilinear coefficients are

$$\alpha_1 = (1 - t_y)(1 - t_x) \ , \ \ \alpha_2 = (1 - t_y)t_x \ , \ \ \alpha_3 = t_y(1 - t_x) \ , \ \ \alpha_4 = t_y t_x \ . \tag{6.9}$$

In light of the definition for a SV PSF in Eq. (6.7) we can compute the convolution of Eq. (6.3) as a sum of four convolutions of the image weighted by coefficients $\alpha_p(x, y)$

$$\begin{aligned}
[Hu](x,y) &= \int u(s,t)h(s,t,x-s,y-t)\ dsdt \ , \\
&= \int u(s,t) \sum_{p=1}^{4} \alpha_p(s,t) h_p(x-s,y-t)\ dsdt \ , \\
&= \sum_{p=1}^{4} \int \left(\alpha_p(s,t)u(s,t)\right) h_p(x-s,y-t)\ dsdt \ , \\
&= \left[\sum_{p=1}^{4} [\alpha_p u] * h_p \right] (x,y) \ .
\end{aligned} \tag{6.10}$$

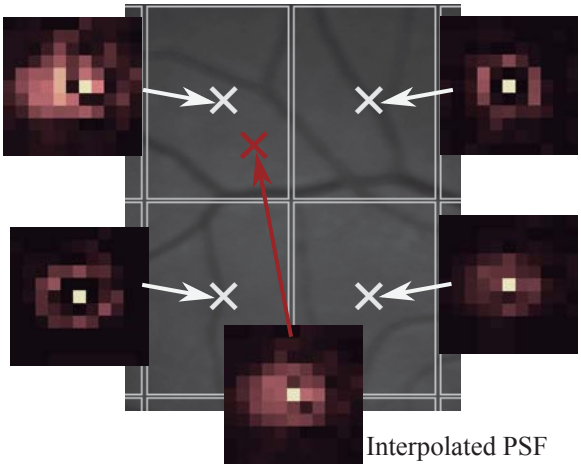


Figure 6.8: Because the blur changes gradually, we can estimate convolution kernels on a grid of positions and approximate the PSF in the rest of the image (bottom kernel) by interpolation from four adjacent kernels.

All first order minimization algorithms (as the one used in the restoration stage Eq. (6.12)) also need the operator adjoint to H (space-variant counterpart of correlation), which is defined as

$$\begin{aligned}
[H^* u](x,y) &= \int u(s,t) h(s-x, t-y, x, y)\, dsdt \,, \\
&= \int u(s,t) \sum_{p=1}^{4} \alpha_p(x,y) h_p(s-x, t-y)\, dsdt \,, \\
&= \sum_{p=1}^{4} \alpha_p(x,y) \int u(s,t) h_p(s-x, t-y)\, dsdt \,, \\
&= \sum_{p=1}^{4} \alpha_p(x,y) [u \circledast h_p](x,y) \;\;, \qquad (6.11)
\end{aligned}$$

where the symbol $\circledast$ represents correlation.

6.3.5 Restoration

Having estimated a reliable SV PSF we proceed to deblur the image. Image restoration is typically formulated within the Bayesian paradigm, in which the restored image is sought as the most probable solution to an optimization problem. The restoration then, can be described as the minimization of the functional

$$E(u) = \min_u \left[\frac{1}{2} \|z - Hu\|^2 + \lambda \int |\nabla u|\, dxdy \right] \;\;, \qquad (6.12)$$

where z is the blurred observed image, H is the blurring operator (Eq. (6.10)), u is the unknown sharp image, and λ is a positive regularization constant, which we have set according to a fine-tuning procedure (Marrugo et al., 2011a). The tuning procedure consists in artificially degrading a retinal image and restoring it with Eq. (6.12) by varying λ. Because the sharp original image is known we can compare it against the restored image using a metric like the peak-signal-to-noise ratio to determine an optimal value of λ. The first term penalizes the discrepancy between the model and the observed image. The second term is the regularization term which, as in Eq. (6.6), serves as a statistical prior. Once again we use total variation, a regularization technique that exploits the sparsity of image gradients in natural images. At present, solving the convex functional of Eq. (6.12) is considered a standard way to achieve close to state-of-the-art restoration quality without excessive time requirements (Campisi & Egiazarian, 2007). We used an efficient method (Chambolle & Lions, 1997) to solve Eq. (6.12) iteratively as a sequence of quadratic functionals

$$u_{i+1} = \arg\min_{u} \left[\frac{1}{2}\|z - Hu\|^2 + \lambda \int \frac{|\nabla u|^2}{2|\nabla u_i|} + \frac{|\nabla u_i|}{2}\, dxdy \right] \ . \tag{6.13}$$

The functional of Eq. (6.13) bounds the original function in Eq. (6.12) and has the same value and gradient in the current u_i, which guarantees convergence to the global minimum. To solve Eq. (6.13) we used the conjugate gradient method (Golub & Van Loan, 1996), in which the adjoint operator (Eq. (6.11)) is used in the gradient of the data term

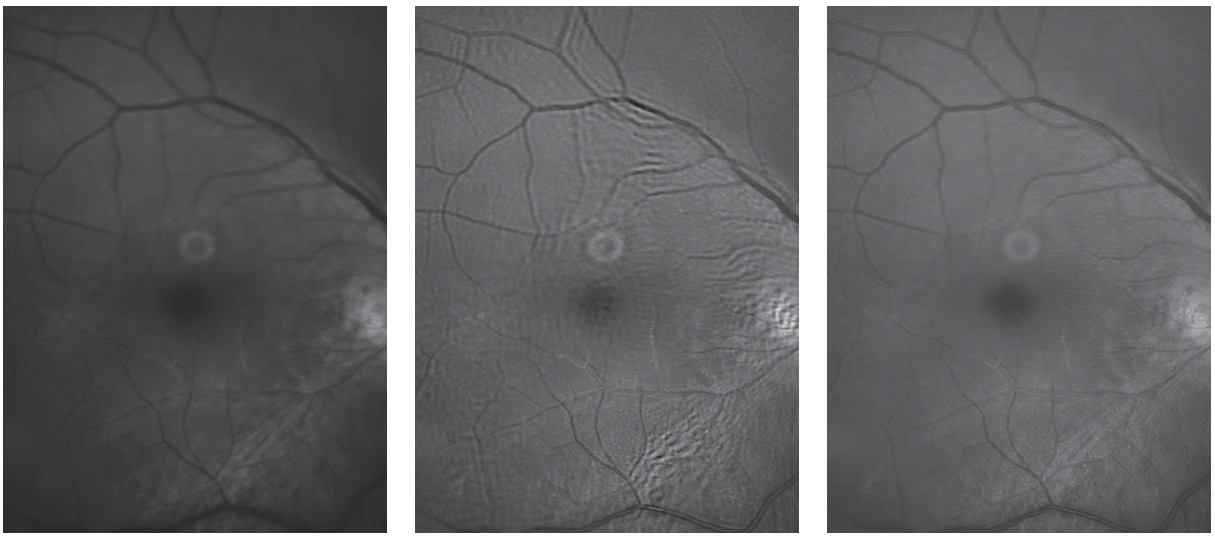
$$\frac{\partial}{\partial u}\frac{1}{2}\|z - Hu\|^2 = H^*(Hu - z) \ . \tag{6.14}$$

Using the operators H and H^*, in forms given by Equations (6.10) and (6.11), for large PSFs can be sped up significantly by computing convolutions and correlations using the Fast Fourier transform. For further details, see the third section in (Chambolle & Lions, 1997).

6.4 Experiments and Results

We performed several experiments to illustrate the appropriateness of the SV approach for restoring blurred retinal images. Moreover, to achieve an artifact free restoration we used our strategy for detecting the non-valid local PSFs and replacing them with a corrected one. All of the images used in the experiments were acquired within the same session, typically with a time span between acquisitions of several minutes.

Initially, to show the limits of the space-invariant approach we restored the blurred retinal image from Fig. 6.2(a) with a single global PSF with the space-invariant method we proposed in (Marrugo et al., 2011a). Recall that



(a) (b) (c)

Figure 6.9: (a) Original degraded image, (b) space-variant restoration with the PSFs of Fig. 6.5(b) which include non-valid elements and (c) space-variant restoration with the corrected PSFs of Fig. 6.7(c). Compare with space-invariant restoration shown in Fig. 6.3(a). The reader is strongly encouraged to view these images in full resolution in `http://www.goapi.upc.edu/usr/andre/sv-restoration/index.html`

this image corresponds to the eye fundus of a patient with strong astigmatism, which induces a SV blur. The restoration is shown in Fig. 6.3(a) and we can clearly observe various artifacts despite an increase in sharpness in a small number of areas. In view of this, it is evident that the the space-invariant assumption does not hold in such cases. In the following we move to the SV approach.

To carry out the SV restoration we estimated the local PSFs on a 5–by–5 grid of image patches as shown in Fig. 6.5. From the estimated PSFs, we notice a clear variation in shape mainly from the top right corner, where they are quite narrow, to the bottom corners where they are more spread and wide. This variation is consistent with the spatial variation of the blur observed in the retinal image of Fig. 6.2(a). We restored the image with these local PSFs that were estimated directly without any adjustment. The restored image is shown in Fig. 6.9(b). One immediately obvious feature is that in several areas the restoration is rather poor, displaying ringing artifacts, whereas in others it is to some extent satisfactory. The local poor-restoration is linked to areas where the PSF estimation failed. By removing and correcting these non-valid local PSFs, we obtained a noteworthy restoration shown in Fig. 6.9(c). Notice the overall improvement in sharpness and resolution with small blood vessels properly defined as shown by the image-

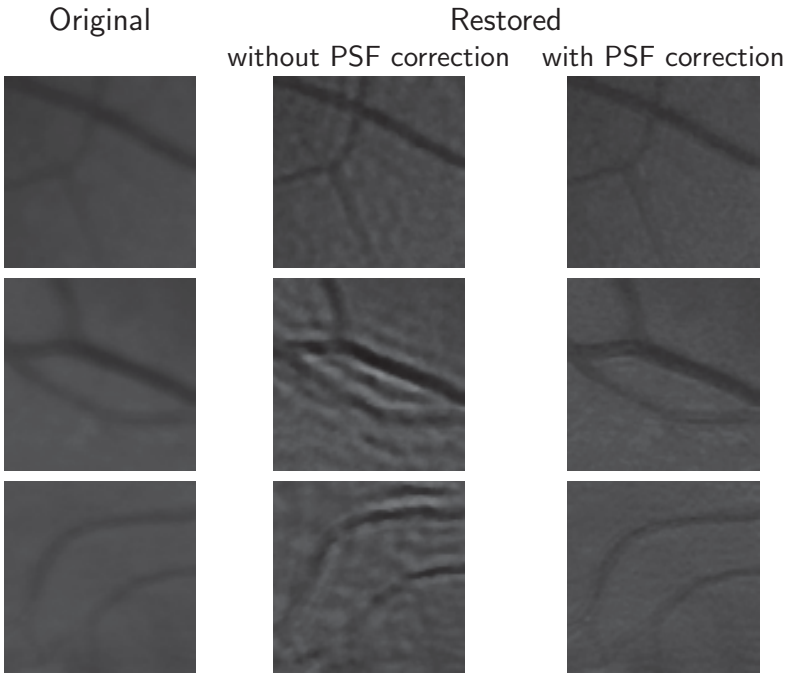


Figure 6.10: Details of restoration. From left to right: the original degraded image, the space-variant restoration without correction of PSFs, and the space-variant restoration with the correction.

details in the third column of Fig. 6.10. It could be said that without the replacement of the non-valid PSFs the image-quality after restoration is certainly worse than the original degraded image (see second column of Fig. 6.10).

To further demonstrate the capabilities of our method, additional restoration results on *real cases* from the clinical practice are shown in the following figures. A typical source of retinal image degradation comes from patients with corneal defects in which the cornea has an irregular structure (Fig. 6.11(a)). This induces optical aberrations, which are mainly responsible for the space-variant blur observed in the retinal image. The image details shown in Fig. 6.11(b) reveal a significant improvement in which the retinal structures are much sharper and enhanced. In Fig. 6.12(a) a full color retinal image is shown, in which three small hemorrhages are more easily discernible in the restored image, along with small blood vessels. Another retinal image, shown in Fig. 6.12(b), reveals a clear improvement in resolution with a much finer definition of blood vessels.

In addition, we processed retinal angiography images to test our method against a different imaging modality. Ocular angiography is a diagnostic test that documents by means of photographs the dynamic flow of dye in the blood vessels of the eye (Saine & Tyler, 2002). The ophthalmologists use

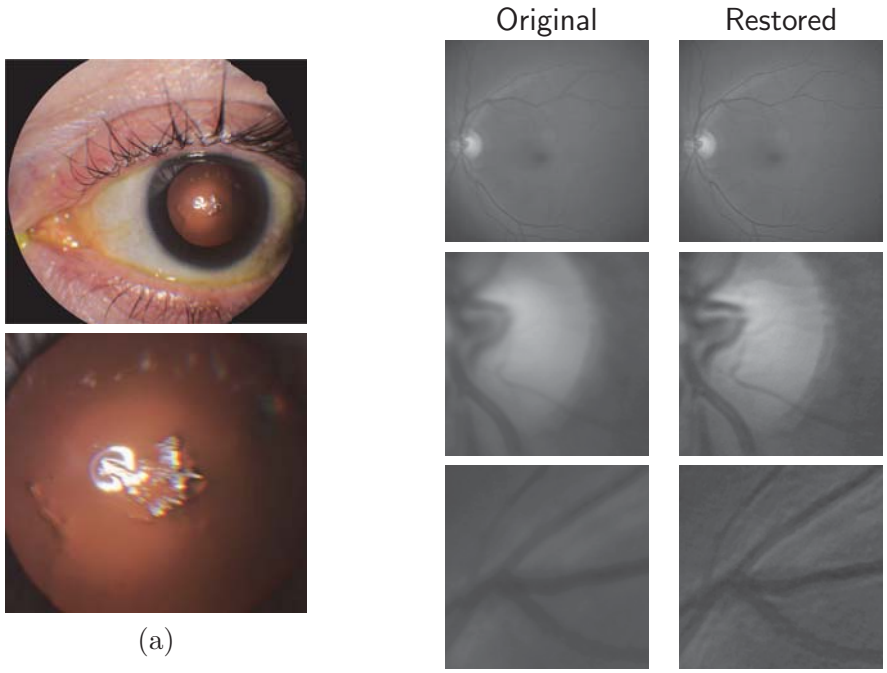


Figure 6.11: (a) Top: Eye with corneal defects that induce retinal images with SV degradation. Bottom: zoomed region. (b) Left column: original image with details. Right column: restored image with details.

these photographs both for diagnosis and as a guide to patient treatment. Ocular angiography differs from fundus photography in that it requires an exciter–barrier filter set (for further details see (Saine & Tyler, 2002)). The retinal angiography shown in Fig. 6.13 is degraded with a mild SV blur that hinders the resolution of small—yet important—details. The restoration serves to overcome this impediment; this can be observed from the zoomed-detail of the restored image. The image enhancement may be useful for the improvement of recent analysis techniques for automated flow dynamics and identification of clinical relevant anatomy in angiographies (Holmes et al., 2012).

Finally, another way to demonstrate the added value in deblurring the retinal images is to extract important features, in this case detection of blood vessels. Such a procedure is commonly used in many automated disease detection algorithms. The improvement in resolution paves the way for a better segmentation of structures with edges. This is in great part due to the effect of the total variation regularization because it preserves the edge information in the image. To carry out the detection of the retinal vasculature we used Kirsch's method (Kirsch, 1971). It is a matched filter algorithm that computes the gradient by convolution with the image and eight tem-

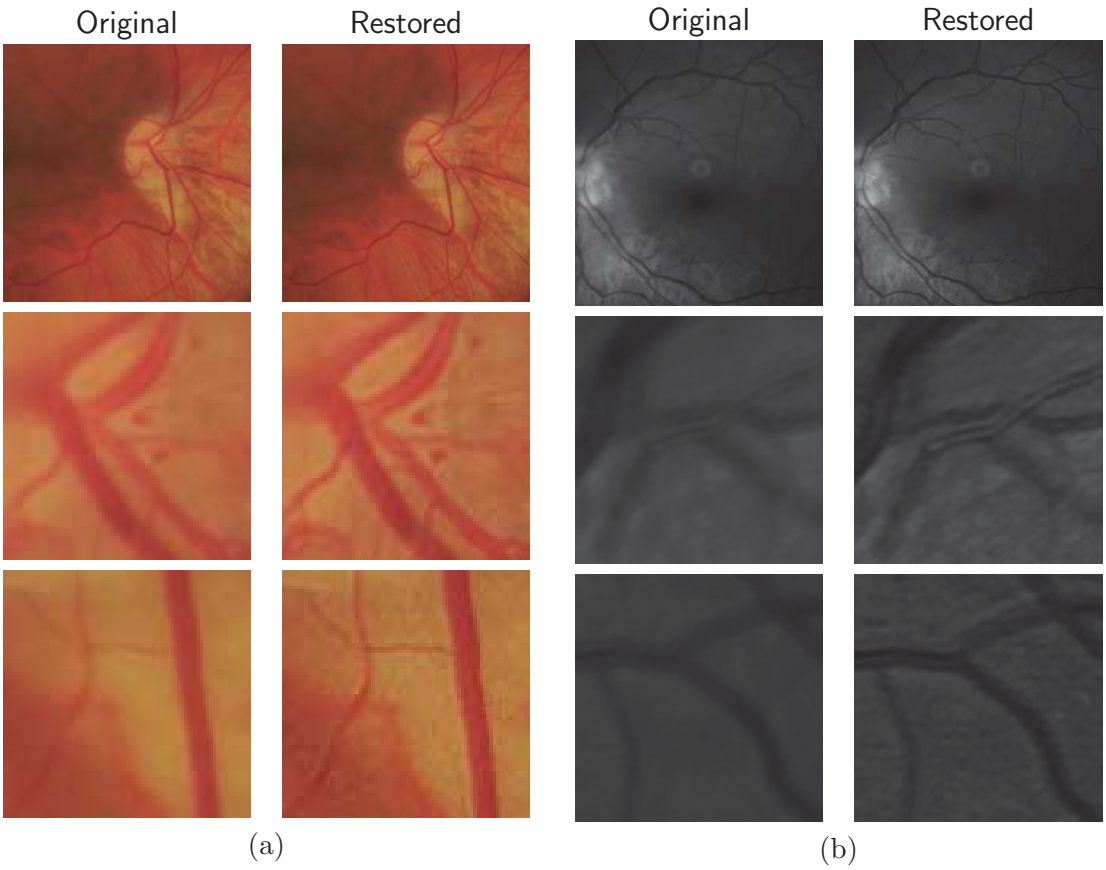


Figure 6.12: Other retinal images restored with the proposed method. First row: original and restored full-size retinal images. Second and third rows: image details.

plates to account for all possible directions. This algorithm has been widely used for detecting the blood vessels in retinal images (Al-Rawi et al., 2007). In Fig. 6.14 we show the detection of the blood vessels from a real image of poor quality image and its restored version using our proposed method. A significant improvement in blood vessel detection is achieved. Smaller blood vessels are detected in the restored image, whereas the detection from the original image barely covers the main branch of the vasculature.

6.5 Discussion

In this chapter we have introduced a method for restoring retinal images affected by space-variant blur by means of blind deconvolution. To do so, we described a spatially variant model of blur in terms of a convolution with a PSF that changes depending on its position. Since the space-variant degradation changes smoothly across the image, we have showed that the

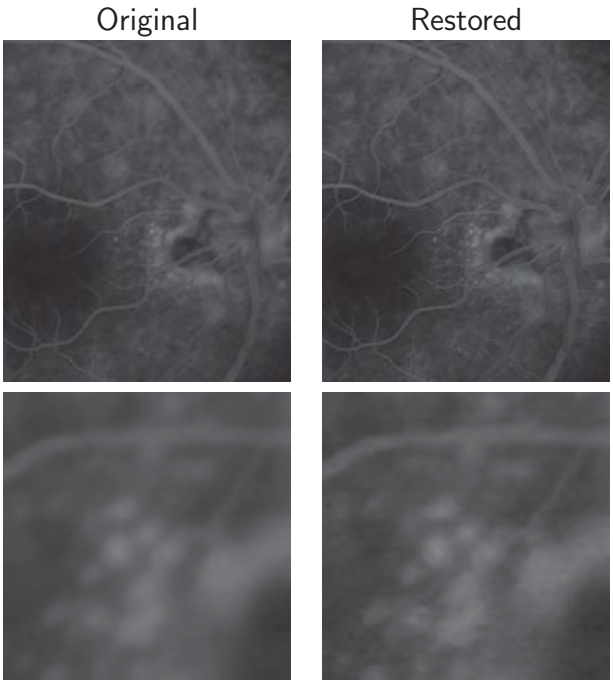


Figure 6.13: Restoration of a retinal angiography. First row: original and restored full retinal images. Second row: image details.

PSF needs not be computed for all pixels but for a small set of discrete positions. For any intermediate position bilinear interpolation suffices. In this way we achieve a space-variant representation of the PSF.

The estimation of accurate local PSFs proved difficult due to the very nature of the images; they usually contain textureless or nearly-homogenous regions that lack retinal structures, such as blood vessels, to provide sufficient information. In this regard we proposed a strategy based on eye-domain knowledge to adequately identify and correct such non-valid PSFs. Without this, the restoration results are artifact-prone. The proposal has been tested on a variety of real retinal images coming from the clinical practice. The details from the restored retinal images show an important enhancement, which is also demonstrated with the improvement in the detection of the retinal vasculature (see, for instance, the example of Fig. 6.14).

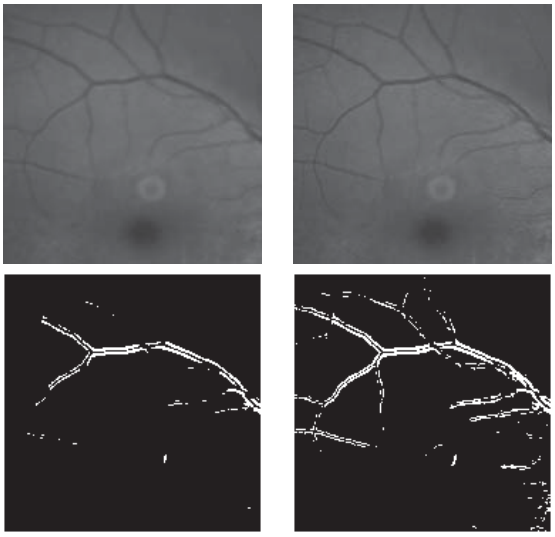


Figure 6.14: First row (from left to right): original and restored retinal images. Second row: detection of blood vessels. Notice how the small blood vessels are better detected in the restored image.

Chapter 7

Conclusions

Retinal image analysis is a constantly growing applied and interdisciplinary field in which any recent development in optics, medical imaging, image and signal processing, statistical analysis, etc. is likely to find its place to further augment the capabilities of the modern clinician. Its primary goal is to have automated image analysis and computer-aided diagnosis ubiquitous and effective enough that ultimately leads to the improvement of the health-care system, thus enhancing the health and quality of the general population.

When we started this thesis the aforementioned goal seemed rather far, today many strong initiatives are being developed to solidify the use of computer-aided diagnosis in many different medical fields. Ophthalmology is a field that is heavily dependent on the analysis of digital images because they can aid in establishing an early diagnosis even before the first symptoms appear. In this way it is easier to stop the development of many ocular diseases when they are in the early stages. As such we have made our best effort to contribute and find solutions to different problems along the imaging pipeline. We have dealt with problems arising in image acquisition (Chapters 3 and 4), poor image quality (Chapters 4, 5, and 6), and extraction of features of medical relevancy (Chapters 2 and 5). By improving the acquisition procedure, sorting the images by quality, or enhancing the images, among the different topics we have covered herein, we leverage the potential use of the images in the clinical setting. This has been our motivation throughout this thesis and the true common denominator of all chapters.

In every chapter we have tried to solve a particular problem, we have proposed a solution and in doing so we have opened new avenues for research that have been, to some extent, dealt with in subsequent chapters. For the most part, in my humble opinion this is how research is done, and if you allow me the simile, it is like trying to find the next piece of a puzzle to which you have little idea what it looks like until you find more pieces. This is why in this concluding chapter we revisit the key pieces of this thesis and

draw specific conclusions that hopefully will help the reader to form and identify the contributions within the overarching story of this thesis.

On Retinal Image Acquisition

Retinal image quality

Whether it be for medical interpretation or for automated analysis, good quality retinal images are *always* required in order to extract meaningful information that is medically relevant. As we discussed in Chapter 3, there are a number of different difficulties in the imaging process that lead to artifacts and images of poor quality. Sorting images in terms of quality is certainly an extremely difficult task due to the complexity of the problem, yet we studied it from a no-reference image quality metric perspective.

1. From the different state-of-the-art no-reference quality metrics that we considered, and their applicability to fundus imaging, we found the metric based on a measure of anisotropy proposed by Gabarda & Cristóbal (2007) to be the best. On top of this metric we proposed a modification that accounts for local quality requirements (by way of a weighting function, see § 3.2) which is commonplace in retinal imaging.

2. The performance of each metric was compared to that of two expert readers. From this we found the anisotropy measure to show the best agreement. It was not unexpected to find that the visibility of retinal structures, like the blood vessels, is one of the most important features to take into account when evaluating retinal image quality. Moreover, we did determine that anisotropy, by which we mean quantifying the fact that when an image is degraded the structures within change in a directional way, is a good indicator of image sharpness. This enabled us to rethink other strategies like those used while focusing for obtaining the sharpest image (Chapter 4).

Robust focusing

5. Following from our preliminary results on no-reference image quality metrics, we designed a focus measure for non-mydriatic fundus imaging that would take into account the anisotropic properties of image degradation (Chapter 4). The proposed focus measure is based on a weighted directional variance of the normalized discrete cosine transform. The normalization is carried out in order to achieve illumination invariance. This is due to the fact that illumination may change during acquisition because of the dynamic properties of the fundus imaging device and the eye itself.

6. Another important aspect that we considered was to achieve sufficient robustness to noise, so that the sharpest image is properly identified. Because the focusing is carried out in the near infrared, the images are noise-prone. To address this issue we optimized the weights with a test set of images corrupted with different types of noise. From this procedure we found that a strong emphasis should be put to mid-frequency coefficients (see § 4.3.3), rather than to high frequencies that are prone to noise contamination. The mid-frequency coefficients contain information of basic structure, but not sharp details in the high frequencies, that seem to be sufficient for correct focusing.

7. The focus measure performance was compared, by means of simulated and real images, to that of other standard focus measures (§ 4.2). The results clearly showed that our measure outperformed the considered measures in robustness and accuracy. Once again, measuring anisotropy proves to be a reliable approach for quantifying image degradation, and may be a practical approach for the design of the autofocus function in modern retinal cameras.

8. In our study we included several young and elderly subjects to assess the limitations of the proposed focus measure. We found no significant differences between the focusing curves (§ 4.4.2), although there was some difficulty in the acquisition of images from the elderly subjects handicapped by unstable fixation.

On Retinal Image Enhancement

There are a number of reasons for wanting to enhance (restore) retinal images, but the most important is that relevant information may be extracted from them. Fundus imaging is a complex process in which many things may go wrong (§ 3.1.2) and the end result is an image of poor quality. Just as there are certain types of degradations, there are also image enhancements tailored to address them.

Non-uniform illumination

Despite controlled conditions, many retinal images suffer from non-uniform illumination and poor contrast given by several factors like the curved surface of the retina or pupil dilation. The curved retinal surface and the geometrical configuration of the light source and camera, lead to a poorly illuminated peripheral part of the retina with respect to the central part.

9. To address this limitation we have developed a strategy based on the knowledge of the luminosity distribution in the retina (§ 3.3). Our strategy is based on the idea that a retinal image can be separated

into a foreground (composed of retinal structures like the vasculature or the optic disc) and a background (the fundus free of any structure). To make the illumination uniform across the whole field of view the background luminosity must be properly estimated. To this end, on top of the pixel classification procedure described by Foracchia et al. (2005), we performed additional *Principal Component Analysis* to correctly identify and leave out the optic disc region from the background estimation so as not to bias the luminosity component. The resulting enhanced images (Fig. 3.15) show a remarkable gain in contrast and a much more uniform illumination distribution, in spite of a minor decrease in overall intensity.

Retinal image deblurring

Blur is one of the main image quality degradations in eye fundus imaging. It significantly hinders the clinical use of the images. Image deconvolution is the appropriate way to address this problem and we have done so in two different—yet complementary—approaches. The most general case is that the blur varies across the whole field of view, which can be described by a space-variant *point-spread function* (PSF). In small regions where the variation of blur is negligible, the PSF is considered space-invariant. This is the so-called *isoplanatic* patch (Bedggood et al., 2008). In some cases if the blur varies little throughout the field of view, the blur of the whole image can be considered to be caused by a space-invariant PSF.

The space-invariant approach. It is the most common constraint of the deconvolution approaches reported in the literature. For this reason this was the first approach we considered when deblurring retinal images (Chapter 5).

10. Because the PSF that blurred the image is unknown we must make use of a blind deconvolution strategy. Specifically, we proposed an image enhancement strategy based on multichannel blind deconvolution (§ 5.2.1). The multichannel aspect refers to the use of two or more images for deconvolution which adds information redundancy and better poses the problem.

11. An important part of our approach is that we have been able to use pairs of images from the same patient acquired even months apart. To the best of our knowledge this has no precedent in the literature. This has been possible because in retinal imaging the acquisition conditions change little. Nonetheless, in our approach we have to preprocess the images to adjust them and have them comply with the degradation model (§ 5.3). Regardless that the acquisitions conditions remain relatively constant in time, the retina can indeed change locally due to

pathological developments (hemorrhages, micro-aneurysms, etc.). To address this problem we developed a strategy for identifying and masking such changes because convolution with a PSF cannot account for them (§ 5.4.3).

12. We tested our approach on a set of synthetically degraded images and obtained significant enhancements. When compared with a state-of-the-art single-channel blind deconvolution, the multi-channel scheme is considerably superior (§ 5.5.1). Results on real images showed a remarkable enhancement evidenced by the increase in visibility of details such as small blood vessels or pathological structures (§ 5.5.2).

The space-variant approach. There are some situations in which the space-invariant approach fails because the blur changes considerably throughout the field of view. For that reason we have proposed a space-variant retinal image enhancement technique (Chapter 6).

13. The technique is based on the fact that the space-variant PSF can be locally considered space-invariant. This means that we can estimate local PSFs by dividing the image in a grid of small patches. However, unlike other works, we do not deconvolve the image per patch and later stitch the patches together to produce the restored image. We do estimate the PSFs per patch, with the approach described in Chapter 5, but later sew them through interpolation to build the space-variant PSF.

14. The estimation of the local PSFs may sometimes fail due to the properties of retinal images; i.e. they usually contain textureless or nearly homogenous regions that lack retinal structures to provide sufficient information. To address this problem we have designed a strategy that takes into account two criteria based on the knowledge of the PSF of the eye and the conditions of acquisition. Without this process, the deconvolution results are artifact-prone.

    - PSFs that are too wide, or have a high energy content concentrated far from the center, are most likely regions where the PSF estimation failed. The distribution is characterized by adding the PSF values along concentric squares and normalizing the resulting vector. Valid kernels should not have most of the energy concentrated in the last position.
    - Measuring the distance from the maximum peak of the local PSF to the geometrical center. Any PSF with its maximum at a distance greater than a predetermined threshold distance from the center is discarded. The threshold distance depends on the PSF size.

15. The deconvolution is carried out with the space-variant PSF in the way detailed in § 6.3.5. We have tested our approach on a variety of real retinal images coming from the clinical practice. The details from the restored images show a remarkable enhancement in visibility of subtle details like small blood vessels (Fig. 6.10). This has also been demonstrated with the improvement in the detection of the retinal vasculature (Fig. 6.14). We believe this to be a significant contribution that can leverage the clinical use of blurred retinal images, that would otherwise be discarded because of low quality.

On Retinal Image Feature Extraction

Optic disc segmentation

As we mentioned in Chapter 2, segmentation of the optic disc (or optic nerve head) was the first problem we have addressed in this thesis because previous efforts had been carried out in our research group.

16. To that end we have analyzed the work of Valencia et al. (2006) and have determined possible drawbacks. Although an important contribution, the authors assumed the optic disc contour of approximately circular shape, which is a limitation when segmenting optic discs of a more elongated shape. In addition, because Valencia et al. (2006) proposed the extraction of the contour by means of a pixel-wise color difference approach, we have studied how lossy compression (classic JPEG and JPEG-2000) may influence the segmentation. Our results have showed that such an approach is not advisable if the images have been lossy compressed with mid-to-high compression ratios. However, by comparing the results of classic JPEG versus JPEG-2000, the latter seems more suitable due to the fact that segmentation results have been more reliable and reproducible.

17. After having analyzed the approach of Valencia et al. (2006) we determined that a more general and robust approach was needed to perform the segmentation of the optic disc. For this reason we have proposed in § 2.3 a segmentation scheme based on active contours and mathematical morphology in the color space. Because one of the main difficulties in segmenting the optic disc lies in the lack of homogeneity in the optic disc region, it is fragmented into multiple subregions by blood vessels, we have carried out a closing operation to remove the blood vessels. This was successfully achieved (Fig. 2.9) by taking into account the color properties of the region. Another important aspect of this preprocessing is that any lossy compression artifact that may be present in the image is removed because of the closing operation that yields a

rather uniform optic disc without blood vessels. Our results showed that if the preprocessing is successful, the active contour locks and segments the optic disc within acceptable accuracy when compared to the hand-labelled ground-truth produced by an expert (§ 2.3.3).

Longitudinal change detection

18. Our proposed approach for deblurring retinal images (Chapter 5) required a number of preprocessing steps, which in turn became a great opportunity to meet one of the main concerns of ophthalmologists when they visually compare retinal images of the same retina over time. Central to the task of determining disease progression is the distinction of true change from variability. The detection of longitudinal changes in retinal images is not an easy task, and even today it is still performed manually (Xiao et al., 2012, Narasimha-Iyer et al., 2006), although the use of computer-aided diagnosis systems is becoming a reality in the clinical setting.

19. By properly registering the images and compensating for illumination distribution differences, we have been able to identify true structural changes pertaining to possible pathological damage or healing. Thus, any change caused by variation of illumination or blur is disregarded. Moreover, the statistical significance test, following the underlying idea that changes must be associated with a group of pixels, proved sufficient for identifying the change and no-change regions (§ 5.4.3). This is a powerful tool for the clinician so that his attention is drawn to the area of interest and further investigates the causes of such changes (Fig. 5.6).

Future Research

The results of this thesis have opened several new avenues for research and applications. Whenever possible we have tried to address such concerns herein, but it is clear that there is still room for improvement. For instance, in the case of retinal image enhancement a possible application is found in the restoration of stereo retinal images for depth estimation. Most stereo images do not satisfy the brightness constancy assumption along with the expected blurring of some parts of the images because photographers find it difficult to focus two images simultaneously. Research can also be conducted to compare the restoration results with deconvolution from wavefront sensing fundus imagers to determine if our methods could be a suitable and inexpensive alternative. Determining a robust approach for estimating a reliable point-spread function (whether space-variant or invariant) is still a open issue and we are currently pursuing such line of work.

Even though we have used real images from the clinical practice, it is evident that the proposed algorithms need further validation on a greater set of patients. This will probably show aspects worth improving and to verify its statistical validity. This requires a collaboration work with other research groups in the field of ophthalmology with access to the clinical practice.

Another line of work is the use of images from multiple modalities. This enables the extraction of information from different sources, even from several not directly related to photographic fundus imaging, like Optical Coherence Tomography, Polarimetric imaging, laser scanning, among others.

Finally, we recognize that the one of the greatest challenges in retinal image analysis comes from the inherent variability of the appearance of the retina throughout the general population. We believe this to be one of the key aspects that offers opportunities for further research. Adapting our methods to such variation is a goal worth pursuing.

Bibliography

Aach, T. & Kaup, A. (1995). Bayesian Algorithms for Change Detection in Image Sequences Using Markov Random Fields. *Signal Processing: Image Communication*, 7(2), 147–160.

Abramoff, M. D., Garvin, M., & Sonka, M. (2010). Retinal Imaging and Image Analysis. *Biomedical Engineering, IEEE Reviews in*, 3, 169–208.

Abramoff, M. D., Niemeijer, M., Suttorp-Schulten, M. S. A., Viergever, M. A., Russell, S. R., & van Ginneken, B. (2008). Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes. *Diabetes care*, 31(2), 193–198.

Abràmoff, M. D., Reinhardt, J. M., Russell, S. R., Folk, J. C., Mahajan, V. B., Niemeijer, M., & Quellec, G. (2010). Automated early detection of diabetic retinopathy. *Ophthalmology*, 117(6), 1147–1154.

Agrawal, A. & McKibbin, M. A. (2003). Technical failure in photographic screening for diabetic retinopathy. *Diabetic medicine : a journal of the British Diabetic Association*, 20(9), 777.

Al-Rawi, M., Qutaishat, M., & Arrar, M. (2007). An improved matched filter for blood vessel detection of digital retinal images. *Computers in Biology and Medicine*, 37(2), 262–267.

Aslantas, V. & Kurban, R. (2009). A comparison of criterion functions for fusion of multi-focus noisy images. *Optics Communications*, 282(16), 3231–3242.

Balicki, M., Han, J.-H., Iordachita, I., Gehlbach, P., Handa, J., Taylor, R., & Kang, J. (2009). Single fiber optical coherence tomography microsurgical instruments for computer and robot-assisted retinal surgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009* (pp. 108–115). Springer.

Bardsley, J., Jefferies, S., Nagy, J., & Plemmons, R. (2006). A computational method for the restoration of images with an unknown, spatially-varying blur. *Optics express*, 14(5), 1767–1782.

Bartling, H., Wanger, P., & Martin, L. (2009). Automated quality evaluation of digital fundus photographs. *Acta Ophthalmologica*, 87(6), 643–647.

Baudoin, C., Lay, B., & Klein, J. (1984). Automatic detection of microaneurysms in diabetic fluorescein angiography. *Revue d'épidémiologie et de santé publique*, 32(3-4), 254.

Bedggood, P., Daaboul, M., Ashman, R., Smith, G., & Metha, A. (2008). Characteristics of the human isoplanatic patch and implications for adaptive optics retinal imaging. *Journal of Biomedical Optics*, 13(2), 024008.

Bennett, T. J. & Barry, C. J. (2009). Ophthalmic imaging today: an ophthalmic photographer's viewpoint - a review. *Clinical & Experimental Ophthalmology*, 37(1), 2–13.

Bernardes, R., Serranho, P., & Lobo, C. (2011). Digital Ocular Fundus Imaging: A Review. *Ophthalmologica*, 226, 161–181.

Bhargava, M., Cheung, C., Sabanayagam, C., Kawasaki, R., Harper, C., Lamoureux, E., Chow, W., Ee, A., Hamzah, H., Ho, M., et al. (2012). Accuracy of diabetic retinopathy screening by trained non-physician graders using non-mydriatic fundus camera. *Singapore medical journal*, 53(11), 715–719.

Bock, R., Meier, J., Nyúl, L. G., Hornegger, J., & Michelson, G. (2010). Glaucoma risk index: Automated glaucoma detection from color fundus images. *Medical Image Analysis*, 14(3), 471–481.

Boucher, M. C., Gresset, J. A., Angioi, K., & Olivier, S. (2003). Effectiveness and safety of screening for diabetic retinopathy with two nonmydriatic digital images compared with the seven standard stereoscopic photographic fields. *Canadian journal of ophthalmology Journal canadien d'ophtalmologie*, 38(7), 557–568.

Bruce, B. B., Lamirel, C., Wright, D. W., Ward, A., Heilpern, K. L., Biousse, V., & Newman, N. J. (2011). Nonmydriatic ocular fundus photography in the emergency department. *New England Journal of Medicine*, 364(4), 387–389.

Bruce, B. B., Newman, N. J., Pérez, M. A., & Biousse, V. (2013). Nonmydriatic ocular fundus photography and telemedicine: Past, present, and future. *Neuro-Ophthalmology*, 37(2), 51–57.

Burns, S. A., Tumbar, R., Elsner, A. E., Ferguson, D., & Hammer, D. X. (2007). Large-field-of-view, modular, stabilized, adaptive-optics-based scanning laser ophthalmoscope. *Journal of the Optical Society of America A, Optics, image science, and vision*, 24(5), 1313–1326.

Campisi, P. & Egiazarian, K. (2007). *Blind image deconvolution: theory and applications*. Boca Raton, Fl, USA: CRC Press.

Can, A., Stewart, C., Roysam, B., & Tanenbaum, H. (2002). A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(3), 347–364.

Catlin, D. & Dainty, C. (2002). High-resolution imaging of the human retina with a Fourier deconvolution technique. *Journal of the Optical Society of America A, Optics, image science, and vision*, 19(8), 1515–1523.

Chambolle, A. & Lions, P. L. (1997). Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2), 167–188.

Chan, T. & Vese, L. (2001). Active contours without edges. *Image Processing, IEEE Transactions on*, 10(2), 266–277.

Chang, C.-C., Chia, T.-L., & Yang, C.-K. (2005). Modified temporal difference method for change detection. *Optical Engineering*, 44(2), 027001.

Chen, W., Er, M. J., & Wu, S. (2006). Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 36(2), 458–466.

Chenegros, G., Mugnier, L., Lacombe, F., & Glanc, M. (2007). 3D phase diversity: a myopic deconvolution method for short-exposure images: application to retinal imaging. *Journal of the Optical Society of America A*, 24(5), 1349–1357.

Choi, K., Lee, J., & Ko, S. (1999). New autofocusing technique using the frequency selective weighted median filter for video cameras. *Consumer Electronics, IEEE Transactions on*, 45(3), 820–827.

Christou, J., Roorda, A., & Williams, D. (2004). Deconvolution of adaptive optics retinal images. *Journal of the Optical Society of America A*, 21(8), 1393–1401.

Clunie, D. A. (2000). Lossless compression of grayscale medical images-effectiveness of traditional and state of the art approaches. *Proceedings SPIE*, 3980, 74–84.

Conrath, J., Erginay, A., Giorgi, R., Lecleire-Collet, A., Vicaut, E., Klein, J.-C., Gaudric, A., & Massin, P. (2007). Evaluation of the effect of JPEG and JPEG2000 image compression on the detection of diabetic retinopathy. *Eye*, 21(4), 487–493.

Costello, T. & Mikhael, W. (2003). Efficient restoration of space-variant blurs from physical optics by sectioning with modified Wiener filtering. *Digital Signal Processing*, 13(1), 1–22.

Ebrahimi, F., Chamik, M., & Winkler, S. (2004). JPEG vs. JPEG2000: An objective comparison of image encoding quality. *Proceedings SPIE*, 5558, 300–308.

Fadzil, M., Nugroho, H., Venkatachalam, P., Nugroho, H., & Izhar, L. (2008). Determination of Retinal Pigments from Fundus Images using Independent Component Analysis. *4th Kuala Lumpur International Conference on Biomedical Engineering*, (pp. 555–558).

Fagin, R., Kumar, R., & Sivakumar, D. (2003). Comparing top k lists. *SIAM Journal on Discrete Mathematics*, 17(1), 134–160 (electronic).

Faust, O., Acharya, R., Ng, E., Ng, K.-H., & Suri, J. S. (2012). Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review. *Journal of medical systems*, 36(1), 145–157.

Feng, P., Pan, Y., Wei, B., Jin, W., & Mi, D. (2007). Enhancing retinal image by the Contourlet transform. *Pattern Recognition Letters*, 28(4), 516–522.

Ferzli, R. & Karam, L. J. (2009). A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). *Image Processing, IEEE Transactions on*, 18(4), 717–728.

Flandrin, P., Baraniuk, R. G., & Michel, O. (1994). Time-frequency complexity and information. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, volume 3 (pp. III–329).: IEEE.

Fleming, I., Balicki, M., Koo, J., Iordachita, I., Mitchell, B., Handa, J., Hager, G., & Taylor, R. (2008). Cooperative robot assistant for retinal microsurgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008* (pp. 543–550). Springer.

Foracchia, M., Grisan, E., & Ruggeri, A. (2005). Luminosity and contrast normalization in retinal images. *MEDICAL IMAGE ANALYSIS*, 9(3), 179–190.

Gabarda, S. & Cristóbal, G. (2007). Blind image quality assessment through anisotropy. *Journal of the Optical Society of America A, Optics, image science, and vision*, 24(12), B42–51.

Giancardo, L., Abramoff, M. D., Chaum, E., Karnowski, T., Meriaudeau, F., & Tobin, K. (2008). Elliptical local vessel density: A fast and robust

quality metric for retinal images. *Annual Int Conf of the IEEE Engineering in Medicine and Biology Society*, 1, 3534.

Giancardo, L., Meriaudeau, F., & Karnowski, T. (2010). Quality Assessment of Retinal Fundus Images using Elliptical Local Vessel Density. *inspire.ornl.gov*.

Golub, G. & Van Loan, C. (1996). *Matrix computations*, volume 3. Johns Hopkins University Press.

Goodman, J. W. (1968). *Introduction to Fourier optics*, volume 2. McGraw-hill New York.

Granville, V., Krivanek, M., & Rasson, J.-P. (1994). Simulated annealing: a proof of convergence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6), 652–656.

Groen, F., Young, I., & Ligthart, G. (1985). A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, 6(2), 81–91.

Gupta, A., Joshi, N., Lawrence Zitnick, C., Cohen, M., & Curless, B. (2010). Single image deblurring using motion density functions. *Computer Vision–ECCV 2010*, (pp. 171–184).

Harizman, N., Oliveira, C., Chiang, A., Tello, C., Marmor, M., Ritch, R., & Liebmann, J. M. (2006). The ISNT rule and differentiation of normal from glaucomatous eyes. *Arch. Ophthalmol*, 124, 1579–1583.

Harmeling, S., Hirsch, M., & Scholkopf, B. (2010). Space-variant single-image blind deconvolution for removing camera shake. *Advances in Neural Inform. Processing Syst.*

Heijl, A., Leske, M. C., Bengtsson, B., Hyman, L., Bengtsson, B., Hussein, M., et al. (2002). Reduction of intraocular pressure and glaucoma progression: results from the early manifest glaucoma trial. *Archives of Ophthalmology*, 120(10), 1268.

Herbert, H. M., Jordan, K., & Flanagan, D. W. (2003). Is screening with digital imaging using one retinal view adequate? *Eye*, 17(4), 497–500.

Holmes, T., Invernizzi, A., Larkin, S., & Staurenghi, G. (2012). Dynamic indocyanine green angiography measurements. *Journal of Biomedical Optics*, 17(11), 116028.

Hubbard, L. D., Brothers, R. J., King, W. N., Clegg, L. X., Klein, R., Cooper, L. S., Sharrett, A. R., Davis, M. D., & Cai, J. (1999). Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study. *Ophthalmology*, 106(12), 2269–2280.

Johnson, G. M. & Fairchild, M. D. (2003). A top down description of S-CIELAB and CIEDE2000. *Color Research & Application*, 28(6), 425–435.

Joshi, G. & Sivaswamy, J. (2008). Colour Retinal Image Enhancement Based on Domain Knowledge. *Computer Vision, Graphics & Image Processing, 2008. ICVGIP '08. Sixth Indian Conference on*, (pp. 591–598).

Kanski, J. (2005). *Diseases of the ocular fundus*. New York, NY: Elsevier/Mosby.

Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4), 321–331.

Kautsky, J., Flusser, J., Zitová, B., & Simberova, S. (2002). A new wavelet-based measure of image focus. *Pattern Recognition Letters*, 23(14), 1785–1794.

Kirsch, R. A. (1971). Computer determination of the constituent structure of biological images. *Computers and biomedical research*, 4(3), 315–328.

Kristan, M., Pers, J., Perse, M., & Kovacic, S. (2006). A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform. *Pattern Recognition Letters*, 27(13), 1431–1439.

Kumar, S., Wang, E.-H., Pokabla, M. J., & Noecker, R. J. (2012). Teleophthalmology assessment of diabetic retinopathy fundus images: smartphone versus standard office computer workstation. *TELEMEDICINE and e-HEALTH*, 18(2), 158–162.

Kundur, D. & Hatzinakos, D. (1996). Blind image deconvolution. *Signal Processing Magazine, IEEE*, 13(3), 43–64.

Larichev, A. V., Irochnikov, N. G., Ivanov, P., & Kudryashov, A. V. (2001). Deconvolution of color retinal images with wavefront sensing. *Adaptive Optical Systems Technology*, 4251, 102.

Lee, N., Smith, R., & Laine, A. (2008). Interactive segmentation for geographic atrophy in retinal fundus images. *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, (pp. 655–658).

Lee, S., Reinhardt, J., Cattin, P., & Abramoff, M. D. (2010). Objective and expert-independent validation of retinal image registration algorithms by a projective imaging distortion model. *MEDICAL IMAGE ANALYSIS*, 14(4), 539–549.

Levin, A., Weiss, Y., Durand, F., & Freeman, W. (2011). Understanding Blind Deconvolution Algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(12), 2354–2367.

Li, H. & Chutatape, O. (2003). Automatic detection and boundary estimation of the optic disk in retinal images using a model-based approach. *Journal of Electronic Imaging*, 12(1), 97–105.

Lian, Z. & Er, M. (2010). Illumination normalisation for face recognition in transformed domain. *Electronics Letters*, 46(15), 1060–1061.

Liatsis, P. & Kantartzis, P. (2005). Real-time colour segmentation and autofocus in retinal images. *ELMAR, 2005. 47th International Symposium*, (pp. 13–18).

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

Maberley, D., Morris, A., Hay, D., Chang, A., Hall, L., & Mandava, N. (2004). A comparison of digital retinal image quality among photographers with different levels of training using a non-mydriatic fundus camera. *Ophthalmic Epidemiology*, 11(3), 191–197.

Marrugo, A. G. & Millán, M. S. (2011). Retinal image analysis: preprocessing and feature extraction. *Journal of Physics: Conference Series*, 274(1), 012039.

Marrugo, A. G., Millán, M. S., & Abril, H. C. (2012a). Implementation of an Image Based Focusing Algorithm for Retinal Imaging. In *X Reunión Nacional de Óptica* (pp. 40–43). Zaragoza.

Marrugo, A. G., Millán, M. S., Cristóbal, G., Gabarda, S., & Abril, H. C. (2012b). Anisotropy-based robust focus measure for non-mydriatic retinal imaging. *Journal of Biomedical Optics*, 17(7), 076021.

Marrugo, A. G., Sorel, M., Sroubek, F., & Millán, M. S. (2011a). Retinal image restoration by means of blind deconvolution. *Journal of Biomedical Optics*, 16(11), 116016.

Marrugo, A. G., Sroubek, F., Sorel, M., & Millán, M. S. (2011b). Multichannel blind deconvolution in eye fundus imaging. In *ISABEL '11-Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies* (pp. 7:1–7:5).: New York, NY, USA.

MATLAB (2010). *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc.

McDonnell, P. J. (2010). Editorial the revenge of the machines:'the retinator'. *Ophthalmology Times*, 35(13), 4.

Meitav, N. & Ribak, E. N. (2012). Estimation of the ocular point spread function by retina modeling. *Optics Letters*, 37(9), 1466–1468.

Mendels, F., Heneghan, C., Harper, P., Reilly, R., & Thiran, J. (1999). Extraction of the optic disk boundary in digital fundus images. *Proc. 1st Joint BMES/EMBS Conf*, (pp. 1139).

Millán, M. S. & Valencia, E. (2006). Color image sharpening inspired by human vision models. *Applied Optics*, 45(29), 7684–7697.

Moscaritolo, M., Jampel, H., Knezevich, F., & Zeimer, R. (2009). An Image Based Auto-Focusing Algorithm for Digital Fundus Photography. *Medical Imaging, IEEE Transactions on*, 28(11), 1703–1707.

Muramatsu, C., Hayashi, Y., Sawada, A., Hatanaka, Y., Hara, T., Yamamoto, T., & Fujita, H. (2010). Detection of retinal nerve fiber layer defects on retinal fundus images for early diagnosis of glaucoma. *Journal of Biomedical Optics*, 15(1), 016021.

Nagy, J. G. & O'Leary, D. P. (1998). Restoring images degraded by spatially variant blur. *SIAM Journal on Scientific Computing*, 19(4), 1063–1082 (electronic).

Narasimha-Iyer, H., Can, A., Roysam, B., Stewart, C., Tanenbaum, H., Majerovics, A., & Singh, H. (2006). Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy. *Biomedical Engineering, IEEE Transactions on*, 53(6), 1084–1098.

Narasimha-Iyer, H., Can, A., Roysam, B., Tanenbaum, H., & Majerovics, A. (2007). Integrated Analysis of Vascular and Nonvascular Changes From Color Retinal Fundus Image Sequences. *Biomedical Engineering, IEEE Transactions on*, 54(8), 1436–1445.

Navarro, R. (2009). The Optical Design of the Human Eye: a Critical Review. *J Optom*, 2, 3–18.

Nayar, S. & Nakagawa, Y. (1994). Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8), 824–831.

Ng Kuang Chern, N., Neow, P. A., & Ang, M. (2001). Practical issues in pixel-based autofocusing for machine vision. In *IEEE Int. Conf. on Robotics and Automation* (pp. 2791–2796).

Osareh, A., Mirmehdi, M., Thomas, B., & Markham, R. (2002). Comparison of colour spaces for optic disc localisation in retinal images. *Pattern Recognition, Proceedings. 16th International Conference on*, 1, 743–746.

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *Systems, Man and Cybernetics, IEEE Transactions on*, 9(1), 62–66.

Patton, N., Aslam, T., MacGillivray, T., Deary, I., Dhillon, B., Eikelboom, R., Yogesan, K., & Constable, I. (2006). Retinal image analysis: Concepts, applications and potential. *Progress in Retinal and Eye Research*, 25(1), 99–127.

Primot, J., Rousset, G., & Fontanella, J. (1990). Deconvolution from wave-front sensing: a new technique for compensating turbulence-degraded images. *JOSA A*, 7(9), 1598–1608.

Qu, Y., Pu, Z., Zhao, H., & Zhao, Y. (2006). Comparison of different quality assessment functions in autoregulative illumination intensity algorithms. *Optical Engineering*, 45, 117201.

Radke, R., Andra, S., Al-Kofahi, O., & Roysam, B. (2005). Image change detection algorithms: a systematic survey. *Image Processing, IEEE Transactions on*, 14(3), 294–307.

Ramirez, J., Garcia, A., Fernandez, P., Parrilla, L., & Lloris, A. (2000). A new architecture to compute the discrete cosine transform using the quadratic residue number system. *IEEE International Symposium on Circuits and Systems*, 5, 321–324 vol.5.

Rényi, A. (1976). Some fundamental questions of information theory. *Selected Papers of Alfred Renyi*, 2(174), 526–552.

Richardson, W. H. (1972). Bayesian-Based Iterative Method of Image Restoration. *J. Opt. Soc. Am.*, 62(1), 55–59.

Rudin, L., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4), 259–268.

Saine, P. & Tyler, M. (2002). *Ophthalmic photography: retinal photography, angiography, and electronic imaging*. Woburn, MA, USA: Butterworth-Heinemann.

Salem, N. & Nandi, A. (2007). Novel and adaptive contribution of the red channel in pre-processing of colour fundus images. *Journal of the Franklin Institute*, 344(3-4), 243–256.

Sang, T.-H. & Williams, W. J. (1995). Renyi information and signal-dependent optimal kernel design. In *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, volume 2 (pp. 997–1000).: IEEE.

Sheppard, C. (2007). Fundamentals of superresolution. *Micron*, 38(2), 165–169.

Sobri, M., Lamont, A., Alias, N., & Win, M. (2003). Red flags in patients presenting with headache: clinical indications for neuroimaging. *British journal of radiology*, 76(908), 532–535.

Sroubek, F. & Flusser, J. (2003). Multichannel blind iterative image restoration. *Image Processing, IEEE Transactions on*, 12(9), 1094–1106.

Sroubek, F. & Flusser, J. (2005). Multichannel blind deconvolution of spatially misaligned images. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 14(7), 874–883.

Sroubek, F. & Milanfar, P. (2012). Robust Multichannel Blind Deconvolution via Fast Alternating Minimization. *Image Processing, IEEE Transactions on*, 21(4), 1687–1700.

Stern, G. A. et al. (1995). Teaching ophthalmology to primary care physicians. *Archives of ophthalmology*, 113(6), 722.

Stewart, C., Tsai, C.-L., & Roysam, B. (2003). The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *Medical Imaging, IEEE Transactions on*, 22(11), 1379–1394.

Stigler, S. (2008). Fisher and the 5 *Chance*, 21(4), 12–12.

Subbarao, M., Choi, T., & Nikzad, A. (1993). Focusing techniques. *Optical Engineering*, 32(11), 2824–2836.

Subbarao, M. & Tyan, J.-K. (1998). Selecting the optimal focus measure for autofocusing and depth-from-focus. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(8), 864–870.

Suppapitnarm, A., Seffen, K., Parks, G., & Clarkson, P. (2000). A Simulated Annealing Algorithm for Multiobjective Optimization. *Engineering Optimization*, 33(1), 59–85.

Tallón, M., Mateos, J., Babacan, S. D., Molina, R., & Katsaggelos, A. K. (2012). Space-variant blur deconvolution and denoising in the dual exposure problem. (pp. doi: 10.1016/j.inffus.2012.08.003).

Trucco, E., Ruggeri, A., Karnowski, T., Giancardo, L., Chaum, E., Hubschman, J. P., al Diri, B., Cheung, C. Y., Wong, D., Abràmoff, M., et al. (in press). Validating retinal fundus image analysis algorithms: Issues and a proposal. *Investigative Ophthalmology & Visual Science*.

Tutt, R., Bradley, A., Begley, C., & Thibos, L. N. (2000). Optical and visual impact of tear break-up in human eyes. *Investigative Ophthalmology & Visual Science*, 41(13), 4117–4123.

Valencia, E. & Millán, M. S. (2008). Color Image Sharpening and Application to Eye Fundus Image Analysis. In N. U. Wetter & J. Frejlich (Eds.), *6th Ibero-American Conference on Optics (RIAO) 9th Latin-American Meeting on Optics, Lasers and Applications (OPTILAS) American Institute of Physics Conference Series* (pp. 39–44).

Valencia, E., Millán, M. S., & Kotynski, R. (2006). Cup-To-Disc Ratio Of The Optic Disc By Image Analysis To Assist Diagnosis Of Glaucoma Risk And Evolution. In G. Cristóbal, B. Javidi, & S. Vallmitjana (Eds.), *5th International Workshop on Information Optics (WIO'06). AIP Conference Proceedings* (pp. 290–299).

Šroubek, F., Flusser, J., & Šorel, M. (2008). Superresolution and blind deconvolution of video. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (pp. 1–4).

Wallace, G. (1992). The JPEG still picture compression standard. *IEEE Trans on Consumer Electronics*, 38(1), xviii–xxxiv.

Wang, Z. & Bovik, A. (2006). *Modern image quality assessment*, volume 2 of *Synthesis Lectures on Image, Video, and Multimedia Processing*. Morgan & Claypool.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 13(4), 1–13.

Watson, A. (1994). Perceptual optimization of DCT color quantization matrices. *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, 1, 100–104 vol.1.

Whyte, O., Sivic, J., Zisserman, A., Ponce, J. C. V., & on, P. R. C. . I. C. (2010). Non-uniform deblurring for shaken images. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (pp. 491–498).

Wild, S., Roglic, G., Green, A., Sicree, R., & King, H. (2004). Global prevalence of diabetes estimates for the year 2000 and projections for 2030. *Diabetes care*, 27(5), 1047–1053.

Williams, R., Airey, M., Baxter, H., Forrester, J., Kennedy-Martin, T., & Girach, A. (2004). Epidemiology of diabetic retinopathy and macular oedema: a systematic review. *Eye*, 18(10), 963–983.

Winder, R., Morrow, P., McRitchie, I., Bailie, J., & Hart, P. (2009). Algorithms for digital image processing in diabetic retinopathy. *Computerized Medical Imaging and Graphics*, 33(8), 608–622.

Xiao, D., Frost, S., Vignarajan, J., Lock, J., Tay-Kearney, M.-L., & Kanagasingam, Y. (2012). Retinal image enhancement and registration for the evaluation of longitudinal changes. In *Adaptive Optical Systems Technology* (pp. 83152O–83152O–8).: SPIE.

Xu, J., Bao, J., Deng, J., Lu, F., & He, J. C. (2011). Dynamic Changes in Ocular Zernike Aberrations and Tear Menisci Measured with a Wavefront Sensor and an Anterior Segment OCT. *Investigative Ophthalmology & Visual Science*, 52(8), 6050–6056.

Xu, L. & Jia, J. (2010). Two-Phase Kernel Estimation for Robust Motion Deblurring. *Computer Vision–ECCV 2010*, (pp. 157–170).

Xu, L. & Luo, S. (2010). Optimal algorithm for automatic detection of microaneurysms based on receiver operating characteristic curve. *Journal of Biomedical Optics*, 15(6), 065004.

Yang, W., Wu, L., Fan, Y., & Wang, Z. (2008). A method of image quality assessment based on region of interest. *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, (pp. 6840–6843).

Zhu, X. & Milanfar, P. (2010). Automatic Parameter Selection for Denoising Algorithms Using a No-Reference Measure of Image Content. *Image Processing, IEEE Transactions on*, 19(12), 3116–3132.

Zitova, B. & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11), 977–1000.

Part III

Compilation of Publications

List of Publications

Articles

Refereed Journals Included in the JCR©

A. G. Marrugo, Michal Šorel, Filip Šroubek, and María S Millán, "Retinal image restoration by means of blind deconvolution", *Journal of Biomedical Optics*, **16**(11):116016, (2011).

A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, and H. C. Abril, "Anisotropy-based robust focus measure for non-mydriatic retinal imaging," *Journal of Biomedical Optics*, **17**(7):076021, (2012).

A. G. Marrugo, María S Millán, Michal Šorel, and Filip Šroubek, "Restoration of retinal images with space-variant blur", *Journal of Biomedical Optics*, submitted.

Technical Articles per Invitation

A. G Marrugo, María S Millán, Gabriel Cristóbal, Salvador Gabarda, Michal Šorel, and Filip Šroubek, "Toward computer-assisted diagnosis and telemedicine in ophthalmology", *SPIE Newsroom*, (2012) (doi: 10.1117/2.1201205.004256).

Refereed Journals not Included in the JCR©

A. G. Marrugo and M. S. Millán, "Optic Disc Segmentation in Retinal Images", *Optica Pura y Aplicada*, **43**(2), 79–86 (2010).

A. G. Marrugo and M. S. Millán, "Retinal image analysis: preprocessing and feature extraction" *Journal of Physics: Conference Series*, **274**(1), 012039, (2011).

Conference Proceedings

Invited Papers

A. G Marrugo, María S Millán, Gabriel Cristóbal, Salvador Gabarda, Michal Šrorel, and Filip Šroubek, "Image analysis in modern ophthalmology: from acquisition to computer assisted diagnosis and telemedicine", in *SPIE Photonics Europe, Proceedings SPIE*, **8436**:84360C, (2012).

M. S. Millán and **A. G. Marrugo**, "Image Analysis and Optics in Ophthalmology", *Lecture Notes of the Iternational Centre of Byocybernetics Seminar*, Polish Academy of Sciences, Warsaw, October, (2009).

International Conferences

A. G. Marrugo, Filip Šroubek, Michal Šorel, and María S Millán. "Multichannel blind deconvolution in eye fundus imaging", In *ISABEL '11-Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*, **7**:1–5. NY, USA, (2011).

A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, and H. C Abril, "No-reference Quality Metrics for Eye Fundus Imaging," in *CAIP'11: Proc. 14th Int. Conf. on Computer Analysis of Images and Patterns, Lecture Notes in Computer Science*, **6854**, 486–493, (2011).

National Conferences

A. G. Marrugo and M. S. Millán, "Efectos de compresión en imágenes de la retina para la evaluación del riesgo glaucomatoso", in *IX Reunión Nacional de Óptica*, pp. 140, Orense (Spain) (2009).

A. G. Marrugo, M. S. Millán, and H. C. Abril, "Implementation of an Image Based Focusing Algorithm for Retinal Imaging," presented at the *X Reunión Nacional de Óptica*, Zaragoza, 40–43, (2012).

Retinal Image Restoration by Means of Blind Deconvolution

Journal of Biomedical Optics, **16**(11):116016, (2011).

Journal of Biomedical Optics 16(11), 116016 (November 2011)

Retinal image restoration by means of blind deconvolution

Andrés G. Marrugo,[a] Michal Šorel,[b] Filip Šroubek,[b] and María S. Millán[a]
[a]Universitat Politècnica de Catalunya, Department of Optics and Optometry, Group of Applied Optics and Image Processing, Violinista Vellsolà 37, Terrassa, Barcelona 08222 Spain
[b]Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, Prague 8, 18208 Czech Republic

Abstract. Retinal imaging plays a key role in the diagnosis and management of ophthalmologic disorders, such as diabetic retinopathy, glaucoma, and age-related macular degeneration. Because of the acquisition process, retinal images often suffer from blurring and uneven illumination. This problem may seriously affect disease diagnosis and progression assessment. Here we present a method for color retinal image restoration by means of multichannel blind deconvolution. The method is applied to a pair of retinal images acquired within a lapse of time, ranging from several minutes to months. It consists of a series of preprocessing steps to adjust the images so they comply with the considered degradation model, followed by the estimation of the point-spread function and, ultimately, image deconvolution. The preprocessing is mainly composed of image registration, uneven illumination compensation, and segmentation of areas with structural changes. In addition, we have developed a procedure for the detection and visualization of structural changes. This enables the identification of subtle developments in the retina not caused by variation in illumination or blur. The method was tested on synthetic and real images. Encouraging experimental results show that the method is capable of significant restoration of degraded retinal images. © *2011 Society of Photo-Optical Instrumentation Engineers (SPIE)*. [DOI: 10.1117/1.3652709]

Keywords: blind deconvolution; image restoration; deblurring; retinal image.

Paper 11248RR received May 18, 2011; revised manuscript received Sep. 17, 2011; accepted for publication Sep. 22, 2011; published online Oct. 31, 2011.

1 Introduction

A fundus imaging device or retinal camera is a specialized low-power microscope with an attached camera designed to photograph the interior of the eye in association with the optical system of the eye. Retinal imaging is acknowledged to be an important tool for both detection and monitoring the progression of diseases affecting the eye, such as diabetic retinopathy, glaucoma, and age-related macular degeneration.[1] The digital format provides a permanent record of the appearance of the retina at any point in time.[2]

The imaging procedure is usually carried in two separate steps: Image acquisition and diagnostic interpretation. Image quality is subjectively evaluated by the person capturing the images, and they can sometimes mistakenly accept a low-quality image.[3] Low-quality image occurrence rate has been reported at 3.7–19.7% in clinical studies,[4–6] which is not a minor fact. A recent study by Abràmoff et al.[7] using an automated system for detection of diabetic retinopathy found that from 10,000 exams 23% had insufficient image quality. A major source of retinal image quality degradation are aberrations of the human eye, imperfections in the fundus camera optics, and improper camera adjustment, flash lighting, or focusing during the exam.[8] Moreover, regardless of how well controlled the aforementioned parameters are, in practice it may not always be possible to obtain good enough image quality as a result of additional factors such as lens opacities in the examined eye, scattering, insufficient pupil dilation or patient difficulty in steady fixating a target in the camera (such as in patients suffering from amblyopia).[3] Out of all possible retinal image degradations, some can be properly compensated via enhancement or restoration techniques (e.g., low-contrast, nonuniform illumination, noise, and blur).[2] However, this compensation is also dependent on the extent of the degradation. Regarding retinal image blurring, its main causes are relative camera-eye motion, inherent optical aberrations in the eye, and improper focusing.

In the past decade, many wavefront technologies—with its origins in astronomy—such as adaptive optics (AO)[9] and deconvolution from wavefront sensing (DWFS),[10] gave rise to the correction of monochromatic aberrations of the eye and also created new opportunities to image the retina at unprecedented spatial resolution. However, AO-corrected and DWFS-based fundus imagers usually aim at resolving details at the level of individual photoreceptors, thus have a field of view (FOV) of a couple degrees and a high resolution on the order of 1 or 2 μm.[11] Greater FOVs can be achieved ($\sim$5 deg)[12,13] with additional hardware constraints, beside the fact that diffraction limited imaging is not guaranteed due to an increase in aberrations.[14] Nevertheless, it is still a considerably narrow FOV and a major disadvantage with clinical subjects because of the need to examine larger areas of the retina. On the other hand, regular non-AO corrected fundus imagers used for routine checkups have a large FOV (typically, 30 deg) at the expense of lower spatial resolution, but still sufficient for practical detection and progression of observable

Address all correspondence to: Andrés G. Marrugo, Universitat Politècnica de Catalunya, Department of Optics and Optometry, Group of Applied Optics and Image Processing, Violinista Vellsolà 37, Terrassa, Barcelona 08222 Spain; Tel: 3493738678; E-mail: andres.marrugo@upc.edu.

1083-3668/2011/16(11)/116016/11/$25.00 © 2011 SPIE

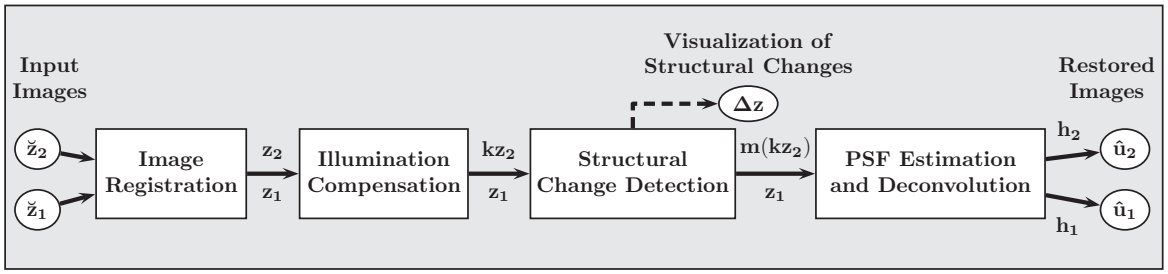


Fig. 1 Block diagram illustrating the proposed method. $\check{z}_i$ are the unregistered degraded input images, and $\hat{u}_i$ are their restored versions. The other variables are intermediate outputs of every stage; their meaning is given in the text.

clinical signs, such as microaneurysms, dot and blot hemorrhages, and exudates, among others. Consequently, large FOV fundus imagers are the major imaging modality available to patients visiting an eye-care clinic. The method proposed herein aims to restore images from conventional large FOV fundus imagers.

Among the normal retinal features, the blood vessel distribution exhibits a unique pattern in each individual and is highly stable in time. It is quite difficult to forge, and most common diseases do not change the pattern in a way that its topology is affected. For that reason, much effort has been put into the development of security systems based on the blood vessel distribution as a biometric signal for authentication purposes.[15] From this consideration, it is reasonable to assume the hypothesis that a pair of fundus images of the same retina, taken at different moments in time, contain enough common information to restore any of them by existing multichannel deconvolution techniques. We will demonstrate this fact later.

1.1 *Overview of Proposed Approach*

In this paper, we propose a new strategy for retinal image deblurring where we consider the most general image degradation case: blurred retinal images acquired in different moments in time, ranging from minutes to months; hence, disease progression is also considered. The main reason for this general image degradation case that considers long time lapses comes from the potential need to restore a degraded image acquired in the past being the only one available at that stage of the disease. This problem arises quite often in clinical practice. A correct assessment of a patient's state evolution requires sharp images from all moments in time; the method proposed here enables such opportunity. Disease progression characterization is embedded in the algorithm with the identification of areas of structural change (see Sec. 3.3).

Our restoration method is based on a technique called blind deconvolution (BD).[16,17] The goal of BD is to recover the original scene from a single image or a set of blurred images in the presence of a poorly determined or unknown point-spread function (PSF). The main assumption is that the blur can be described by a convolution of a sharp image with the unknown PSF. Restoration by deconvolution improves contrast and resolution of digital images (i.e., it is easier to resolve and distinguish features in the restored image). To avoid confusion with super-resolution, we briefly describe what we mean by resolution improvement. Digital deconvolution can be described as any scheme that sharpens up the PSF, while the spatial frequency bandwidth remains unchanged. This means that the spatial frequency response and the two-point resolution is improved, but the cutoff frequency is unchanged;[18] in the super-resolution context, the goal is to increase the cutoff frequency.

BD algorithms can be of single input [single-image blind deconvolution (SBD)] or of multiple images [multichannel blind deconvolution (MBD)]. Despite the fact that SBD is one of the most ill-posed problems, there are several reliable SBD algorithms,[19] although most of them require that the blurred image be governed by relatively strong edges, which is not case here. In Sec. 4.1 we compare our approach to a recent state-of-the-art SBD method.[20] The computational overhead from MBD (all of the preprocessing to adjust the time-sequence of images) in comparison to SBD is practically negligible, and the robustness of MBD is far superior and worth applying because SBD fails to produce a valid restoration. By the same token, the additional processing enables the identification of structural changes in the retina over time—a central task in medical practice. As a result, we have chosen a multichannel approach for the restoration of blurred retinal images.

An overview of the proposed approach is described in Fig. 1. We consider as input two-color retinal images acquired with a conventional fundus camera within a time lapse that can span from several minutes to months given by routine patient checkups. The images correspond to the same retina but can differ with respect to illumination distribution, blur, and local structural changes given by pathological developments. These differences cannot solely be accounted for by the convolutional model described in Sec. 2. For that reason, the images must be preprocessed before the blind deconvolution stage can take place. We register the images and compensate for inter-image illumination variation and structural changes. In fact, this preprocessing work becomes a great opportunity to meet one of the main concerns of ophthalmologists when they visually compare fundus images of the same retina over time: To identify true structural or morphological changes pertaining to possible pathological damage and, consequently, disregarding other changes merely caused by variation of illumination or blur. Ours is a two-stage blind deconvolution strategy. The first stage consists in the estimation of the PSFs following a multichannel scheme, and the second stage is the image deconvolution, where we restore every image with its corresponding PSF, independently. This has several advantages that will be explained in detail Sec. 3.5. The multichannel scheme is based on the method described in Ref. 21, which has proved to work well

in practice with sufficient experimental data. It is an alternating minimization scheme based on a maximum *a posteriori* (MAP) estimation, with *a priori* distribution of blurs derived from the multichannel framework and *a priori* distribution of the ideal sharp image defined by regularization with the total variation of the image.[22] MAP is formulated as an optimization problem, where regularization terms are directly related to priors. Regularization involves the introduction of additional information in order to solve an ill-posed problem in the form of a penalty or restriction in the minimization routine (see Sec. 3.4). This provides good quality of restoration—significantly better than, for example, Lucy–Richardson algorithm,[23] still widely used in biomedical applications. We have modified the algorithm in Ref. 21 to leave out regions where the eye fundus has structurally changed (it only takes into account one image in these regions) with the use of a masking operator, similarly to the solution proposed in Ref. 24 within the super-resolution context. This enabled us to restore both degraded input images.

In this work, our novel contributions to the retinal image processing task are twofold. First, we propose a degradation model for time-series retinal images, which captures the underlying distortions resulting from instrument limitations and changes between patient visits; we are also able to identify and highlight such changes. Second, we propose a restoration strategy based on blind deconvolution that is able to obtain image enhancement and resolution improvement using inexpensive digital methods applied to images acquired with a conventional fundus camera.

2 Mathematical Model of Image Degradation

The unregistered input images, as shown in Fig. 1, are $\check{z}_1$ and $\check{z}_2$. After registration, we obtain two degraded registered images z_1 and z_2, which we model as originating from an ideal sharp image u. Mathematically, the degradation model is stated as

$$z_1 = u * h_1 + n_1,$$
$$z_2 = (uk^{-1}) * h_2 + n_2 \quad , \tag{1}$$

where the asterisk is the standard convolution, h_i are called convolution kernels or PSFs, and k is a function accounting for relative local illumination change between images z_1 and z_2. For pixels where no illumination changes occur, $k \approx 1$. The noise n_i is assumed Gaussian additive with zero mean in both images. In our case, the PSFs and k comprise all radiometric degradations described above except structural changes in the eye, which is treated in Sec. 3.3. Despite the fact that we consider the PSFs to vary in time between the two image acquisitions, we assume them to be spatially invariant within each image. Because the FOV is of 30 deg or less, this assumption can be accepted in the first approach. This ideal sharp image u is actually unknown, and its estimation is the purpose of this paper. Thus to avoid confusion, the estimated (restored) image is denoted by $\hat{u}$. In Sec. 4.1, we test the performance of our method with synthetically degraded images, which means that we know u.

3 Description of the Method

In this section, we describe every stage of the proposed method. To illustrate each stage we use the images shown in Fig. 2. They were acquired using a nonmydriatic digital fundus camera system with conventional xenon flash lighting source (in the visible spectrum). The fundus images are from a patient that suffered from age-related macular degeneration and were captured within a seven-month time lapse. They are color RGB 24 bit-depth fundus images of size 1500 × 1200 digitized in TIFF format. This is a general example where both images do not correspond exactly to the same object field, the illumination distribution across both images is not exactly the same, and there are some structural differences between them given by the pathological development in the macula (centered yellowish region).

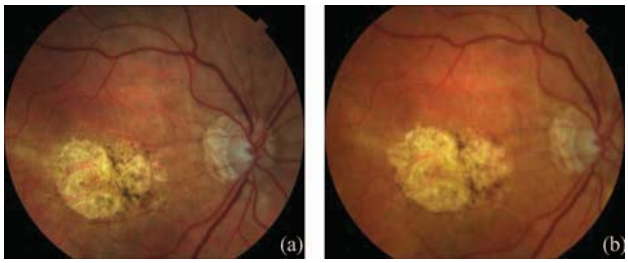


Fig. 2 Color fundus images of a human eye affected by age-related macular degeneration. Images (a) and (b) were captured within a seven-month time lapse, and (a) was captured before (b).

3.1 *Image Registration*

Image registration is a procedure that consists of spatial alignment of two or more images. General and application-specific image registration, such as in retinal imaging, has been investigated from the beginning of image-processing research. The interested reader is referred to the image registration review of Zitová and Flusser[25] and the recent work by Lee et al.[26] for objective validation of several retinal image registration algorithms. Image-registration techniques are usually divided into two groups: intensity-based and feature-based methods. Intensity-based methods have the drawback of poor performance under varying illumination conditions. Feature-based methods are robust to such effects but rely on accurate and repeatable extraction of the features. The retinal vasculature is known to provide a stable set of features for registration.

For registering the images, we use the robust dual-bootstrap iterative closest-point algorithm. We briefly describe it here; for a full description, of the method the reader is referred to Ref. 27. The vasculature from each image is automatically traced; starting from initial seed points extracted from a 1-D edge detection and, later, recursively tracking the vessels using directional templates. The vessel branching and crossover points are used as landmarks to register the images to subpixel accuracy. The registration algorithm starts from initial low-order estimates that are accurate only in small image regions called bootstrap regions. The transformation is then refined using constraints in the region, and the bootstrap region is expanded iteratively. The algorithm stops when the bootstrap region expands to cover the overlap between the images, and uses 12-dimensional quadratic mapping. This transformation model includes rotation, scale, translation, a shearing term, and a quadratic term that describes the spherical shape of the retina. We refer the interested reader to Ref. 28 for details on the model derivation. This registration algorithm is very robust to local changes and low overlap between

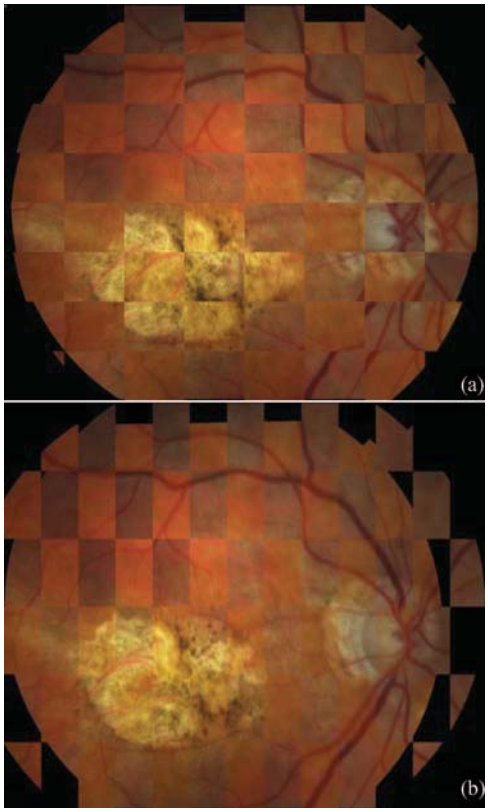


Fig. 3 Registration of images from Fig. 2 in checkerboard representation. (a) Before and (b) after registration.

images as demonstrated by its high success rate on test images with at least one common landmark point and overlaps even as low as 35%.[27] Even though the reported accuracy in Ref. 27 is of subpixel accuracy, in our case of degraded images this can be slightly worse without compromising the outcome. Minor local misregistration errors may occur when landmark points do not match precisely, but they will not be taken into account in the restoration because they will be masked out before the PSF estimation and image deconvolution stages (see Sec. 3.3).

To confirm the registration outcome, the pair of images before and after registration are shown in Fig. 3 in checkerboard representation, where the images are merged together in a chesslike pattern, where each square alternates information from one image to the other. Note how after registration the images have been correctly aligned, especially the blood vessel distribution.

3.2 *Compensation of Uneven Illumination*

Despite controlled conditions in retinal image acquisition, such as optical stops to prevent glares and provide a diffuse illumination, there are many patient-dependent aspects that are difficult to control and mainly affect the illumination component with gradual nonuniform spatial variations. Some of the contributing factors are (i) the curved surface of the retina (as a consequence, all regions cannot be illuminated uniformly); (ii) imaging requires either a naturally or an artificially dilated pupil (The degree of dilation is highly variable across patients); (iii) unexpected movements of the patient's eye; and (iv) presence of diseases. This nonuniform illumination across the image results in shading artifacts and vignetting. This effect hinders both quantitative image analysis and the reliable operation of subsequent global operators.

In our model, described by Eq. (1), the relative changes in intensity between the two fundus images cannot be described exclusively by convolution with different PSFs and must be compensated by k. A number of general-purpose techniques have been investigated to attenuate the variation of illumination. However, most techniques are oriented toward single-image compensation,[2] for instance, using the red channel to estimate background illumination.[29] Therefore, no consistency between two images is guaranteed. For our case, this uneven illumination can be compensated by properly adjusting the intensity values on one image to approximately match that of the other while satisfying a predetermined illumination model. This can be carried out if the blurring is not too large and the illumination changes smoothly, which is usually the case for fundus images. This assumption can be expressed mathematically as

$$(k^{-1} \cdot u) * h \approx k^{-1}(u * h).$$

The illumination of the fundus is formed by a slowly varying light field over a smooth surface, thus it can be modeled by a low-order parametric surface. In Ref. 30 they used a fourth-order polynomial to effectively model the light pattern formed by an illumination source passing through the attenuating ocular media. Here, we use a similar approach, but fitting the surface with respect to both images. The parametric surface fitting equation can then be formulated as

$$\arg\min_{k} \|z_1(x, y) - k(x, y)z_2(x, y)\|, \tag{2}$$

where $k(x, y) = \alpha_{15}y^4 + \alpha_{14}y^3x + \cdots + \alpha_2 y + \alpha_1$, and z_1, z_2 are the registered fundus images. We minimize Eq. (2) in the least-squares sense to estimate the 15 parameters. This procedure can be both carried out using the luminance channel or the green channel as usual in retinal image processing.[31] Here, we have used the green channel. Owing to the fact that the illumination can be compensated globally by the polynomial function k, it is important to realize that the structural changes remain unaffected. The interpretation of k from Eq. (2) is straightforward. If the registered images z_1 and z_2 had neither illumination changes nor structural changes, then $k \approx 1$ throughout the common object field. In Fig. 4, we show the resulting $k(x, y)$ for the images in Fig. 2. The different shades of gray indicate the average contrast and intensity difference between the two images. From the image, it can be seen that most areas have similar intensity values except for the upper left part (dark region).

3.3 *Segmentation of Areas with Structural Changes*

The pathological region is actually a structural change and cannot be taken as a variation of illumination. Image change analysis is of interest in various fields and many algorithms have been developed for change detection.[32,33] A survey of change detection methods can be found in Ref. 34. An initial step in order to identify these changes comes from computing the difference from the two registered images including the illumination compensation as

$$\Delta z(x, y) = z_1(x, y) - k(x, y)z_2(x, y) \ . \tag{3}$$

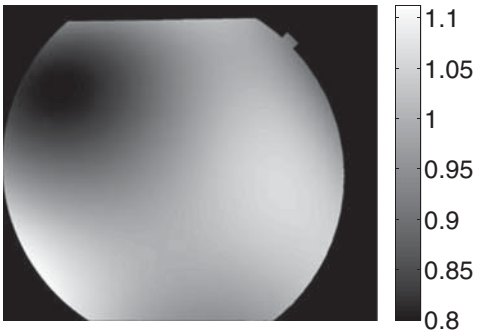


Fig. 4 Illumination compensation function $k(x, y)$.

The difference image is shown in absolute value Fig. 5(a). To better understand this result, in Fig. 5(b) we show one of the retinal images in gray scale, where the pixels related to structural changes are highlighted in pseudocolor. This image constitutes an important output of our algorithm. The structural changes can now be visualized and detected from the difference image $\Delta z(x, y)$ by taking a statistical significance test, in the same fashion as in Ref. 30. First, structural changes are often associated with a group of pixels; thus, the change decision at a given pixel j should be based on a small block of pixels in the neighborhood of j denoted as w_j. Second, in the absence of any change, the difference can be assumed to be due to noise alone. Therefore, the decision as to whether or not a change has occurred corresponds to choosing one of two competing hypothesis: the null hypothesis $\mathcal{H}_0$ or the alternative hypothesis $\mathcal{H}_1$, corresponding to no-change and change decisions, respectively. Assuming a Gaussian distribution for the difference values, the changes can be identified by comparing the normalized sum square of the differences within the neighborhood w_j to a predetermined threshold τ as described by Aach and Kaup.[32] The test is carried out as follows:

$$\Omega_j = \frac{1}{\sigma_n^2} \sum_{(x,y)\in w_j} \Delta z(x, y)^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau, \tag{4}$$

where σ_n is the noise standard deviation of the difference in the no-change regions. The threshold τ is derived from the fact that Ω_j follows a χ^2 distribution with N degrees of freedom, where N is the number of pixels in the window w_j. It can be obtained for a particular false-positive rate α from the χ^2 tables. The choice of an appropriate α is both guided by mathematical considerations (a 5% level for statistical significance is commonplace[35]) and the consequences that false alarms and misses might have. In this case, the effect of false alarms is unimportant because there would still be a large number of remaining pixels from where to compute the PSFs. On the other hand, misses do have a considerable impact in view of the fact that these pixels do not fulfill the convolutional model. As a result, α values of <0.05 might yield a more accurate change detection at the expense of possible undesirable misses. For all experiments, we use a 3×3 window ($N = 9$) and set $\alpha = 0.05$. The parameter σ_n was estimated by manually picking out no-change regions from a training set of images, computing Eq. (3) and the standard deviation inside these regions. Using Eq. (4) at each pixel, we can determine a change mask between the images or conversely a no-change mask. Given that, for the MBD procedure, we are interested in estimating the PSF from the no-change regions, the masking function m is obtained directly from the no-change mask of the significance test. The mask is shown in Fig 5(c). Note that the pathological region is the main cause of structural changes.

3.4 *Point-Spread Function Estimation*

In this section, we describe the basic principles of the blind deconvolution method used for the estimation of the PSFs. For this purpose, we have chosen one of the best working MBD methods.[21] MATLAB implementation of this method is available on the web of the authors.[36] The algorithm can be viewed as a Bayesian MAP estimation of the most probable sharp image and blur kernels. For our purposes, we used a modification of the original method that ignores regions affected by structural changes, which improves stability and precision of the computation. Without this modification, represented by the mask m in Eq. (5), the algorithm does not work reliably. The algorithm can be described as a minimization of the functional

$$\arg\min_{u,h_1,h_2} \left(\frac{1}{2}\|u * h_1 - z_1\|^2 + \frac{1}{2}\|m(u * h_2 - kz_2)\|^2 + \lambda_u \int |\nabla u| \, dx \, dy + \lambda_h \|m(z_1 * h_2 - kz_2 * h_1)\|^2 \right), \quad h_1, h_2 \geq 0, \tag{5}$$

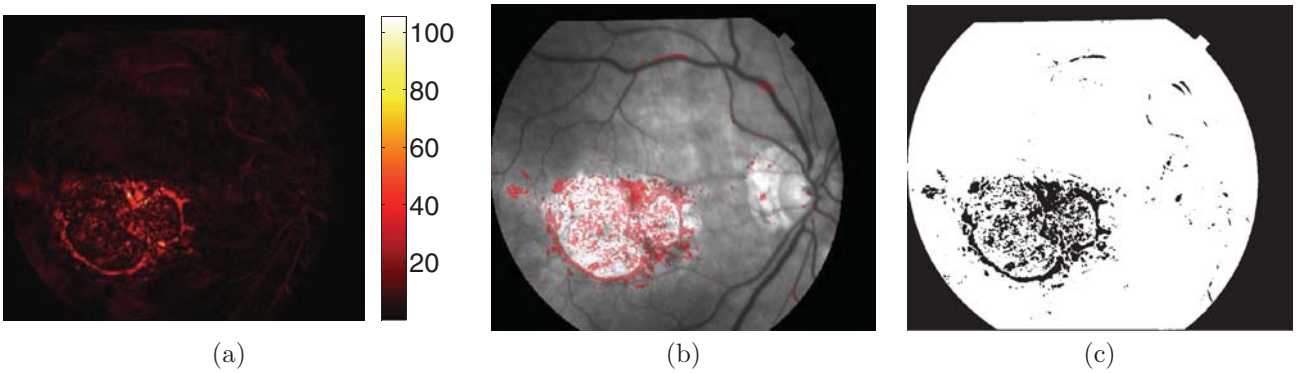


Fig. 5 Intermediate outputs from the algorithm: (a) image difference $\Delta z(x, y)$ in absolute value, (b) image difference in pseudocolor on top of gray-scale fundus image, and (c) mask m for avoiding areas with structural changes.

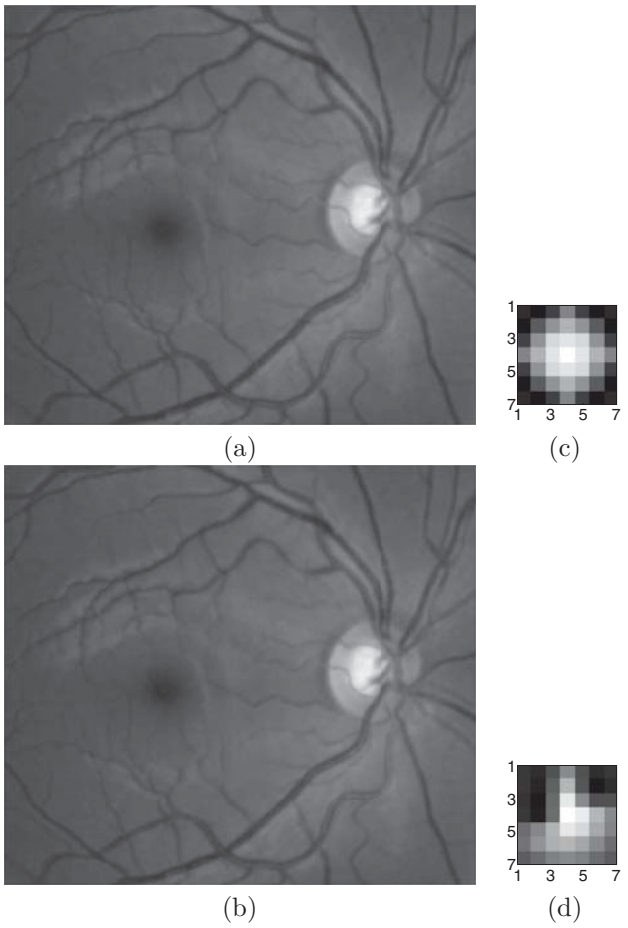


Fig. 6 (a, b) Degraded images (BSNR = 40 dB) and (c, d) PSFs.

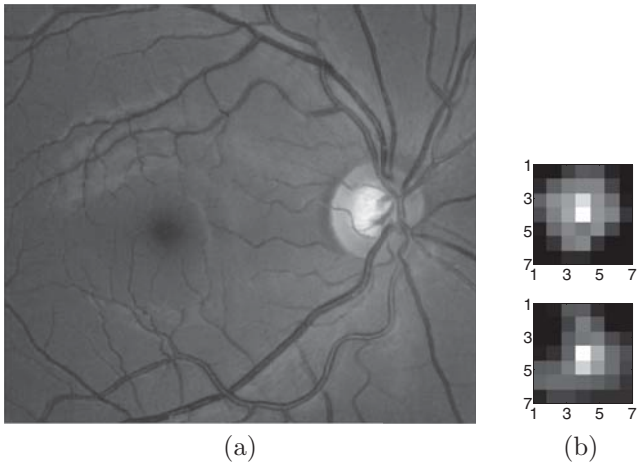


Fig. 7 (a) Restored image (ISNR = 4.45 dB) and (b) Estimated PSFs.

with respect to the latent image u and blur kernels h_1 and h_2. The first and second terms measure the difference between the input blurred images and the searched image u blurred by kernels h_1 and h_2. The size of this difference is measured by L_2 norm $\|.\|$ and should be small for the correct solution; ideally, it should correspond to the noise variance in the given image. Function k compensates for uneven illumination as described in Sec. 3.2. The value of the masking function m is 1 in the valid points [white in Fig. 5(c)] and 0 in the pixels where the eye fundus has structurally changed. Any of the first two terms could be masked, but not both at the same time. This is because the latent image u cannot have pixels with no value at all; hence, these pixels must take values from any of the two images. In this case, z_2 is masked. As a result, these pixels take values from the first term. The two remaining terms are regularization terms with positive weighting constants λ_u and λ_h. The third term is nothing else than the total variation of image u. It improves stability of the minimization and from the statistical viewpoint incorporates prior knowledge about the solution. The last term is a condition linking the PSFs h_1 and h_2 of both images, which also improves the numerical stability of the minimization.

The functional is alternately minimized in the subspaces corresponding to the image and the PSFs. The advantage of this scheme lies in its simplicity, this alternating minimization approach is actually a variation of the steepest-descent algorithm. The minimization in the PSF subspace is equivalent to the solution of a system of linear equations in the least-squares sense with the non-negativity constraint, in our implementation solved by the MATLAB fmincon function. The nonblind deconvolution realized by the minimization in the image subspace, is solved by half-quadratic iterative scheme,[37] replacing the total variation by $\int \sqrt{|\nabla u|^2 + \epsilon^2}$, where ϵ is an auxiliary variable in the

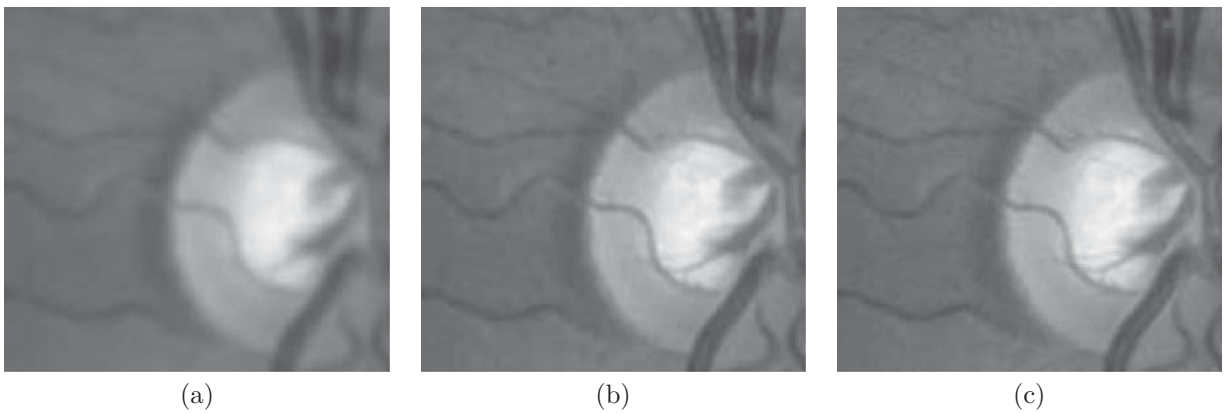
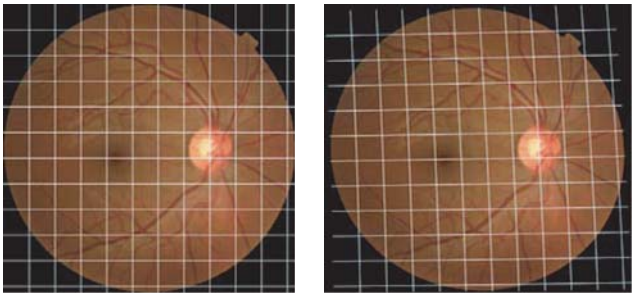


Fig. 8 Details from (a) degraded image, (b) restored image, and (c) original image.



(a) (b)

Fig. 9 (a) Original image and (b) geometrically distorted image.

range $0 < \epsilon \ll 1$. It is a small relaxation parameter that makes total variation differentiable around zero. A typical value for ϵ is 10^{-1}.

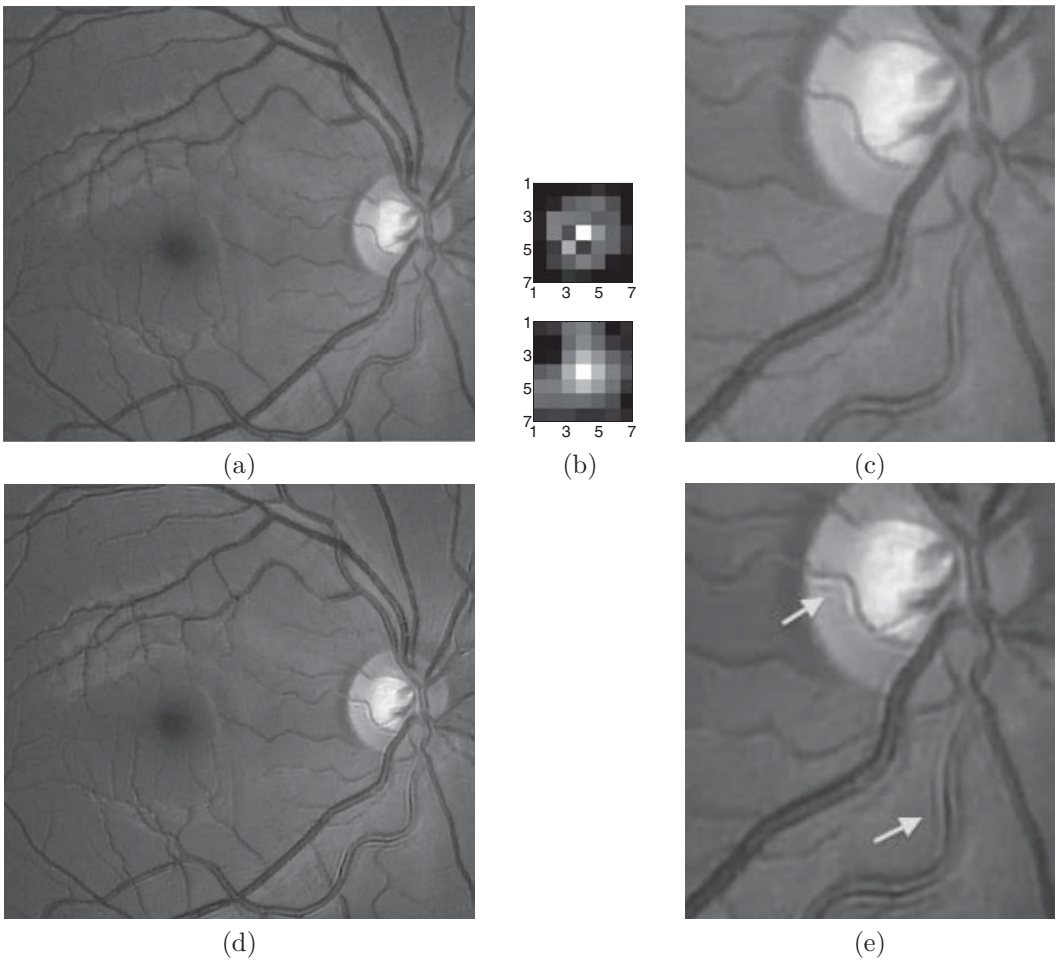
The main difference with respect to the original method[21] is the introduction of the masking function m, which is computed in the beginning of the algorithm as described in Sec. 3.3. During the minimization, the multiplication by m is included in the operator corresponding to the convolution with u (in the PSF minimization step) and in the operator corresponding to the convolution with h_2 (in the image minimization step). Because of the simplicity of this masking operation, the speed is practically the same as the speed of the original algorithm. In addition, even though we work with a complicated set of pixels, we can use the standard operation of convolution, which can eventually be speeded up using Fast Fourier transform (FFT).

3.5 *Image Restoration*

The aim of our algorithm is to restore both images as much as possible. Note that from Eq. (5) the restored version of z_1 ($\hat{u}_1$) is obtained because z_2 is masked; $\hat{u}_2$ could be obtained by minimizing Eq. (5) again with fixed PSFs and masking z_1. This procedure has the disadvantage that both images are restored only within the common object field. Therefore, an appropriate solution is to restore each image z_i via single-channel deconvolution with their corresponding PSF h_i (estimated from the previous step) by the minimization of the functional

$$\arg\min_{u_i} \left(\|u_i * h_i - z_i\|^2 + \lambda_u \int |\nabla u_i| \, \mathrm{d}x \, \mathrm{d}y \right). \tag{6}$$

This approach provides a further advantage in that the PSF estimation can be computed from a relatively small area of the common object field, provided that there are retinal structures



(a) (b) (c)

(d) (e)

Fig. 10 Image restoration from degraded and geometrically distorted images: (a) restored image by the proposed method (ISNR = 4.11 dB); (b) estimated PSFs; and (c) image detail, restored image by the method in Ref. 20 (ISNR = −0.72 dB); and (d) image detail.

within, thus greatly reducing the computational cost of the combined PSF estimation plus image deconvolution.

Finally, it should also be noted that the whole process of PSF estimation plus deconvolution can be computed for every channel of the RGB fundus image. However, in spite of the increase in computational burden, tests showed no real advantage to estimate the PSF for each channel. Moreover, the most suitable channel for PSF estimation is the green because it provides the best contrast. Whereas the blue channel encompasses the wavelengths most scattered and absorbed by the optical media of the eye; hence, the image has very low energy and a relatively high level of noise. As a result, the RGB deconvolved fundus image was computed by deconvolving every R, G, and B channel from the green channel PSF.

4 Experiments and Results

4.1 *Synthetic Images*

In this section, we use synthetically degraded retinal images to test the performance of the proposed method. We use blurred signal-to-noise ratio (BSNR) to measure the noise contained in the degraded image, and improvement in signal-to-noise ratio (ISNR) to measure the quality of restored images.[38] They are defined as follows:

$$\mathrm{BSNR} = 20\log_{10}\left(\frac{\|z\|}{\|n\|}\right),$$

$$\mathrm{ISNR} = 20\log_{10}\left(\frac{\|u-z\|}{\|u-\hat{u}\|}\right),$$

where u, z, $\hat{u}$, and n are the original image, degraded image, restored image, and noise vector, respectively. For ISNR, higher means better restoration; whereas for BSNR, lower means noisier degraded image. These metrics are mainly used to provide an objective standard for comparison to other techniques and they can only be used for simulated cases.

The first example is shown in Fig. 6, where the degraded images are synthesized from a sharp real image and the kernels shown in Fig. 6(c) and 6(d) plus Gaussian noise with zero mean and variance $\sigma^2 = 10^{-6}$ (BSNR=40 dB). The recovered image and PSFs are shown in Fig. 7. The restoration provides an ISNR=4.45 dB. In this case, for synthetically degraded images the masking operation of Sec. 3.3 was not applied. Visual inspection of the details shown in Fig. 8 clearly reveal the accuracy of the method. Under these circumstances, the algorithm is able to produce a significant restoration of fine details like small blood vessels around the optic disc.

To further test our approach under a more realistic degradation, we produced an initial geometrical distortion, via a quadratic model[26,28] as the one used for registration (Fig. 9). After the geometric distortion, the degradation (blur plus noise) is produced on both images (BSNR=40 dB). They are then registered, and the restored image is recovered via MBD. The restored image and the estimated PSFs are shown in Fig. 10. The ISNR is slightly less (4.11 dB) than in the previous case, but still sufficient to produce a significant restoration. To corroborate our assumption that MBD methods seem better suited for this type of images, we tried to restore the image with a recent SBD method proposed in Ref. 20. The result is shown in Fig. 10(e) and visually reveals that it does not follow the true nature of the blurring with artifacts around the blood vessels, thus being prone to produce a poor restoration evidenced by an ISNR= − 0.72 dB.

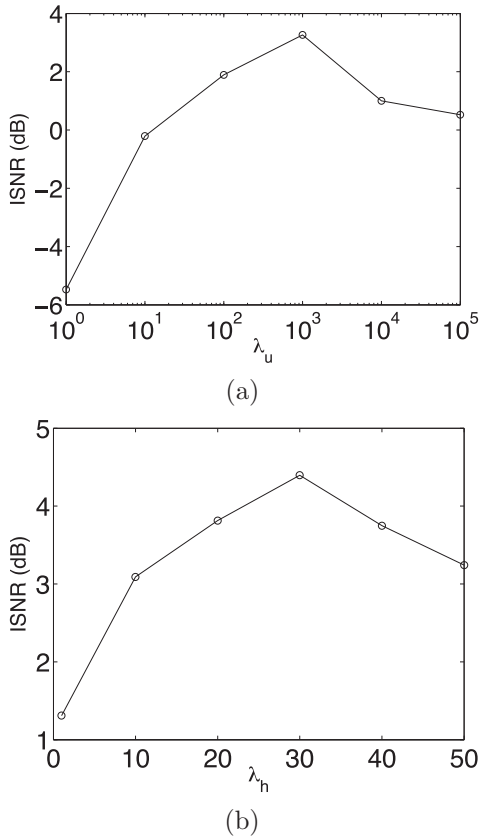


Fig. 11 Test on parameter setting (BSNR = 40 dB). Average ISNR with respect to different initial values of (a) λ_u and (b) λ_h.

Concerning parameter setting, in Fig. 11 we show the sensitivity of the two parameters λ_u and λ_h for the minimization of Eq. (5) in ISNR of the restored images. In Fig. 11(a), we fix the value of λ_h to 10 and check the ISNR of the restored images for different initial values of $\lambda_u = \{10^0, 10^1, 10^2, 10^3, 10^4, 10^5\}$. The best restoration is obtained with $\lambda_u = 10^3$; thus, in Fig. 11(b) we carried out the same procedure by fixing the value of λ_u to 10^3 and checking the ISNR of the restored image for different values of $\lambda_h = \{1, 10, 20, 30, 40, 50\}$. The best restoration was obtained with an initial value of $\lambda_h = 30$. For this type of image, when scaled to the interval $\langle 0, 1\rangle$, we find $20 < \lambda_h < 40$ to be a suitable range to produce an optimal restoration.

4.2 *Real Images*

The experiments shown in this section aim to demonstrate the applicability of the proposed method for retinal image deblurring in real scenarios. Three different cases are shown in Fig. 12, including the retinal images that were used to illustrate the method (Fig. 2). The estimated PSFs are shown at the bottom of the restored images. All images contain some pathological damage and have been acquired within considerable lapses of time (several months). In all examples, the resolution improvement can be visually assessed by the clear distinction of details, such as small blood vessels or the increase in sharpness

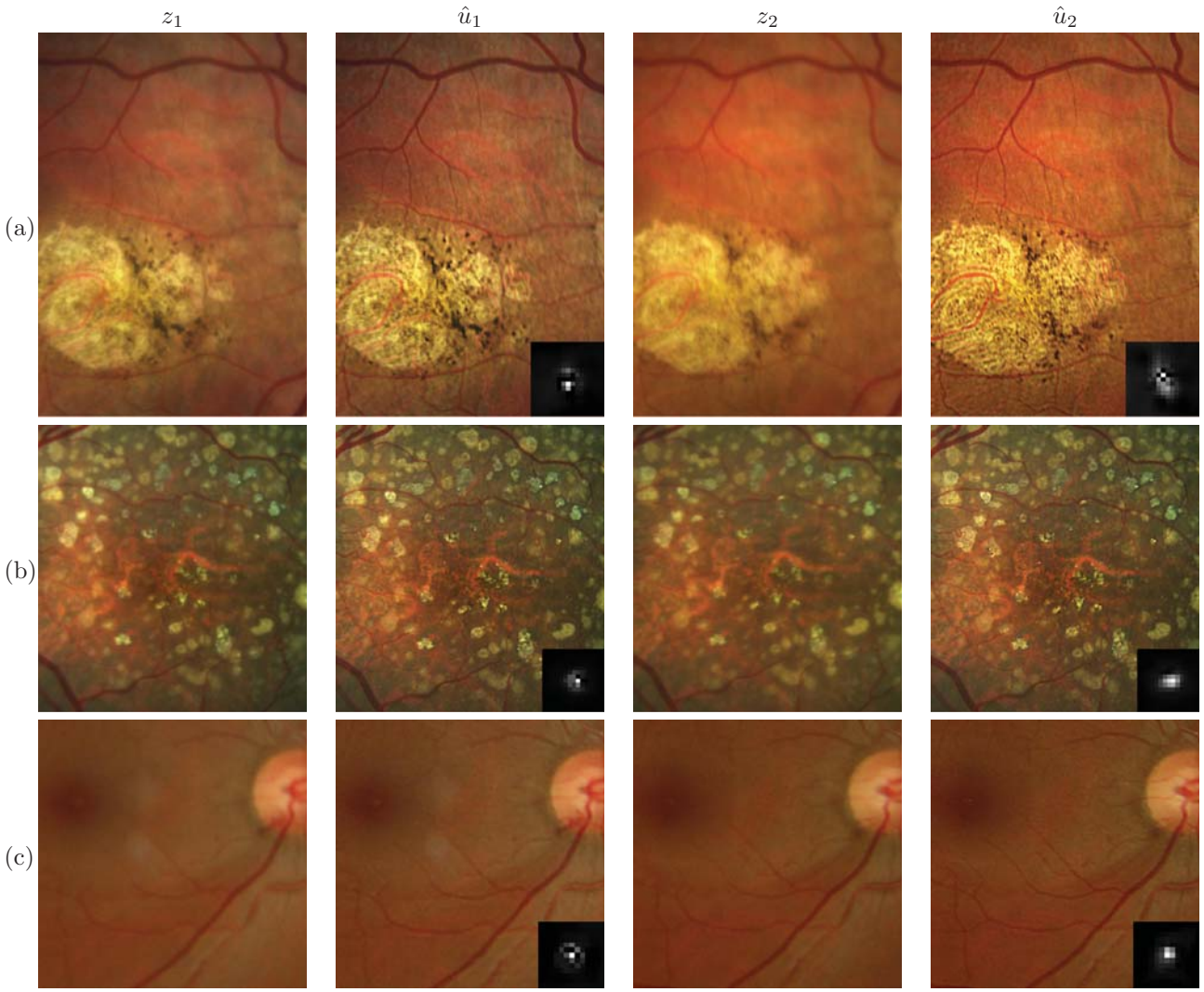


Fig. 12 Original and restored color retinal images; (a–c) indicate three separate cases arranged from left to right following our notation for degraded (z_i) and restored ($\hat{u}_i$) images. The images are cropped to represent the region of interest given by the pathological area. The estimated PSF is shown at the bottom of the restored image. Video files are also included for change detection in cases (a) and (b). (Video 1, Quicktime, 0.5 MB) [URL: http://dx.doi.org/10.1117/1.3652709.1]; (Video 2, Quicktime, 0.4 MB) [URL: http://dx.doi.org/10.1117/1.3652709.2]

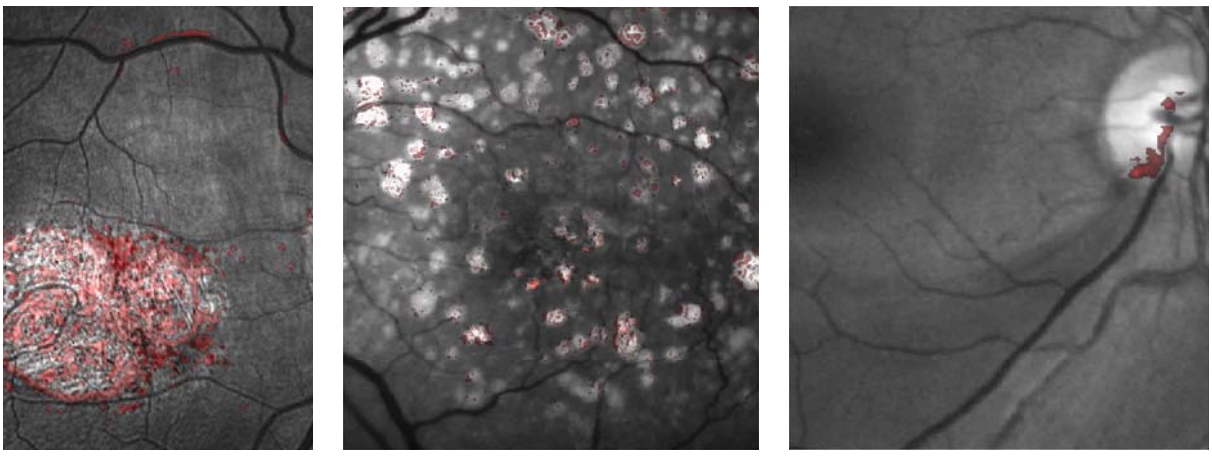


Fig. 13 Visualization of structural changes in pseudo-color for the images of Fig. 12.

of edges, especially in the pathological areas. We emphasize the fact that these images correspond to real routine patient follow-up and were not intentionally degraded. From a clinical viewpoint, the enhancement can be used for a more precise assessment of a patient's state. Likewise, the images are more suitable for subsequent processing such as for the detection of retinal pathology.[29,39]

In Fig. 13, the same images are shown but in gray scale to highlight the areas of structural change in pseudocolor. As mentioned earlier, this is an important result for its potential impact in medical practice. Subtle changes can be identified by this approach, such as the ones in Fig. 13(b) and the hemorrhage in the region of the optic disk in Fig. 13(c). Another technique to rapidly identify changes from the two images is by alternating both restored images in a video sequence. Videos 1 and 2 (Fig. 12) correspond to the first two real cases.

5 Conclusion

The main purpose of this paper has been to investigate a new approach for retinal image restoration based on multichannel blind deconvolution. In addition, we developed a strategy for identifying and highlighting areas of structural change with possible relation to pathological damage. We have verified that fundus images of the same retina over time contain enough common information to be restored with the proposed method. The method consists of a series of preprocessing steps to adjust the images so they comply with the convolutional model, followed by the final stages of PSF estimation and deconvolution. The synthetically degraded images enabled us to test the performance of the proposed approach and also to compare with a state-of-the-art single-channel blind deconvolution method. Results showed a remarkable enhancement evidenced by the increased visibility of details such as small blood vessels or pathological areas. The proposed method provides a novel practical approach for retinal image enhancement and, equally important the analysis of retinal changes over time. Central to the task of determining disease progression is the distinction of true change from variability.

The results of this study open several new avenues for research and applications. A possible application is found in the restoration of stereo retinal images for depth estimation. Most stereo images do not satisfy the brightness constancy assumption along with the expected blurring of some parts of the images because photographers find it difficult to focus two images simultaneously. Finally, research can also be conducted to compare to deconvolution from wavefront-sensing fundus imagers to determine if our method could be a suitable and inexpensive alternative.

Acknowledgments

This research has been partly funded by the Spanish Ministerio de Ciencia e Innovación y Fondos FEDER (Project No. DPI2009-08879). Financial support was also provided by the Czech Ministry of Education under the Project No. 1M0572 (Research Center DAR). The authors are also grateful to the ophthalmologist Jordi Monés, M. D., from the Institut de la Màcula i la Retina, Barcelona, for providing the images. The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

References

1. J. Kanski, *Diseases of the Ocular Fundus*, Elsevier/Mosby, New York (2005).
2. R. Winder, P. Morrow, I. McRitchie, J. Bailie, and P. Hart, "Algorithms for digital image processing in diabetic retinopathy," *Comput. Med. Imag. Graphics* **33**(8), 608–622 (2009).
3. H. Bartling, P. Wanger, and L. Martin, "Automated quality evaluation of digital fundus photographs," *Acta Ophthalmol.* **87**(6), 643–647 (2009).
4. A. Agrawal and M. A. McKibbin, "Technical failure in photographic screening for diabetic retinopathy," *Diabet. Med.* **20**(9), 777–780 (2003).
5. M. C. Boucher, J. A. Gresset, K. Angioi, and S. Olivier, "Effectiveness and safety of screening for diabetic retinopathy with two nonmydriatic digital images compared with the seven standard stereoscopic photographic fields," *Can. J. Ophthalmol.* **38**(7), 557–568 (2003).
6. H. M. Herbert, K. Jordan, and D. W. Flanagan, "Is screening with digital imaging using one retinal view adequate?" *Eye* **17**(4), 497–500 (2003).
7. M. D. Abràmoff, M. Niemeijer, M. S. A. Suttorp-Schulten, M. A. Viergever, S. R. Russell, and B. van Ginneken, "Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes," *Diabetes Care* **31**(2), 193–198 (2008).
8. A. V. Larichev, N. G. Irochnikov, P. Ivanov, and A. V. Kudryashov, "Deconvolution of color retinal images with wavefront sensing," *Proc. SPIE* **4251**, 102–116 (2001).
9. J. C. Christou, A. Roorda, and D. R. Williams, "Deconvolution of adaptive optics retinal images," *J. Opt. Soc. Am. A* **21**, 1393–1401 (2004).
10. J. Primot, G. Rousset, and J. C. Fontanella, "Deconvolution from wavefront sensing: a new technique for compensating turbulence-degraded images," *J. Opt. Soc. Am. A* **7**, 1598–1608 (1990).
11. D. Catlin and C. Dainty, "High-resolution imaging of the human retina with a Fourier deconvolution technique," *J. Opt. Soc. Am. A* **19**, 1515–23 (2002).
12. W. Wang, M. C.W. Campbell, M. L. Kisilak, and S. R. Boyd, "Optical design considerations when imaging the fundus with an adaptive optics correction," *Proc. SPIE* **7099**, 70990C (2008).
13. S. A. Burns, R. Tumbar, A. E. Elsner, D. Ferguson, and D. X. Hammer, "Large-field-of-view, modular, stabilized, adaptive-optics-based scanning laser ophthalmoscope," *J. Opt. Soc. Am. A* **24**, 1313–1326 (2007).
14. P. Bedggood, M. Daaboul, R. Ashman, G. Smith, and A. Metha, "Characteristics of the human isoplanatic patch and implications for adaptive optics retinal imaging," *J. Biomed. Opt.* **13**, 024008 (2008).
15. M. S. Millán, E. Pérez-Cabré, and B. Javidi, "Multifactor authentication reinforces optical security," *Opt. Lett.* **31**(6), 721–723 (2006).
16. D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.* **13**(3), 43–64 (1996).
17. E. P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*, CRC Press, Boca Raton (1996).
18. C. Sheppard, "Fundamentals of superresolution," *Micron* **38**(2), 165–169 (2007).
19. A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1964–1971 (2009).
20. L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," *Lect. Notes Comp. Sci.* **6311**, 157–170 (2010).
21. F. Šroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE Trans. Image Process.* **14**(7), 874–883 (2005).
22. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, 259–268 (1992).
23. W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.* **62**(1), 55–59 (1972).
24. F. Šroubek, J. Flusser, and M. Šorel, "Superresolution and blind deconvolution of video," in *Proc. of IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1–4 (2008).
25. B. Zitová and J. Flusser, "Image registration methods: a survey," *Image Vis. Comput.* **11**(21), 977–1000 (2003).

26. S. Lee, J. Reinhardt, P. Cattin, and M. Abramoff, "Objective and expert-independent validation of retinal image registration algorithms by a projective imaging distortion model," *Med. Image Anal.* **14**(4), 539–549 (2010).
27. C. Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE Trans. Med. Imaging* **22**(11), 1379–1394 (2003).
28. A. Can, C. Stewart, B. Roysam, and H. Tanenbaum, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Trans. Pattern Anal.* **24**(3), 347–364 (2002).
29. C. Muramatsu, Y. Hayashi, A. Sawada, Y. Hatanaka, T. Hara, T. Yamamoto, and H. Fujita, "Detection of retinal nerve fiber layer defects on retinal fundus images for early diagnosis of glaucoma," *J. Biomed. Opt.* **15**, 016021 (2010).
30. H. Narasimha-Iyer, A. Can, B. Roysam, C. Stewart, H. Tanenbaum, A. Majerovics, and H. Singh, "Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy," *IEEE Trans. Biomed. Eng.* **53**(6), 1084–1098 (2006).
31. M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med. Image Anal.* **9**(3), 179–190 (2005).
32. T. Aach and A. Kaup, "Bayesian algorithms for change detection in image sequences using markov random fields," *Signal Process. Image Commun.* **7**, 147–160 (1995).
33. C.-C. Chang, T.-L. Chia, and C.-K. Yang, "Modified temporal difference method for change detection," *Opt. Eng.* **44**, 027001 (2005).
34. R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Trans. Image Process.* **14**(3), 294–307 (2005).
35. S. Stigler, "Fisher and the 5% level," *Chance* **21**, 12 (2008).
36. MATLAB application for multichannel blind deconvolution, http://zoi.utia.cas.cz/download (2008).
37. A. Chambolle and P. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.* **76**(2), 167–188 (1997).
38. Liyakathunisa and V. K. Ananthashayana, "Super resolution blind reconstruction of low resolution images using wavelets based fusion," *World Acad. Sci. Eng. Technol.* **40**, 177–181 (2008).
39. L. Xu and S. Luo, "Optimal algorithm for automatic detection of microaneurysms based on receiver operating characteristic curve," *J. Biomed. Opt.* **15**, 065004 (2010).

Anisotropy-Based Robust Focus Measure for Non-Mydriatic Retinal Imaging

Journal of Biomedical Optics, **17**(7):076021, (2012).

Journal of Biomedical Optics 17(7), 076021 (July 2012)

Anisotropy-based robust focus measure for non-mydriatic retinal imaging

Andrés G. Marrugo,[a] María S. Millán,[a] Gabriel Cristóbal,[b] Salvador Gabarda,[b] and Héctor C. Abril[a]
[a]Universitat Politècnica de Catalunya, Group of Applied Optics and Image Processing, Department of Optics and Optometry, Violinista Vellsolà 37, 08222 Terrassa, Spain
[b]Instituto de Óptica, Consejo Superior de Investigaciones Científicas, Serrano 121, 28006 Madrid, Spain

Abstract. Non-mydriatic retinal imaging is an important tool for diagnosis and progression assessment of ophthalmic diseases. Because it does not require pharmacological dilation of the patient's pupil, it is essential for screening programs performed by non-medical personnel. A typical camera is equipped with a manual focusing mechanism to compensate for the refractive errors in the eye. However, manual focusing is error prone, especially when performed by inexperienced photographers. In this work, we propose a new and robust focus measure based on a calculation of image anisotropy which, in turn, is evaluated from the directional variance of the normalized discrete cosine transform. Simulation and experimental results demonstrate the effectiveness of the proposed focus measure.
 [DOI: 10.1117/1.JBO.17.7.076021]

Keywords: focus measure; eye fundus; retinal image; anisotropy; discrete cosine transform.

Paper 12148 received Mar. 2, 2012; revised manuscript received Jun. 19, 2012; accepted for publication Jun. 19, 2012; published online Jul. 17, 2012.

1 Introduction

Ocular fundus imaging has long played a key role in the documentation, diagnosis, and progression assessment of ophthalmic diseases. With the advent of digital cameras, ophthalmic imaging changed dramatically. Among the advantages of digital imaging are the ease and speed of access to data, fast and exact duplication, archiving and transmission, and digital image analysis techniques. Altogether, these advantages set the foundations for modern ophthalmology in the framework of telemedicine.

Fundus cameras can be mydriatic or non-mydriatic. Mydriatic fundus cameras require pharmacological dilation, while non-mydriatic cameras use a near-infra-red (NIR) viewing system to exploit the patient's natural dilation in a dark room.[1] Infra-red light is used to preview the retina on a video monitor. Once the monitor's image is focused and aligned, a flash of visible light from a Xenon arc lamp is fired and the image is captured. Non-mydriatic fundus cameras are equipped with a focusing mechanism that displaces a compensation lens. It is basically an aspheric objective lens design that, when combined with the optics of the eye, matches the image plane to the eye fundus. The focus control of the fundus camera is used to compensate for refractive errors in the subject's eye. Until recently,[2] these cameras were entirely operated manually with the focusing mechanism assisted by a split-line visual aid. Manual focusing is error prone, especially when performed by inexperienced photographers, and may lead to images that require additional restoration or enhancement.[3] The auto-focus feature offered in new retinal cameras is a significant advance that ultimately leads to a more robust imaging system, especially for medical screening purposes. On the other hand, the auto-focus feature still relies on the split-line mechanism, whereas in this work we propose a passive focus measure (FM) completely based on image analysis. For further details on fundus imaging, the reader is referred to Refs. 1, 4, and 5.

In Ref. 6 we studied the performance of several state-of-the-art no-reference image quality metrics for eye fundus imaging. The most interesting finding relates to the importance of directional properties with image quality. In other words, the measure of anisotropy as a quality metric. This was proposed by two co-authors of this paper (Gabarda and Cristóbal) in Ref. 7 and it represents an important step forward in the area of no-reference quality metrics. However, given the properties of the NIR fundus focusing system, the FM should be robust to noise (spatial and temporal) and to changes in illumination and contrast. Furthermore, real-time imaging is constrained by the overhead required to compute the directional Rényi entropy as described in Ref. 7. Therefore, in this work we made use of our findings on the directional dependency of eye fundus images against defocus to define a new and robust FM based on the directional variance of the normalized discrete cosine transform (DCT). The FM proposed here could impact the design of portable retinal cameras with autofocus function or the manufacturing of low-cost retinal cameras because they would require fewer optical components.

1.1 *Focusing*

In a single-lens optical imaging system operating within the paraxial regime, the process of focusing consists in adjusting the relative position of the object, the lens, the image sensor, or a certain combination of the three to obtain a focused image. Let $f(x,y)$ be the focused image of a planar object and $g_i(x,y)$ a sequence of images recorded for a sequence of camera parameter settings. The eye fundus is actually a curved surface; however, in our case $f(x,y)$ corresponds to a small region of the fundus so it can be considered as an isoplanatic

Address all correspondence to: Andrés G. Marrugo, Universitat Politècnica de Catalunya, Group of Applied Optics and Image Processing, Department of Optics and Optometry, Violinista Vellsolà 37, 08222 Terrassa, Spain. Tel: +34937398678; E-mail: andres.marrugo@upc.edu

0091-3286/2012/$25.00 © 2012 SPIE

patch.[8] We consider the variation of only one camera parameter at a time—either the lens position or the focal length. The acquired set of images can be expressed by convolution

$$g_i(x,y) = (f * h_i)(x,y), \qquad i = 1, \ldots, m, \tag{1}$$

where $h_i(x,y)$ is the point spread function (PSF) of the blur in the ith observation. In a practical imaging system, the image magnification and mean image brightness may change while focusing even if nothing has changed in the scene. Normalization with respect to these two parameters can be carried out. However, illumination normalization is more easily performed. Image magnification may be neglected because in most practical applications the magnification is less than 3%.[9] Ideally, the best possible case occurs when $h_i(x,y) = \delta(x,y)$, therefore $g_i(x,y) = f(x,y)$. In practice, all $h_i(x,y)$ have an unknown low-pass filter effect.

An FM may be understood as a functional defined on the image space which reflects the amount of blurring introduced by $h_i(x,y)$. Let S be the FM with which we look for the "best" image by maximizing/minimizing $S(g_i)$ over $i = 1, \ldots, m$. A reasonable FM should be monotonic with respect to blur and robust to noise. Groen et al.[10] used eight different criteria for the evaluation of focus functions. Ideally, the focus function should be unimodal, but in practice it can present various local maxima that can affect the convergence of the auto-focus procedure. Moreover, the focus curve should ideally be sharp at the top and long tailed, which can accelerate the convergence of the screening procedure.

2 Related Works

Various FMs have been reported in the literature.[2,9,11–13] They mainly consist of a focus-measuring operator that estimates the sharpness of the image. The image that yields a maximum FM is considered as the focused one. Almost all FMs depend directly on the amount of high-frequency information in the image. The high-frequency components correspond to edge information. On the other hand, their accuracy can deviate depending on the content of the processed images. Because well-focused images have sharper edges, they are expected to have higher-frequency content than blurred ones.[13] The common FMs are based on norm of gradient or second derivative of the image, gray level variance, and Laplacian energy. Surprisingly, little is known about the performance of these methods for fundus imaging, and the literature on this subject is scarce.

To the best of our knowledge, there exist only two published works that deal with autofocusing in retinal imaging;[2,14] however, they use conventional mydriatic imaging in the visible spectrum, which is not our case. In Ref. 14 they do not propose a FM; instead they use several preprocessing operations to improve the performance of traditional FMs for segmentation purposes. On the other hand, in the recent work by Moscaritolo et al.,[2] they propose a filtering technique to assess sharpness of optic nerve head images; however, they do not compare with other methods. In this section we briefly summarize five notable approaches—including that of Moscaritolo et al.[2]—for later comparison with our proposed method.

The first FM S_1 was proposed in Ref. 2. It may be defined mathematically as

$$S_1 = \mathrm{Var}\left(z_{\mathrm{med}} |g_i - z_{\mathrm{lp}}(g_i)| \right), \tag{2}$$

where z_{lp} is a low-pass filtering of $g_i(x,y)$, z_{med} is a nonlinear median filter of the absolute value $|.|$ of the difference for removing noise, and Var(.) is the variance. Another important measure is the l_2-norm of image gradient, also called the energy of gradient, defined as

$$S_2 = \sum_x \sum_y \left[\frac{\partial g_i(x,y)}{\partial x} \right]^2 + \left[\frac{\partial g_i(x,y)}{\partial y} \right]^2. \tag{3}$$

The third measure is the Laplacian energy. It can analyze high frequencies associated with image edges and is calculated as

$$S_3 = \sum_x \sum_y [\nabla^2 g_i(x,y)]^2. \tag{4}$$

Nayar and Nakagawa[15] proposed a noise-insensitive FM based on the summed modified Laplacian operators. When two second partial derivatives with respect to horizontal and vertical directions have different signs, one offsets the other and the evaluated focus value is incorrect. The method is a modification to obtain the absolute value of each second partial derivative as

$$S_4 = \sum_x \sum_y \left[\left| \partial^2 \frac{g_i(x,y)}{\partial x^2} \right| + \left| \partial^2 \frac{g_i(x,y)}{\partial y^2} \right| \right]. \tag{5}$$

The frequency-selective weighted median (FSWM) filter[16] is a high-pass nonlinear filter based on the difference of medians. It is well known as a nonlinear edge detector that removes impulsive noise effectively. The FSWM uses several nonlinear subfilters having a weight according to the frequency acting like a bandpass filter as

$$z_F(x) = \sum_j^P \beta_j \hat{z}_j(x), \tag{6}$$

where $z_F(x)$ is the FSWM filter, P is the number of subfilters, $\beta_j \in R$, and $\hat{z}_j(x)$ is the weighted median filter. The FM is produced by summing FSWM results, Fx and Fy, applied to an image along the horizontal and vertical directions as

$$S_5 = \sum_x \sum_y (F_x^2 + F_y^2). \tag{7}$$

Subbarao and Tyan[17] analyzed the robustness of three FMs: the image variance (not included here), S_2, and S_3. They recommended using S_3 because of its tolerance to additive noise; however, the differences among individual measures were not significant. There are many other FMs, such as the wavelet-based FM proposed in Ref. 12 or the mid-frequency discrete cosine FM in Ref. 11, but they were not included in our study either because of their lack of robustness to noise or their complex implementation. For a review and evaluation of FMs in natural images, the reader is referred to Refs. 9 and 13.

3 Discrete Cosine Transform

DCT is an invertible, linear transformation $\mathcal{T}: \mathbb{R}^N \to \mathbb{R}^N$. An image is transformed to its spectral representations by projection onto a set of of orthogonal two-dimensional (2-D) basis functions. The amplitudes of these projections are called the DCT

coefficients. Let $g(x,y)$, for $x = 0, 1, 2, \ldots, M-1$ and $y = 0, 1, 2, \ldots, N-1$, denote an $M \times N$ image and its DCT denoted by $\mathcal{T}[g(x,y)]$: $G(u,v)$, given by the equation

$$G(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x,y)\alpha(u)\alpha(v) \cos\left[\frac{(2x+1)u\pi}{2M}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right], \quad (8)$$

where

$$\alpha(\xi; \mathrm{A}) = \begin{cases} \sqrt{\frac{1}{A}} & \xi = 0, \\ \sqrt{\frac{2}{A}} & \text{otherwise}, \end{cases} \quad (9)$$

where $A = \{M, N\}$ depending on variables u and v, respectively. Low-order basis functions represent low spatial frequencies, while those of higher orders represent the higher spatial frequencies (Fig. 1). Therefore, low-order coefficients depict slow spatial variations in image intensity, while those of higher orders depict rapid variations.

The DCT is closely related to the discrete Fourier transform (DFT), a standard tool in signal processing, and has been reported as a suitable transform for spectral-based focusing algorithms.[18] However, the DCT has a greater energy compaction property than the DFT, i.e., most of the image information tends to be concentrated in a few low-frequency DCT coefficients. This is also why the JPEG compression standard is based on the DCT. In addition, many efficient schemes for the computation of DCT exist,[19] and hardware implementations are commonly available.[20]

3.1 *Normalized DCT*

The normalized DCT[21] of an image is defined as

$$\tilde{G}(u,v) = \tilde{\mathcal{T}}[g](u,v) = \frac{|\mathcal{T}[g](u,v)|}{\sum_{(u,v)} |\mathcal{T}[g](u,v)|}. \quad (10)$$

This normalization is important because it leads to invariance to changes in the contrast of the image. This can be shown with the following: let $g'(x,y) = cg(x,y)$, where c is a non-zero scaling factor. Given that the DCT is linear, the normalized DCT of g' is

$$\tilde{\mathcal{T}}[g'](u,v) = \frac{c|\mathcal{T}[g](u,v)|}{c\sum_{(u,v)} |\mathcal{T}[g](u,v)|} = \tilde{\mathcal{T}}[g](u,v), \quad (11)$$

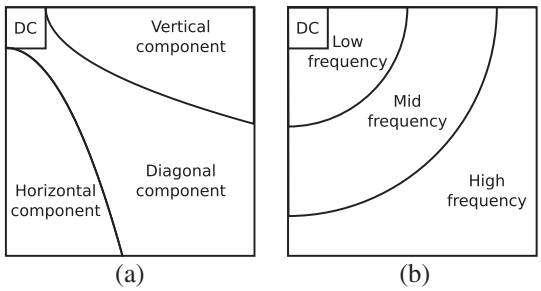


Fig. 1 Relationship between DCT coefficients and frequency components of an image.

which implies that the normalized DCT is contrast invariant and any measure based on this transform as well.

For illustrating the nature of blurring and the behavior of the DCT, we take the red channel from a sharp RGB fundus image (because it resembles more the NIR image) and simulate the imaging system as a linear shift-invariant system to acquire a sequence of images by varying the lens position. This was carried out by means of Fresnel propagation. In Fig. 2, we show the original sharp image, image patches of both the sharp and blurred images, and their DCT spectra (in the same log scale). Notice how the spectrum changes—there is less high- and mid-frequency content in the blurred image spectrum. In addition, in the original spectrum there are some favored orientations in the mid- and low-frequency coefficients, while in the blurred spectrum they seem to become more uniformly distributed. Another important feature is that in the blurred spectrum the coefficients related to high frequency have decreased significantly, and, as described in Sec. 2, many FMs are actually based on the idea of emphasizing high frequencies. While this may be true in theory, in practice there will always be noise contributing to the high-frequency content due to different acquisition conditions. Furthermore,

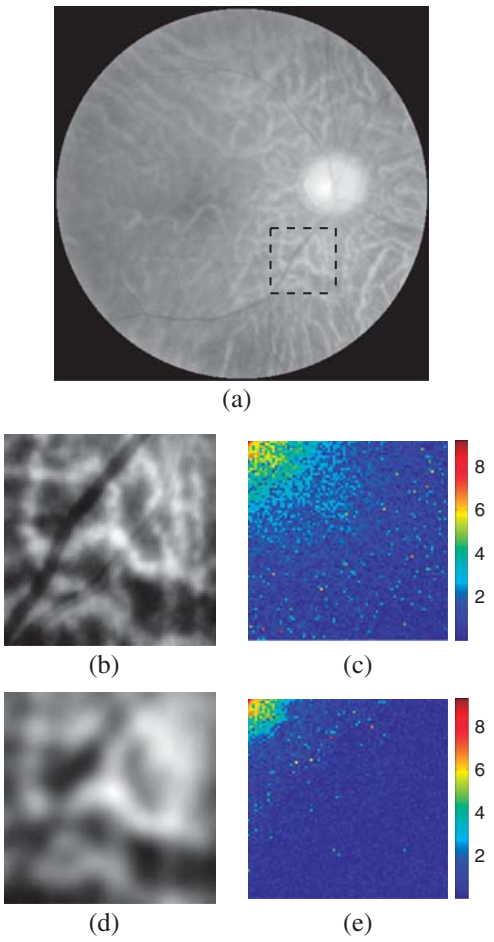


Fig. 2 (a), Original sharp fundus image (R channel from RGB fundus image). (b), ROI from sharp image, and (c), its DCT spectrum. (d), ROI from blurred image, and (e), its DCT spectrum. For visualization purposes both spectra are shown in log scale. Coefficients with higher values are shown in red and those with lower values are shown in blue. The blurred image spectrum is dominated by low-order coefficients.

given that the focusing mechanism involves acquiring a sequence of images, there will be spatial and temporal variations of noise.

4 Focus Measure

4.1 *Measure of Anisotropy*

As we have seen in the previous example, the overall nature of blurring can be described as a low-pass filtering that tends to break down the characteristic anisotropy of the original image. The FM proposed here aims to quantify this anisotropic dependence based on the normalized DCT of the image.

To define our measure, we introduce some notation. From Eq. (10), $\tilde{G}(u,v)$ is the normalized DCT of $g(x,y)$ of size $N \times N$, and λ_j, for $j = 1,2,3$, is a vector along one of the three main orientations of the spectrum depicted in Fig. 3. We will restrict our study to angular partitions of the spectrum roughly equivalent to vertical, diagonal, and horizontal components of the image space. Our measure of anisotropy mainly consists in calculating a difference of weighted coefficients along these orientations. Let $\tilde{G}_j = \{\tilde{G}(u,v) : \theta = \arctan(\frac{v}{u}), \theta_j \le \theta < \theta_{j+1}, j = 1,2,3\}$ be the set of DCT coefficients located between θ_j and θ_{j+1} angles, for $\theta_j \in \{0 \text{ deg}, 30 \text{ deg}, 60 \text{ deg}, 90 \text{ deg}\}$. The function $\psi_{\lambda_j}(.)$ takes as input $\tilde{G}_j$, performs orthogonal projection of all its elements along vector λ_j, and averages the elements that after projection fall on the same discrete (u,v) coordinates. With $\psi_{\lambda_j}(.)$ we seek to compact the information around the three main orientations in a one-dimensional vector of N elements. To illustrate, let us compute $\psi_{\lambda_1}(\tilde{G}_1) = [\psi^1_{\lambda_1}, \psi^2_{\lambda_1}, \ldots, \psi^N_{\lambda_1}]^T$, where $\tilde{G}_1$ is the set of coefficients located between $\theta_1 = 0$ deg and $\theta_2 = 30$ deg. In Fig. 3(b), we show the projection of the coefficient with coordinates (4,2) along λ_1. After projection, this coefficient has coordinates (4,1). Therefore, the element $\psi^4_{\lambda_1} = \text{mean}[\tilde{G}(4,1), \tilde{G}(4,2)]$. Consequently, we can stack all ψ_{λ_j} to form the following matrix,

$$\mathbf{\Psi} = \begin{bmatrix} \psi^1_{\lambda_1} & \psi^1_{\lambda_2} & \psi^1_{\lambda_3} \\ \psi^2_{\lambda_1} & \psi^2_{\lambda_2} & \psi^2_{\lambda_3} \\ \vdots & \vdots & \vdots \\ \psi^N_{\lambda_1} & \psi^N_{\lambda_2} & \psi^N_{\lambda_3} \end{bmatrix}.$$

Note that the first element of each vector corresponds to the dc coefficient. This coefficient does not convey any directional information of the image; however, we decided to keep it in the matrix for the sake of completeness. To obtain a measure of anisotropy—the FM itself—from $\mathbf{\Psi}$ we compute the variance of the weighted sum of the columns, computed as the matrix product $\mathbf{w\Psi}$,

$$S_a(g) = \text{Var}(\mathbf{w\Psi}) = \text{E}[(\mathbf{w\Psi} - \mu)^2], \tag{12}$$

where $\mathbf{w} = [w_1, w_2, \ldots, w_N]$, E is the expected value, and μ is the mean of the matrix product $\mathbf{w\Psi}$. The vector $\mathbf{w}$ can be regarded as a weighting procedure and with it we aim to achieve robustness to noise and illumination variation.

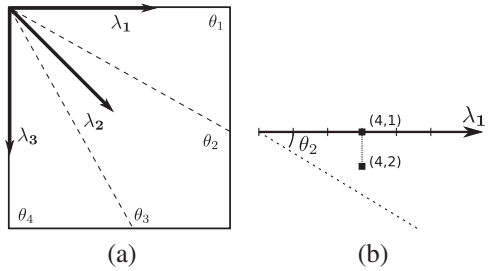


Fig. 3 (a) Vectors along the main directions of the DCT. (b) Projection of a coefficient along λ_1.

4.2 *DCT Coefficient Weighting*

The first issue to address is the selection of a suitable $\mathbf{w}$. In DCT-based pattern recognition, robustness is achieved by means of coefficient truncation.[22] It is known that low frequencies are related to illumination variation and smooth regions, and high frequencies represent noise as well as small variations (like edge and details) of the image. The middle-frequency coefficients contain useful information of basic structure; therefore these are suitable candidates for recognition.[23] Consequently, a trade-off between low-frequency and high-frequency truncation should be achieved to obtain a robust FM that is monotonic with respect to blur, unimodal, and at the same time robust to noise and illumination variations.

We decided to find a $\mathbf{w}$ that meets our requirements based on a training set of m images. This can be formulated as an optimization problem. The goal would be to find the vector $\mathbf{w} = [w_1, w_2, \ldots, w_N]$ that simultaneously optimizes K objective values $\{J_1(\mathbf{w}), J_2(\mathbf{w}), \ldots, J_K(\mathbf{w})\}$. Every objective value $J_k(\mathbf{w})$ is formulated so that the FM S_a decreases with respect to blur, $S_a(g^k_i) > S_a(g^k_{i+1}) \; \forall \, i = 1, \ldots, m$. There are K subsets of $g_i(x,y)$ all generated in the same way as described in Eq. (1), but they differ in that every k stands for a different kind of noise degradation, except for $k = 1$, the noise-free case. In other words, we want to find a $\mathbf{w}$ that guarantees monotonicity of S_a with respect to blur under different types of noise. The objective values are implicitly defined in terms of permutations of the ordered set $H = \{S_a(g_1), S_a(g_2), \ldots, S_a(g_m)\}$. Thus, the reference permutation is $\pi_r = \{1, 2, \ldots, m\}$, and any other arbitrary permutation of H violates the decreasing property of S_a with respect to blur. As a result, our goal is to find a $\mathbf{w}$ that produces permutations π_k for all K types of noise equal to that of π_r. The objective value is defined as the l_1-norm of the difference between π_r and π_k,

$$J_k(\mathbf{w}) : \sum_j^m |\pi_r(j) - \pi_k(j)|. \tag{13}$$

It is zero for two identical permutations, and approaches zero as π_k approaches π_r. This is the same for all $J_k(\mathbf{w})$; hence our single aggregate objective function[24] is the weighted linear sum of all $J_k(\mathbf{w})$, where all weights are equal to 1.

The solution to this problem is not a straightforward task, as the search space is multivariate and a unique global optimum cannot be guaranteed to exist. Therefore, we solved it using a probabilistic metaheuristic approach called simulated annealing.[25] It provides an acceptably good solution in a fixed amount of time. Each step of the algorithm replaces the current solution by a random nearby solution, chosen with a probability that depends both on the difference between the corresponding function values and on a global parameter T (called the temperature),

that is gradually decreased during the process. The dependency is such that the current solution changes almost randomly when T is large, but increasingly downhill as T goes to zero. (For further details see Ref. 24.)

4.3 *Implementation*

Typically, FMs are applied to a region called the focusing window, which is much smaller than the image. To achieve real-time computation, we decided to implement our measure by dividing the focusing window into subwindows. The measure is computed in the following manner:

1. The focusing window is divided into non-overlapping sub-images of size 16×16. This is chosen so that the most basic structures of the image fit in the subwindows.
2. Each subwindow image is transformed with the normalized DCT, and the FM S_a is computed.
3. An overall FM $\bar{S}_a$ is computed by taking the mean of all S_a values from the subwindows.

According to this implementation, the parameter $\mathbf{w}$ consists of 16 elements. The considered noise degradations for the procedure described in Sec. 4.2 are Gaussian noise ($\sigma^2 = 0.001$), speckle noise ($\sigma^2 = 0.001$), and impulsive noise ($d = 0.01$). The resulting $\mathbf{w}$ is shown in Fig. 4. As expected, the first two coefficients are practically zero. Observe the distribution of $\mathbf{w}$ instead of the individual values per coefficient. This means that a strong emphasis should be put to mid-frequency coefficients. It is perhaps not surprising that the distribution resembles a bandpass filter. This finding is consistent with the work in Ref. 9, where they showed that bandpass filtering causes the FMs to have sharp peaks while retaining monotonicity and unimodality. Interestingly, these weights also resemble the band pass response of the contrast sensitivity function of the human visual system. In the DCT domain, different approaches have been considered for computing visually optimized coefficients for a given image.[26] A major feature of our approach is the fast computation of the FM. The average execution time per frame, in MATLAB implementation on a PC with a 2.66-GHz Intel Core 2 Duo processor, is 40 ms. In most cases this is sufficient; however, if needed, implementation in a low-level programming language could significantly reduce the execution time. In addition, because we divide the focusing window into subwindows, our implementation could be further improved by taking advantage of large parallel architectures such as in graphics processor unit computing.

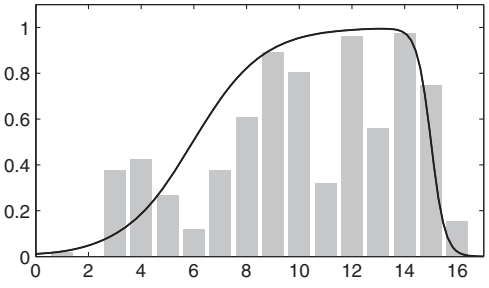


Fig. 4 DCT coefficient weights obtained from the optimization procedure. The distribution resembles a bandpass filter.

5 Results

5.1 *Simulated Images and Robustness Assessment*

To evaluate the robustness of our proposed FM S_a we simulated the focusing procedure. We generated a sequence $g_i(x, y)$ for $i = 1, \ldots, m$ from the red channel of a sharp RGB fundus image and propagated it at different distances through a linear imaging system of fixed focal length by means of Fresnel propagation. This is equivalent to displacing the lens or the sensor to look for the optimal focus position. From this noise-free sequence, we generated six additional sequences by corrupting it with two levels of three different types of noise: Gaussian, speckle, and impulse noise. We carried out this procedure for 20 retinal images for a total of 140 focusing sequences. Ideally, a noise-robust FM should produce the same (or similar) focusing curve for both the noise-free and corrupted sequences. To quantify the similarity between two focusing curves S_r and S_c we used the zero-lag normalized cross-correlation defined as

$$R(S_r, S_c) = \frac{\sum_i S_r(i) \cdot S_c(i)}{\sqrt{\sum_i S_r{}^2(i) \cdot \sum_i S_c{}^2(i)}}, \tag{14}$$

where r stands for the reference curve computed from the noise-free sequence and c the curve computed from the noise-corrupted sequence. The output is 1 in the case of perfect correlation and 0 for no correlation at all. The reason for the zero-lag calculation, as opposed to the regular cross-correlation by sliding dot product, is that we need the maxima of the curves to coincide in the horizontal position as well as the matching of the profiles.

All FMs were computed using a focusing window of 128×128 pixels located over retinal structures. In Fig. 5, we show an example to illustrate the robustness assessment of the FMs. The FM curves represent the normalized measure value over the search space for different lens positions. The highest value should be obtained when the lens is in the optimal focus position identified by the dashed vertical line. As the lens gets farther from the optimal position, the measure value should decrease proportionally to the distance. It comes as no surprise that all measures performed sufficiently well in the noise-free sequence shown in Fig. 5(a), where all curves follow a typical bell shape with a unique maximum. However, in the curves shown in Fig. 5(b)–5(d) where the focusing sequence is corrupted by different types of noise, the proposed FM S_a clearly outperforms the other measures in terms of monotonicity and unimodality. Notice that under Gaussian and speckle noise [Fig. 5(b)–5(c)], the S_a curves are nearly identical to the noise-free S_a curve in Fig. 5(a). Without jumping to conclusions, this result is interesting because it graphically shows the robustness of the proposed FM. The results for all of the 140 sequences are summarized in Table 1. Each value represents the average cross-correlation obtained for all 20 sequences corrupted with a specified type and level of noise for a particular FM. The overall average for each FM is shown in the last column. These results provide further evidence that the proposed FM S_a has a considerable robustness to noise, with an overall performance value of 0.929 and an exceptional 0.996 for the sequence corrupted with Gaussian noise with $\sigma^2 = 0.001$. The second and third best FMs were S_4 and S_1, with overall values of 0.781 and 0.502, respectively. In comparison with S_a these values represent a moderate

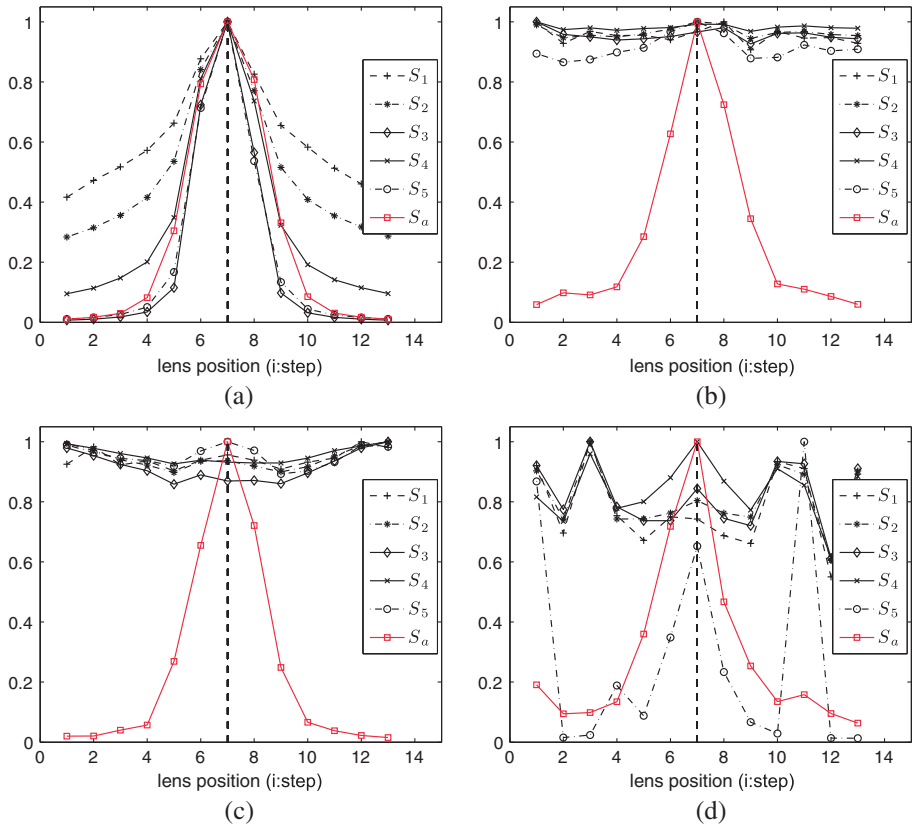


Fig. 5 Focus measures curves for the simulated images. The dashed vertical line indicates the correct focused position. (a) Noise-free images. (b) Gaussian noise ($\sigma^2 = 0.001$). (c) Speckle noise ($\sigma^2 = 0.001$). (d) Impulse noise ($d = 0.01$).

Table 1 Average normalized cross-correlation results for noise robustness assessment of focus measures from 140 sequences generated from 20 retinal images corrupted with different types and levels of noise. The three best FMs are in bold type.

	Gaussian ($\sigma^2 = 0.001$)	Gaussian ($\sigma^2 = 0.005$)	Speckle ($\sigma^2 = 0.001$)	Speckle ($\sigma^2 = 0.005$)	Impulse ($d = 0.01$)	Impulse ($d = 0.05$)	Overall average
S_1	0.554	0.486	0.635	0.422	0.477	0.438	**0.502**
S_2	0.524	0.499	0.468	0.408	0.476	0.462	0.473
S_3	0.449	0.444	0.370	0.359	0.420	0.417	0.410
S_4	0.784	0.782	0.750	0.746	0.836	0.791	**0.781**
S_5	0.495	0.380	0.495	0.304	0.795	0.362	0.472
S_a	0.996	0.939	0.997	0.992	0.979	0.667	**0.929**

to mild noise robustness. In the following section, we use these two FMs to compare with S_a in real images.

5.2 *Real Images*

In this subsection we show the results obtained from real NIR-focusing eye fundus images. The images have a relatively low signal to noise ratio (SNR) which justifies the need for a robust FM. All images were acquired using a digital fundus camera system (TRC-NW6S, Topcon, Tokyo, Japan) taking the video output from the infrared focusing system with a resolution of 640×480. The focusing system enables a compensation range of $-13D : 12D$ in normal operation. For strong myopia or hyperopia, two additional compensation lenses are available to compensate the ranges: $-12D : -33D$ and $+9D : +40D$, respectively. The image sequences analyzed here were acquired by means of an in-house assembled motor mechanism for the displacement of the compensation lens across the whole range for normal operation.

It is well known that as a person ages the crystalline lens of the eye gradually gets opacified, obstructing the passage of light. This is called a cataract. A complete loss of transparency is observed only in advanced stages in untreated patients. In early stages of cataracts retinal examination is considered

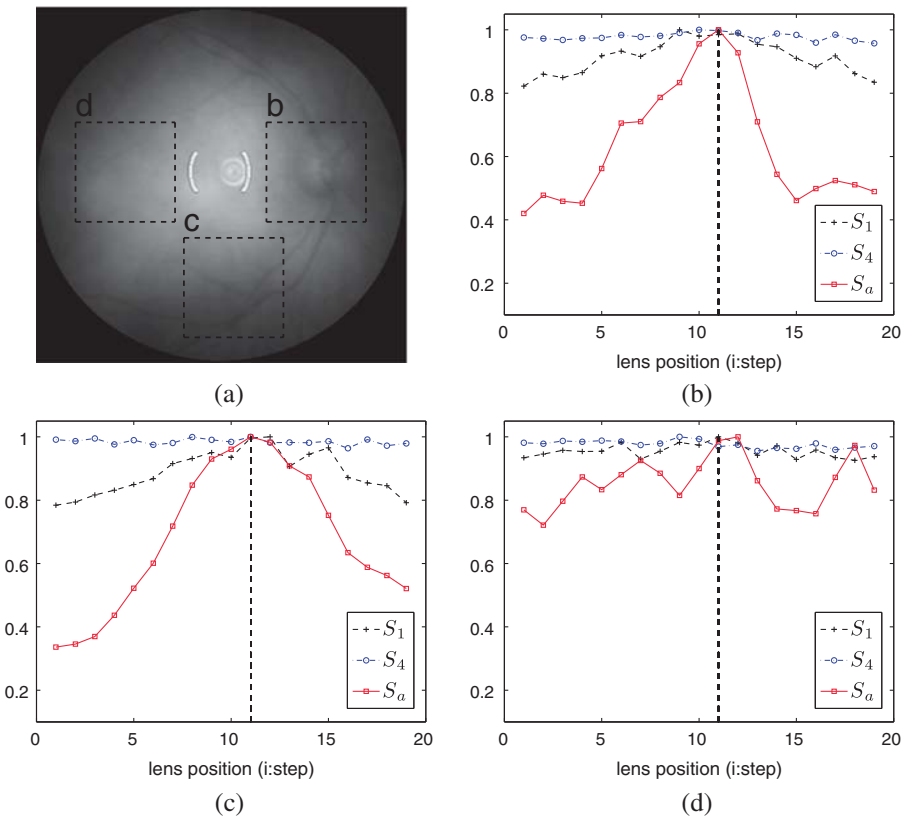


Fig. 6 Focus measure curves obtained by placing the focusing window over different regions of (a) the retinal image. The dashed vertical line indicates the correct focused position. Areas (b) and (c) are located over prominent retinal structures, whereas (d) is located over a relatively uniform region.

practicable—however, it is not without difficulty. For this reason, we decided to test our focusing method on healthy young subjects and elderly subjects with first signs of cataracts, not only to demonstrate its applicability on real images, but to assess its limitations as well. In this work we show results from five representative subjects of ages 27, 40, 68, 70, and 81 years for a total number of 10 eye fundi.

First we show the effects of placing the focusing window on different regions of the retinal image. A retinal image has distinct sharp structures such as the blood vessels and the optic disk, as opposed to the relatively uniform background. No FM is reliable without placing the focusing window on top of structures with edges, a fact easily appreciable from the three focusing curves shown in Fig. 6, which were computed from the right eye fundus of the 27-year-old subject. The optimal focus position identified by the dashed vertical line was verified via the split-line focusing mechanism. The S_a curves computed from regions (b) and (c) are clearly reliable in terms of monotonicity and unimodality and coincide on the optimal focus position. Conversely, the S_1 and S_4 curves fail to produce a reliable profile against the S_a curves that display a steeper peak at the optimal focus position, evidence of the measure's robustness to noise. In contrast, all measures computed from region (d) are unusable because they are mainly given by noise.

To illustrate the link between the focusing curves and the image quality, in Fig. 7 we show three image details depicting the optic disk region for three different focusing positions. The image detail in Fig. 7(b) corresponds to the focused image (optimal focus position 11 in the S_a curve Fig. 6). Notice how this image is properly focused: it has sharp details such as the blood vessels. The other two images are blurred, demonstrating the consistency of the S_a curves with image quality or sharpness. The result that emerges from this example is that to effectively locate the best-focused image, homogeneous regions should be avoided. An adaptive technique, based in an edge detector for example, could prove useful for detecting such prominent structures and therefore candidate regions for applying the focusing technique automatically. The focusing curves shown hereafter, however, were all computed from a focusing window located manually over retinal structures.

To further analyze the performance of the FM in Fig. 8, we show the focusing curves obtained from four of the five subjects; the ages are shown in the figure caption. From the four cases shown only in one [Fig. 8(c)], the S_a measure peak did not coincide precisely with the optimal focus position. However, the error is no more than a single position. The FMs curves of S_1 and S_4 are generally flatter than those of S_a which in a focus search strategy is not wanted because of the difficulty to properly distinguish the optimum position in a coarse or initial search. From the curves in Fig. 8, we can also note that there appears to be little difference between the curves from young

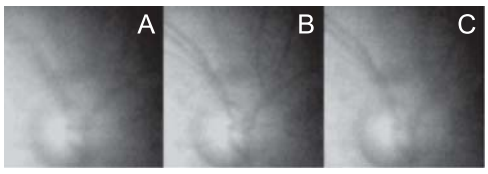


Fig. 7 Image detail from Fig. 6 for different focusing positions. (a) 6, (b), 11 (optimal focus), and (c), 15. The positions are in reference to Fig. 6(b)–6(c).

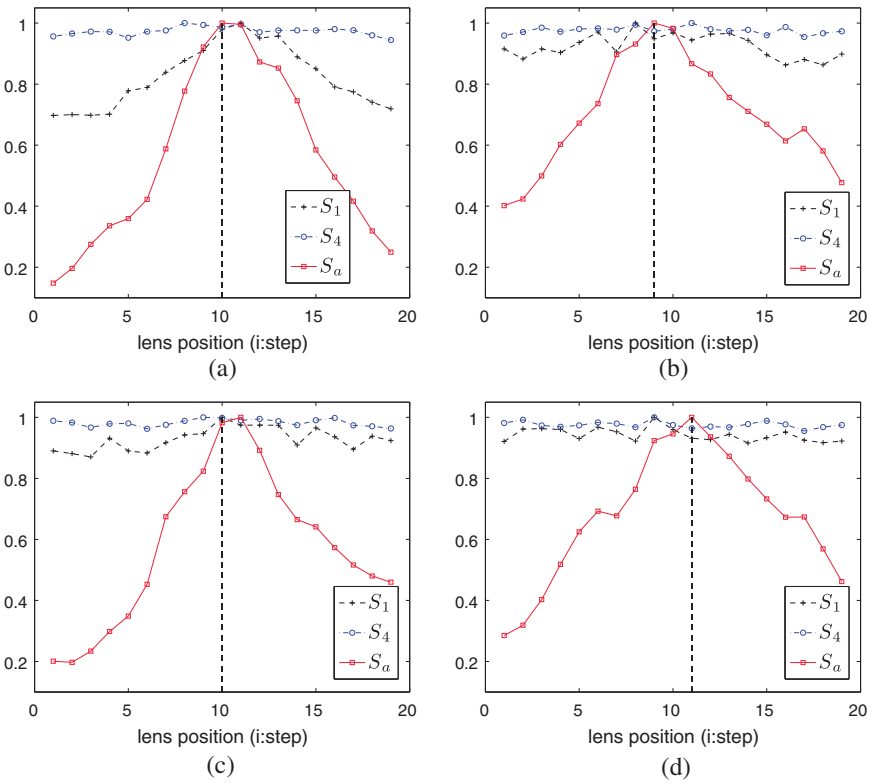


Fig. 8 Focusing curves obtained from four subjects with ages 27 (a) 40 (b) 68 (c), and 70 (d) years. The dashed vertical line indicates the correct focused position.

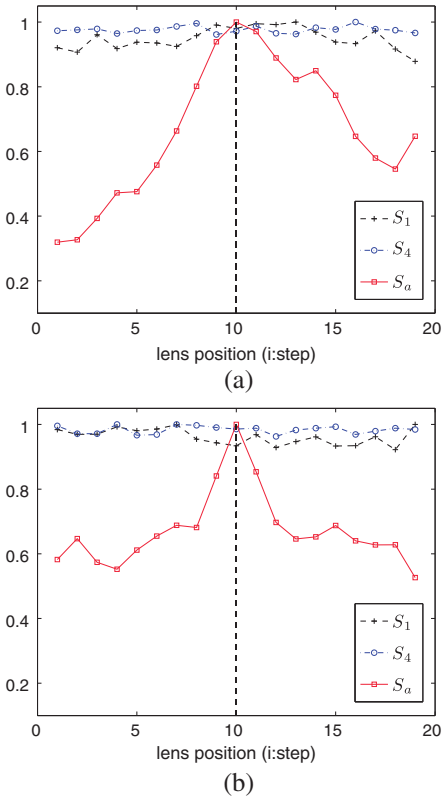


Fig. 9 Focusing curves obtained from the 81-year-old subject for each eye fundus. In the right eye (a), the crystalline lens has been extracted and replaced with an intraocular lens. The left eye (b), is in an early stage of cataract. The dashed vertical line indicates the correct focused position.

and elderly subjects. In Fig. 9, we show the focusing curves obtained from the 81-year-old subject for both eye fundi. This case is interesting on its own because in the right eye [Fig. 9(a)], the crystalline lens has been extracted and replaced with an intraocular lens, whereas the left eye [Fig. 9(b)], is in an early stage of cataract. While both focusing curves are able to successfully identify the optimal focus position, the curve in Fig. 9(b) is certainly flatter throughout most of the search space. This is most likely due to the difference in visibility and clarity from both eyes. In general, from the comparison against S_1 and S_4 it can clearly be stated that the proposed FM S_a outperforms them in the considered cases.

A close examination of the results reveal that the shape of the focusing curve is not exclusively given by the degree of defocus, but by the state of the subject's eye and the analyzed region of the fundus as well. This is important because it conditions the strategy for searching the optimal focus position. Finally, even though the results seem to indicate that the FM could be successfully applied to both young and elderly subjects, further research on a higher number and variety of subjects is necessary. Additionally, we report here that we encountered some difficulty in the procedure with the elderly subjects related to sustaining fixation during the acquisition procedure. From an initial number of six subjects one was excluded from all calculations due to this condition. Patient inability to successfully establish fixation is a true challenge in fundus photography, and dealing with it is out of the scope of this work.

6 Conclusion

In this paper, a new robust FM for nonmydriatic retinal imaging has been proposed. It is based on a measure of anisotropy, mainly the weighted directional variance of the normalized

DCT. The weights were calculated by means of an optimization procedure to maximize the noise robustness of the FM. Not only were the resulting weights in agreement with previous works,[9] but they also provide a key insight into the design of noise-invariant FMs. By both simulation and real fundus imaging, we demonstrated the robustness and the accuracy of the proposed FM, which clearly outperformed the other considered measures. The findings presented here may have a number of implications for the design and operation of autofocusing in modern retinal cameras. Finally, in this study we included several young and elderly subjects to assess the limitations of the proposed FM. Even though we found no significant differences between the focusing curves, there was some difficulty in the acquisition of images from the elderly mainly given by inability to sustain fixation. As with all such studies, there are limitations that offer opportunities for further research. Adapting our method to these variations within the patient population is a goal worth pursuing.

Acknowledgments

This research has been partly funded by the Spanish Ministerio de Economía y Competitividad and Fondos FEDER (project DPI2009-08879) and projects TEC2010-09834-E and TEC2010-20307. The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship. The authors especially thank the Pérez-Cabré and the Sisquella-Cabré families for their cooperation in the experimental session.

References

1. T. J. Bennett and C. J. Barry, "Ophthalmic imaging today: an ophthalmic photographer's viewpoint—a review," *Clin. Exp. Ophthalmol.* **37**(1), 2–13 (2009).
2. M. Moscaritolo et al., "An image based auto-focusing algorithm for digital fundus photography," *IEEE Trans. Med. Imag.* **28**(11), 1703–1707 (2009).
3. A. G. Marrugo et al., "Retinal image restoration by means of blind deconvolution," *J. Biomed. Opt.* **16**(11), 116016 (2011).
4. R. Bernardes, P. Serranho, and C. Lobo, "Digital ocular fundus imaging: a review," *Ophthalmologica* **226**(4), 161–181 (2011).
5. M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Rev. Biomed. Eng.* **3**, 169–208 (2010).
6. A. G. Marrugo et al., "No-reference quality metrics for eye fundus imaging," in *CAIP 2011, Lect. Notes Comput. Sci.* **6854**, 486–493 (2011).
7. S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy," *J. Opt. Soc. Am. A* **24**(12), B42–B51 (2007).
8. P. Bedggood et al., "Characteristics of the human isoplanatic patch and implications for adaptive optics retinal imaging," *J. Biomed. Opt.* **13**(2), 024008 (2008).
9. M. Subbarao, T. S. Choi, and A. Nikzad, "Focusing techniques," *Opt. Eng.* **32**(11), 2824–2836 (1993).
10. F. C. A. Groen, I. T. Young, and G. Ligthart, "A comparison of different focus functions for use in autofocus algorithms," *Cytometry* **6**(2), 81–91 (1985).
11. S. Y. Lee et al., "Enhanced autofocus algorithm using robust focus measure and fuzzy reasoning," *IEEE Trans. Circuits System. Video Technol.* **18**(9), 1237–1246 (2008).
12. J. Kautsky et al., "A new wavelet-based measure of image focus," *Pattern Recogn. Lett.* **23**(14), 1785–1794 (2002).
13. V. Aslantas and R. Kurban, "A comparison of criterion functions for fusion of multi-focus noisy images," *Opt. Commun.* **282**(16), 3231–3242 (2009).
14. P. Liatsis and P. Kantartzis, "Real-time colour segmentation and autofocus in retinal images," in *ELMAR, 47th International Symposium*, pp. 13–18, IEEE, Zadar, Croatia (2005).
15. S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(8), 824–831 (1994).
16. K. S. Choi, J. S. Lee, and S. J. Ko, "New autofocusing technique using the frequency selective weighted median filter for video cameras," *IEEE Trans. Consum. Electron.* **45**(3), 820–827 (1999).
17. M. Subbarao and J.-K. Tyan, "Selecting the optimal focus measure for autofocusing and depth-from-focus," *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(8), 864–870 (1998).
18. N. N. K. Chern, P. A. Neow, and M. H. Ang, "Practical issues in pixel-based autofocusing for machine vision," in *IEEE Int. Conf. Robotics Automation*, Vol. **3**, pp. 2791–2796, IEEE, Singapore (2001).
19. G. K. Wallace, "The jpeg still picture compression standard," *IEEE Trans. Consum. Electron.* **38**(1), 18–34 (1992).
20. J. Ramirez et al., "A new architecture to compute the discrete cosine transform using the quadratic residue number system," in *IEEE Int. Symp. on Circuits and Systems*, Vol. **5**, pp. 321–324, IEEE, Geneva, Switzerland (2000).
21. M. Kristan et al., "A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform," *Pattern Recogn. Lett.* **27**(13), 1431–1439 (2006).
22. Z. Lian and M. J. Er, "Illumination normalisation for face recognition in transformed domain," *Electron. Lett.* **46**(15), 1060–1061 (2010).
23. W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. Syst. Man Cybern. B Cybern.* **36**(2), 458–466 (2006).
24. A. Suppapitnarm et al., "A simulated annealing algorithm for multiobjective optimization," *Eng. Optimiz.* **33**(1), 59–85 (2000).
25. V. Granville, M. Krivanek, and J.-P. Rasson, "Simulated annealing: a proof of convergence," *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(6), 652–656 (1994).
26. A. B. Watson, "Perceptual optimization of DCT color quantization matrices," in *ICIP94, IEEE Int. Conf. Image Proc.*, Vol. **1**, pp. 100–104, IEEE, Austin, TX (1994).

Restoration of Retinal Images with Space-Variant Blur

Journal of Biomedical Optics, submitted.

Restoration of retinal images with space-variant blur

Andrés G. Marrugo,[1, a)] María S. Millán,[1, b)] Michal Šorel,[2, c)] and Filip Šroubek[2, d)]

[1)] *Group of Applied Optics and Image Processing*

Department of Optics and Optometry, Universitat Politècnica de Catalunya

Violinista Vellsolà 37, 08222 Terrassa, Spain.

[2)] *Institute of Information Theory and Automation*

Academy of Sciences of the Czech Republic

Pod Vodárenskou věží 4, 18208 Prague 8, Czech Republic.

(Dated: 10 January 2013)

Retinal images are essential clinical resources for the diagnosis of retinopathy and many other ocular diseases. Because of improper acquisition conditions or inherent optical aberrations in the eye, the images are often degraded with blur. In many common cases the blur varies across the field of view. Most image deblurring algorithms assume a space-invariant blur, which fails in the presence of space-variant blur. In this work we propose a novel strategy for the restoration of retinal images in which we consider the blur to be both unknown and space-variant. We model the blur by a linear operation interpreted as a convolution with a point-spread function (PSF) that changes with the position in the image. To achieve an artifact-free restoration, we propose a framework for a robust estimation of the space-variant PSF based on an eye-domain knowledge strategy. The restoration method was tested on a variety of real degraded retinal images. The results show an important enhancement, significant enough to leverage the images' clinical use.

Keywords: Blind deconvolution, space-variant restoration, image restoration, deblurring, retinal image.

[a)] andres.marrugo@upc.edu

[b)] millan@oo.upc.edu

[c)] sorel@utia.cas.cz

[d)] sroubekf@utia.cas.cz

Toward Computer-Assisted Diagnosis and Telemedicine in Ophthalmology

SPIE Newsroom, (2012) (doi: 10.1117/2.1201205.004256).

SPIE Newsroom

10.1117/2.1201205.004256

Toward computer-assisted diagnosis and telemedicine in ophthalmology

Andrés G. Marrugo, María S. Millán, Gabriel Cristóbal, Salvador Gabarda, Michal Šorel, and Filip Šroubek

Recent developments in retinal image analysis enable the possibility of advanced computerized medical systems.

Digital retinal imaging is an important element of modern ophthalmology. It provides visual documentation and, most importantly, the ability to extract and process information automatically. It is difficult to conceive modern health care practices without digital imaging and electronic health records. They both have led to a significant improvement in health care quality, but at the expense of lower physician productivity. In the words of Michael Abramoff,[1] a leading specialist in the field, "health-care automation has made physicians maybe do better, but not more." However, this seems to be merely the beginning of what is possible. A general lack of resources alongside ever-increasing health care costs is bound to stagnate this impetus for further improving the quality of health care, or most likely continue to leave an incredible amount of patients undiagnosed or untreated. In the long term, this translates into further increases in health care expenditures because, if a patient is left untreated, the cost of medical care in advanced stages of a disease increases dramatically. This situation could be avoided by investing in proper screening mechanisms.

It is within this overwhelming context that digital image analysis techniques can be employed to overcome most of the problems associated with eye disease screening, management, and progression assessment, among others. Computer-aided diagnosis (CAD) and telemedicine use has risen as the integration of different technological efforts is aimed at overcoming these difficulties. Retinal image analysis is a continuously growing research field with newly completed results being translated into clinical use. However, there are still many barriers to overcome before a definitive successful clinical deployment occurs. Retinal image illumination compensation, restoration, and change detection are only three of the many required techniques for retinal image analysis and CAD.

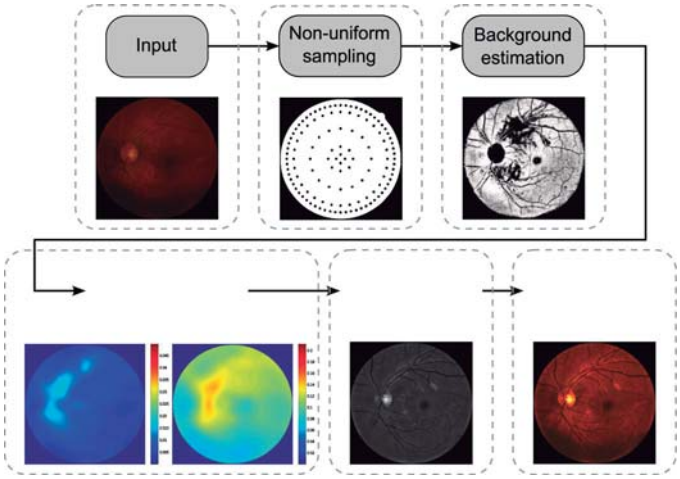


Figure 1. *Illumination compensation algorithm.*

Despite controlled conditions, many retinal images suffer from non-uniform illumination, mainly due to the curved retinal surface and the geometrical configuration of the light source and camera. Several approaches have been proposed to address this topic.[2] They can be loosely classified into single image or multiple image compensation techniques. An example of a single-image illumination compensation algorithm is depicted in Figure 1. Our main idea is that the original degraded image can be enhanced by estimating the background luminosity and contrast distribution in order to compensate for uneven illumination.[3] Thus, contrast and luminosity can be understood in terms of gain and offset. Our objective is to achieve illumination uniformity by performing the inverse operations on the image. The

Continued on next page

resulting image is significantly improved in terms of visibility, making it more suitable for visualization or for further processing tasks such as disease detection and segmentation.

One of the main concerns of ophthalmologists when they visually compare retinal images of the same patient over time is identifying true structural changes associated with possible pathological damage, while disregarding other changes merely caused by illumination variation or blur. To identify these changes,[4] we used a procedure that computed the pixel-by-pixel difference from the two registered retinal images with previous illumination compensation (see Figure 2c). From this difference image, the structural changes can be visualized and detected by taking a statistical significance test.[5] The two important aspects are first, that structural changes are mainly associated with a group of pixels so that the decision at a pixel is based on its neighborhood, and second, that in the absence of any change the difference is most likely given by noise. Figure 2d shows the structural changes detected from the images in Figures 2a and 2b, where most of the changes are located over the pathological region in the center.

Retinal images often suffer from blurring, which hinders their medical use. We have proposed a method for the restoration of retinal images by means of blind deconvolution (BD).[4] Using this method, we can recover the original scene from a single or set of blurred images in the presence of an unknown point spread function (PSF). We used a multi-image deconvolution algorithm,[6] where the PSF and the restored image are estimated in an alternation scheme and can be viewed as a Bayesian maximum, a posteriori estimation of the most probable sharp image and blur kernels.[4] We performed the restoration with a single image and with two images to demonstrate the advantages of using multiple images (see Figure 3). The intensity profiles from the restored images reveal a noticeable enhancement with more details properly resolved. The multi-channel deconvolution overcomes the limitations of single-channel deconvolution due to information redundancy. The improvement in resolution is evidenced by the gain in contrast and steeper slopes.

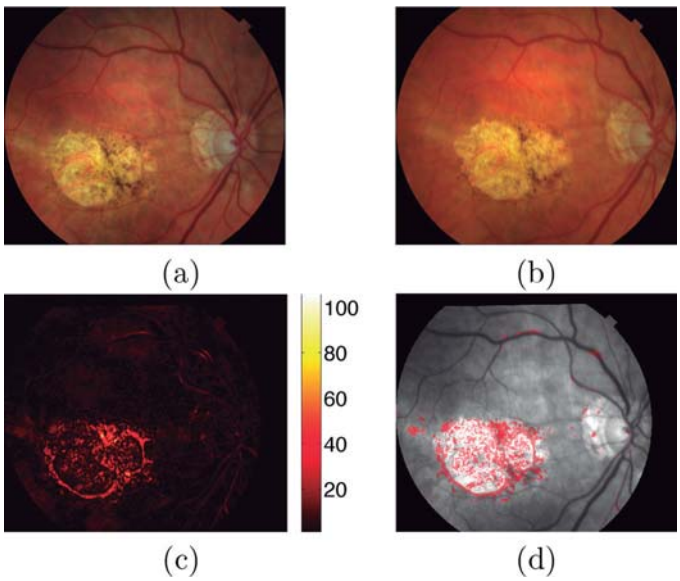


Figure 2. *Change detection. (a)–(b) Images from a patient acquired within a six-month lapse. (c) Image difference in absolute value. (d) Image change map in red on top of gray-scale retinal image.*

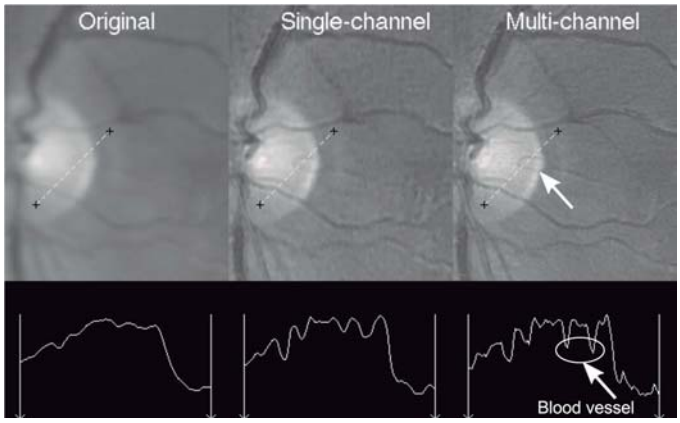


Figure 3. *Original degraded image, restored version using single-channel deconvolution, and multi-channel deconvolution (top, from left to right). Intensity profiles corresponding to the dashed segment (bottom).*

Digital retinal image analysis can be effective and cost-efficient for disease management, diagnosis, and screening. The increasing need for early detection and screening, along with the ever-increasing costs of health care, are likely to be the driving forces for the rapid adoption and translation of research findings into clinical practice. Future progress in this field must focus on solving data analysis problems, such as when the amount of data to be analyzed is excessively large and when the analysis is complex and requires quantification, as opposed to the more qualitative nature of the human expert.

Our next step is to continue developing techniques for retinal image analysis, especially for longitudinal change detection and image enhancement, to aid in early disease detection.

The authors are grateful for the support of the Spanish Ministry of Education, the Ministry of Science and Innovation, and the Czech

Continued on next page

Ministry of Education.

Author Information

Andrés G. Marrugo and María S. Millán
Applied Optics and Image Processing Group
Polytechnic University of Catalunya
Barcelona, Spain

María S. Millán is a full professor and optics chair whose research involves image processing, optical technologies, and applications. She is a SPIE Fellow and member of the European Optical Society.

Gabriel Cristóbal and Salvador Gabarda
Institute of Optics
Science Research Council
Madrid, Spain

Michal Šorel and Filip Šroubek
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Prague, Czech Republic

References

1. M. D. Abramoff, *Automated detection of retinal disease: When Moore's law meets Baumol's cost disease*, **SPIE Newsroom**, 2012. doi:10.1117/2.3201202.17
2. R. Winder, P. Morrow, I. McRitchie, J. Bailie, and P. Hart, *Algorithms for digital image processing in diabetic retinopathy*, **Comput. Med. Imag. Graph. 33** (8), pp. 608–622, 2009. doi:10.1016/j.compmedimag.2009.06.003
3. A. G. Marrugo and M. S. Millán, **Retinal image analysis: Preprocessing and feature extraction 274** (1), p. 012039, 2011. doi:10.1088/1742-6596/274/1/012039
4. A. G. Marrugo, F. Sroubek, M. Sorel, and M. S. Millán, *Retinal image restoration by means of blind deconvolution*, **J. Biomed. Opt. 16** (11), p. 116016, 2011. doi:10.1117/1.3652709
5. H. Narasimha-Iyer, A. Can, B. Roysam, C. Stewart, H. Tanenbaum, A. Majerovics, and H. Singh, *Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy*, **IEEE Trans. Biomed. Eng. 53** (6), pp. 1084–1098, 2006. doi:10.1109/TBME.2005.863971
6. F. Šroubek and J. Flusser, *Multichannel blind deconvolution of spatially misaligned images*, **IEEE Trans. Image Process. 14** (7), pp. 874–883, 2005. doi:10.1109/TIP.2005.849322

Optic Disc Segmentation in Retinal Images

Óptica Pura y Aplicada, **43**(2), 79–86 (2010).

Sección Especial: IX Reunión Nacional de Óptica

Optic disc segmentation in retinal images

Segmentación del disco óptico en imágenes de la retina

Andrés G. Marrugo[(*,S)] and María S. Millán[(S)]

Grup d'Òptica Aplicada i Processament d'Imatge, Departament d'Òptica i Optometria, Univesitat Politècnica de Catalunya, Terrassa.

[(*)] Email: andres.marrugo@upc.edu S: miembro de SEDOPTICA / SEDOPTICA member

Recibido / Received: 30/11/2009. Versión revisada / revised versión: 29/03/2010. Aceptado / Accepted: 30/03/2010

ABSTRACT:

Retinal images are widely used for diagnostic purposes by ophthalmologists. Therefore, these images are suitable for digital image analysis for their visual enhancement and pathological risk or damage detection. Here, we discuss two different approaches towards optic disc segmentation. We review and analyze a prior work based on a pixel-wise contour extraction for optic nerve head segmentation along with the effects of processing retinal images under lossy compression. This analysis revealed the need for a more general and robust approach. Finally, we introduce a different strategy for optic disc segmentation based on the use of active contours and color mathematical morphology.

Key words: Retinal Image, Optic Disc, Eye Fundus, Medical Imaging, Segmentation, Image Analysis.

RESUMEN:

Las imágenes de la retina son ampliamente utilizadas para el diagnóstico clínico por oftalmólogos. Por tanto, estas imágenes pueden ser procesadas digitalmente, tanto para su realce como para detección de riesgo patológico. En este trabajo se presentan dos vías para la segmentación del disco óptico. Partimos de un trabajo previo sobre la extracción del contorno del disco óptico, y analizamos los efectos producidos por un procesado de la imagen retiniana bajo compresión con pérdidas. De este análisis resulta evidente la necesidad de un algoritmo con un enfoque más general y robusto. Por último, presentamos una estrategia diferente para la segmentación del disco óptico basado en el uso de contornos activos y morfología matemática en el espacio de color.

Palabras clave: Imagen Retiniana, Fondo de Ojo, Disco Óptico, Imagen Médica, Segmentación, Análisis de Imagen.

REFERENCES AND LINKS

[1] N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. Dhillon, R. H. Eikelboom, K. Yogesana, I. J. Constable, "Retinal image analysis: Concepts, applications and potential", *Prog. Retin. Eye Res.* **25**, 99–127 (2006).

[2] N. Harizman, C. Oliveira, A. Chiang, C. Tello, M. Marmor, R. Ritch, J.M. Liebmann, "The ISNT rule and differentiation of normal from glaucomatous eyes", *Arch. Ophthalmol*. **124**, 1579-1583 (2006).

[3] T. Walter, J. C. Klein, P. Massinm, A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy – detection of exudates in color fundus images of the human retina", *IEEE T. Med. Imaging* **21**, 1236-1243 (2002).

[4] H. Li, O. Chutatape, "Boundary detection of optic disk by a modified ASM method", *Pattern Recogn.* **36**, 2093-2104 (2003).

[5] E. Valencia, M. S. Millán, R. Kotynski, "Cup-to-disc ratio of the optic disc by image analysis to assist diagnosis of glaucoma risk and evolution," in *Information Optics: 5th International Workshop*, G. Cristóbal, B. Javidi, S.Vallmitjana Edts., American Institute of Physics, CP860, 290-299 (2006).

[6] E. Valencia, M. S. Millán, "Color image sharpening and application to eye fundus image analysis," RIAO/OPTILAS 2007, edited by N.U. Wetter, J. Freijlich, American Institute of Physics CP992, 39-44 (2008).

[7] A. G. Marrugo, M. S. Millán, "Efectos de compresión en imágenes de la retina para la evaluación del riesgo glaucomatoso", IX Reunión Nacional de Óptica, Sociedad Española de Óptica, Orense, 140 (2009).

[8] F. Mendels, C. Heneghan, P.D. Harper, R.B. Reilly, J. P. Thiran, "Extraction of the optic disk boundary in digital fundus images", Proc. First Joint BMES/EMBS Conference, **2**, 1139 (1999).

[9] A. Osareh, M. Mirmehdi, B. Thomas, R. Markham, "Comparison of colour spaces for optic disc localisation in retinal images", 16th Int. Conf. on Pattern Recognition **1**, 743–746 (2002).

[10] H. Zhou, G. Schaefer, T. Liu, F. Lin, "Segmentation of optic disc in retinal images using an improved gradient vector flow algorithm", *Multimed. Tools Appl.* (2010) Online: doi 10.1007/s11042-009-0443-0

[11] M.S. Millán, E. Valencia, "Color image sharpening inspired by human vision models", *Appl. Opt.* **45**, 7684–7697 (2006).

[12] G. M. Johnson, M.D. Fairchild, "A top down description of the S-CIELAB and CIEDE2000", *Color Res. Appl*. **28**, 425–435 (2003).

[13] X. Zhang, B. A. Wandell, "A spatial extension to CIELAB for digital color image reproduction", *Soc. Fo Info. Disp. Symp. Tech. Digest* **27**, 731–734 (1996).

[14] T. F. Chan, L. A. Vese, "Active contours without edges", *IEEE T. Image Process.* **10**, 266–277 (2001).

[15] J. Conrath, "Evaluation of the effect of JPEG and JPEG2000 image compression on the detection of diabetic retinopathy", *Eye* **21**, 487-493 (2006).

[16] R. J. Winder, P. J. Morrow, I. N. McRitchie, J. R. Bailie, P. M. Hart, "Algorithms for digital image processing in diabetic retinopathy", *Comput. Med. Imag. Grap.* **33**, 608-622 (2009).

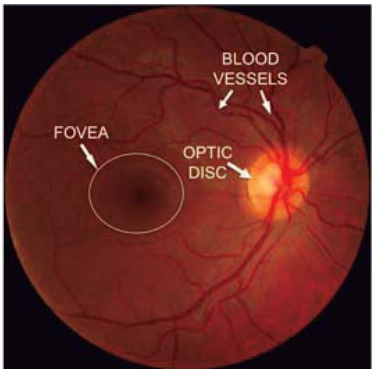
[17] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", *IEEE T. Image Process.* **13**, 1-13 (2004).

1. Introduction

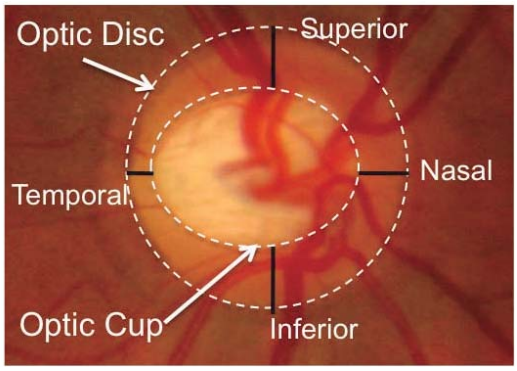
Computer vision and image processing techniques are found today in all fields of medical science, and are especially relevant to modern ophthalmology. Retinal images are widely used for diagnostic purposes by ophthalmologists. They provide vital information about the health of the sensory part of the visual system. Several diseases that can lead to blindness, such as glaucoma, manifest as anomalies in the retinal image [1]. Automatic segmentation and analysis of retinal images can therefore be used to detect pathological risk or damage, as well as to assist in diagnosis. Such segmentation normally includes the extraction of normal and abnormal features. The normal features of fundus images include the Optic Disc (OD), the fovea and blood vessels (Fig. 1(a)). Exudates and hemorrhages are among the main abnormal features. In the center of the optic disc, Fig. 1(b), one can distinguish a white cup-like area called the Optic Cup (OC).

Glaucoma is a disease characterized by progressive loss of peripheral vision. The decrease of optic nerve fibers caused by ocular hypertension may be visualized in the base of the optic nerve (optic disc). Two commonly used parameters for glaucomatous risk assessment are the cup-to-disc ratio (CDR) and the ISNT rule [2]. The former is estimated as the area occupied by the optic cup within the OD area, whereas the latter is an acronym to indicate the normal distribution of the neuroretinal rim (Inferior ≥ Superior ≥ Nasal ≥ Temporal). The dubbed ISNT rule is widely used in clinical practice. In recent years there has been an increasing interest for semi or automatic determination of the glaucomatous risk as well as the segmentation of the optic nerve head [1]. Li and Chutatape [3] used principal component analysis to locate the OD. Walter *et al* [4] detected the OD by means of morphological filtering techniques and the watershed transformation. In [5,6] they estimate the CDR on a pixel-based segmentation for glaucomatous risk assessment.

In this work we focus on OD segmentation. In the first section we review a previous approach and analyze possible effects in the segmentation of lossy compressed retinal images [7]. In section 3 we introduce a different strategy for OD segmentation based on color mathematical morphology and active contours. We evaluate the results on 20 retinal images and finally outline the conclusions.



(a)



(b)

Fig. 1. (a) Eye fundus image. (b) Optic nerve head region.

2. Optic nerve head segmentation: a previous approach

The OD is one of the main components on fundus images, and its localization is a relevant issue in retinal image analysis. The shape of the OD is an important indicator of various ophthalmic pathologies [1]. Several methods for localizing the OD boundary have been reported in the literature [5,8-10]. In [8] a two-stage method was reported: First, the image was processed using gray-level mathematical morphology to remove the blood vessels. Then a snake was manually placed to detect the OD boundary. The algorithm in [9] uses gradient vector flows to localize the OD. In this section we review an algorithm, previously developed in our research group, for OD feature extraction [5,6]. We analyze its strengths and drawbacks, as well as possible effects in the segmentation of lossy compressed retinal images.

We will briefly describe the main steps of the algorithm developed in [5,6]. The color images of the eye fundus are pre-processed using a sharpening algorithm [11] to smooth noise and sharpen edges by means of the Laplacian of Gaussian operator (LoG). Subsequently, the region of interest (ROI) is manually selected and transformed to polar coordinates for OD boundary extraction. This manual selection is a downside to the algorithm because the user is prompted to assign a manual initial guess for the OD's diameter. In the following stage, the image is transformed to polar coordinates and $\Delta E00$ color differences [12] are calculated between neighbor pixels in all radial directions. Pixels with highest color difference along each radial direction are sought to mark the OD boundary. The algorithm assumes a nearly circular shape as a general hypothesis and seeks on a pixel-based criterion for local distortions of the contour. The OD contour is extracted by interpolation and a back-conversion to Cartesian coordinates. In this approach the algorithm also deals with optic cup extraction by means of color seed selection, thresholding, and boundary extraction in the polar coordinate system. From this double segmentation of both the OD and the OC, they calculate the Cup-to-Disc area ratio and determine whether the ISNT rule is fulfilled.

The resulting segmentation is compared against manually labeled ground-truth produced by an expert. Fig. 2(a) shows a plausible satisfactory segmentation for a circular-shaped OD. There is great similarity as shape is concerned. A drawback of this approach is in the extraction of not so circular-shaped ODs. When the OD tends to a more elliptical shape (which could be a sign of potential risk of glaucoma at early stages) we show that the output may significantly differ from the ground-truth. The segmentation of an elliptical-shaped OD is shown in Fig. 2(b). The arrows indicate regions given by the elliptical shape, which cannot be correctly segmented due to the assumption that the OD is approximately circular.

(a) (b)

Fig. 2. Examples for OD segmentation using the algorithm in [5]. Ground-truth in white and contour produced by algorithm in black. (a) nearly circular shaped OD, (b) an elliptical shaped OD.

2.1 Compression effects in segmentation

Nowadays there are many medical instruments that acquire images and, with the purpose of saving memory space, most of these images are saved by default under lossy compression standards, such as classic JPEG. Until recently, the governing premise for lossy compression was that the average human observer would not notice the difference. However, the same cannot be said for computers. So far, standards for a reliable diagnosis have been established in radiology and pathology, whereas in ophthalmology this is still an open issue [15,16]. Therefore, our aim is to identify any compression effects in OD feature extraction, so as to determine possible problems in processing retinal images under compression and establish a safe-ground.

For this analysis we have used the algorithm of [5, 6]. We have used two of the most common lossy compression standards, classic JPEG and JPEG2000, to determine the effect in OD and OC segmentation. A set of 20 color retinal images were compressed for both standards under ratios of 1:2, 1:8, 1:11, 1:22, 1:31 and 1:47. The low compression ratios are used as a reference and the higher ratios correspond to the ones used in a subjective study for the assessment of diabetic retinopathy [15]. Fig. 3(a) shows the segmentation of both OD and OC from the original image without compression in TIFF format. An example of the effects of compression in segmentation of OD and OC is shown in Fig. 3(b)-(g). From these figures we can see that OC segmentation varies considerably under the effects of classic JPEG compression. On the contrary, the OC segmentation under JPEG2000 is more stable. A commonly used parameter to illustrate the corruption in images is the peak signal to noise ratio (PSNR). Although, there exists other parameters that correlate better with perceptual assessment such as the Structural Similarity Index (SSIM) [17], the PSNR gives enough information to set a basis for the comparison of our results. The PSNR was calculated for all compression ratios under JPEG and JPEG2000 standards. In Fig. 4(a) we show the average PSNR for the ROI. The standard JPEG2000 slightly outperforms JPEG.

To appropriately assess the effect in the segmentation we derived a measure based on the l_2-norm of the CDR and ISNT parameters. We recall that I, S, N, and T parameters correspond to the widths of the neuroretinal rim in the inferior,

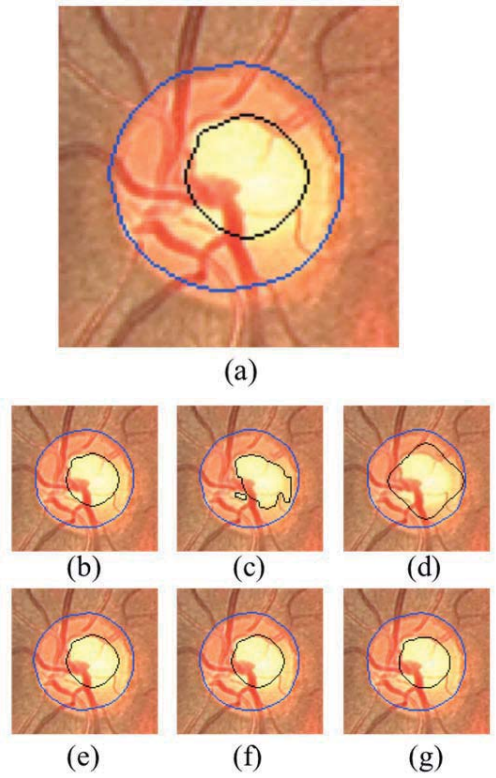


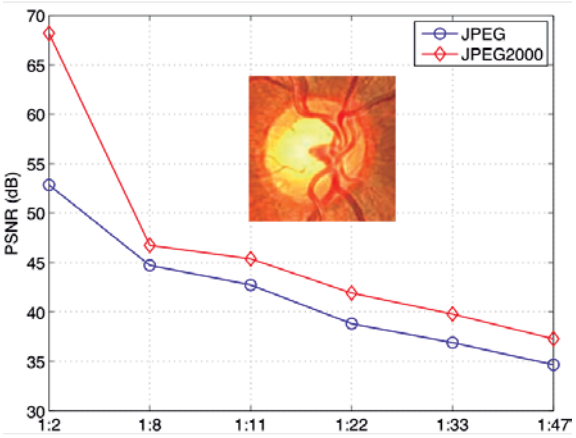
Fig. 3. (a) OD and OC segmentation in image without compression. Segmentation with JPEG compression ratios of (b) 1:2, (c) 1:8, (d) 1:11 and JPEG2000 (e) 1:2, (f) 1:8, and (g) 1:11.

superior, nasal and temporal directions, respectively. In Fig. 1(b), for instance, these quantities correspond to the black segment lengths measured in pixels. For each image, we take a 1-dimmensional vector expressed as $\mathbf{S}^j = \{S_i^j\}$, j=A,B, i=1,…4, $S_1^j = \mathrm{CDR}$, $S_2^j = \mathrm{I}$, $S_3^j = \mathrm{S}$, $S_4^j = \mathrm{N}$; where A and B correspond to the original and the compressed images, respectively. In other words, the elements of each vector correspond to the CDR and I, S, N parameters, where the last three components are normalized to T. Therefore, the normalized distance measure can be calculated as:

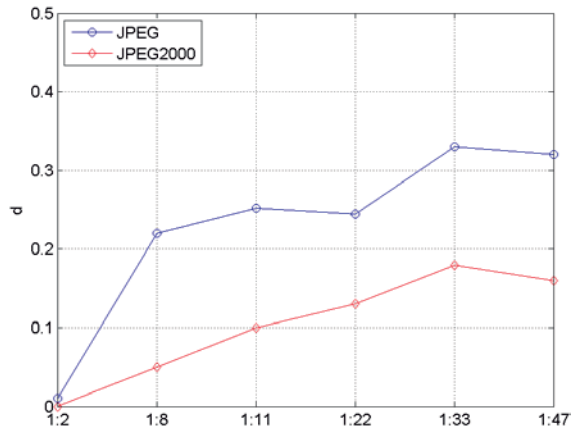
$$d = \frac{\left|\mathbf{S}^A - \mathbf{S}^B\right|}{\left|\mathbf{S}^A\right|} = \frac{\sqrt{\sum_{i=1}^{n}\left|S_i^A - S_i^B\right|^2}}{\sqrt{\sum_{i=1}^{n}\left|S_i^A\right|^2}}. \qquad (1)$$

The average distance measure for all 20 images of the set is shown in Fig. 4(b). It can be seen clearly that for very low compression ratios there is a negligible difference, whereas for mid and higher ratios the difference does become significant, particularly for JPEG compression. It is a well

known fact that JPEG and JPEG2000 lossy compression standards introduce random noise given by precision round off, but what might seem to affect a pixel-wise segmentation are the blocking artifacts from classic JPEG. Any algorithm that performs segmentation based on color seed selection is prone to produce inadequate segmentation under high ratio lossy compression. JPEG2000 compression appears to be more reliable than classic JPEG for OD and OC feature extraction.



(a)



(b)

Fig. 4. (a) PSNR in dB for the different compression ratios under JPEG and JPEG2000 standards. (b) Normalized distance measure for CDR and ISNT.

3. Optic disc localization by means of active contours

In this section we develop a strategy for OD boundary extraction in ocular fundus images. The pre-processing stage consists in performing color mathematical morphology to remove the blood vessel regions. Subsequently, an active contours approach is used to determine the OD boundary. An active contour is an energy-minimizing spline guided by external constraint forces influenced by image forces that pull it toward features such as lines and edges [14]. Mendels *et al* [8] presented a technique to localize the OD based on active contours formulated on gray-level images. In this work we formulate our approach in the Lab color space to take full advantage of the color features available for the pre-processing and feature extraction stages.

The segmentation algorithm is fully automatic. We have processed 20 color 24 bit-depth RGB fundus images of size 768 × 576 pixels. All images were acquired using a Topcon TRC-NW6S retinograph and a 3CCD Sony DXC-990P camera. The accuracy of the method is compared to ground-truth manual segmentation produced by an expert. The manual segmentation was carried out by an ophthalmologist with the use of an assisted graphical user interface developed in our research group.

3.1 Color mathematical morphology

Active contour methods generally work by locking onto homogeneous regions of a given image. This task is made extremely difficult since the OD region is fragmented into multiple subregions by blood vessels. Furthermore, the blood vessels enter the OD from different directions with a general tendency to concentrate around the nasal side of the OD region. Mathematical morphology can extract important shape characteristics and also remove irrelevant information. It typically probes an image with a small shape or template known as a structuring element. Using gray-level morphology, the operation can be applied to the intensity or lightness channel. Osareh *et al* [9] showed that in retinal images color morphology outperforms gray-level morphology, which results in more homogeneous regions and better preservation of the OD edges. They used a definition of color morphology within the CIELAB [12] color space (from now on Lab space) based on a color difference metric. We performed a closing operation, i.e. dilation to first remove the blood vessels and then an erosion to approximately restore the boundaries to their former position.

In color morphology, each pixel must be considered as a vector of color components.

Definitions of maximum and minimum operations on ordered vectors are necessary to perform basic operations. Hence, for each arbitrary point x in the color space, the definitions for dilation (I_d) and erosion (I_e) by structuring element B are:

$$I_d(x) = \{I(y): I(y) = \max[I(z)], z \in B_x\}, \quad (2)$$

$$I_e(x) = \{I(y): I(y) = \min[I(z)], z \in B_x\}. \quad (3)$$

Osareh et al. [9] introduced a lexicographical order to color morphology in the Lab space such that basic morphological operations could be performed. This is a problem-oriented formulation based on the knowledge that the OD region contains contrasting pixels: bright, almost saturated regions crossed by dark blood vessel regions. These color differences will reside in well-separated regions of the Lab color space. Given that color differences in the Lab space correspond to the metric distance between them, the basic morphological operations of dilation and erosion can be defined using the color difference of all pixels within the structuring mask to a certain reference point. The color difference within the Lab color space can be obtained using the Euclidean norm, and the reference point is established at the origin (0, 0, 0). The dilation is the furthest point from the origin, and the erosion is the point closest to the origin. The closing operation involves a dilation followed by an erosion. An example of closing using this formulation with a disc type-structuring element is shown in Fig. 5(a)-(b). It is evident that this approach produces a more homogeneous region while approximately preserving the OD edges Fig. 5(c).

3.2 Optic disc boundary extraction

The OD boundary is determined by fitting a geometric active contour model, namely the Chan-Vese [14] model. The performance of the methodology is evaluated by fitting the active contour onto the OD and comparing the resulting region against hand-labeled ground truth information. In general the active contour consists of a set of points placed near the contour of interest, which are gradually brought closer to the exact shape of the desired region in the image. This is carried out through iterative minimization of an energy function. The Chan-Vese model [14] establishes the following energy function for an image u_0:

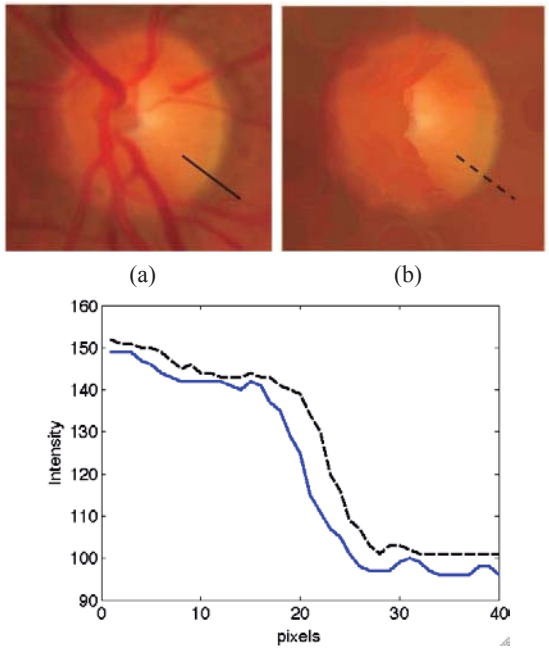


Fig. 5. Color morphology closing. (a) Original ROI (150x150 pixels) with OD inside. (b) Lab closing with a 25x25 disc-type structuring element. (c) Intensity profile cuts from (a) solid and (b) dashed.

$$F(c_1, c_2, C) = \int_{outside(C)} |u_0(x,y) - c_1|^2 dxdy + \int_{inside(C)} |u_0(x,y) - c_2|^2 dxdy + g(C), \quad (4)$$

where C is a piecewise parameterized curve (contour), g is any function evaluated at C, and c_1 and c_2 represent the average intensity value of u_0 inside and outside the curve, respectively. Minimizing the fitting error in (4), the model looks for the best partition of u_0 taking only two values, namely c_1 and c_2, and with one edge C, the boundary between these two regions, given by $\{u_0 \approx c_1\}$ and $\{u_0 \approx c_2\}$. Now, let us define a signed distance function ϕ that is zero exactly at C, that increases in absolute value with respect to the distance from C and that is positive inside and negative outside, as shown in Fig. 6. By doing so, we have defined implicitly the curve as, $C = \{(x, y) \mid \phi(x, y) = 0\}$. Therefore, the energy function can be expressed as:

$$F(c_1, c_2, \phi) = \int_{\phi>0} (u_0(x,y) - c_1)^2 H(\phi) dxdy + \int_{\phi<0} (u_0(x,y) - c_2)^2 (1 - H(\phi)) dxdy, \quad (5)$$

where $H(.)$ is the Heaviside function. Keeping c_1 and c_2 fixed, and minimizing F with respect to ϕ we obtain the associated Euler-Lagrange equation for ϕ. Parameterizing the descent direction by an artificial time $t \geq 0$ (or number of iterations), the equation in $\phi(t,x,y)$ (with $\phi(0,x,y) = \phi_0(x,y)$ defining the initial contour) is:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\left[div\left(\frac{\nabla \phi}{|\nabla \phi|}\right) - (u_0 - c_1)^2 - (u_0 - c_2)^2 \right] . \quad (6)$$

where $\delta(.)$ is the Dirac function. This partial differential equation can be solved numerically using a finite difference scheme. In relation to the problem at hand, we take the initial contour to be a circle big enough to fully contain the OD. From this circle a signed distance map is built for ϕ_0, fulfilling the condition to be positive inside the contour, zero exactly at the boundary, and negative outside. The iterative process consists in calculating the force from the image information, from the curvature penalty, and later evolving the curve (i.e. calculating ϕ_{n+1}).

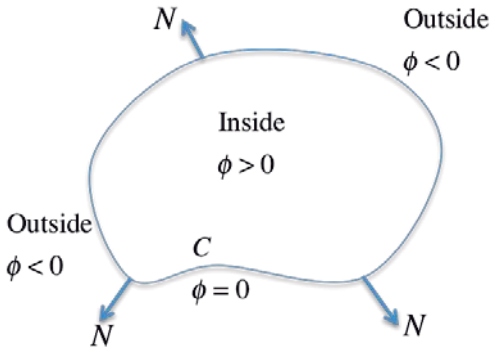


Fig.6. Curve C propagating in the normal direction.

3.3 Results

The ROI is selected manually as a window of 150×150 pixels, with the whole OD inside Fig. 5(a). We applied the Lab closing to all images using a symmetrical 25×25 pixels disc-structuring element since the blood vessels were determined not to be wider than 20 pixels. The Lab closing allowed to remove the blood vessels cleanly and provided the required uniform OD region Fig. 5(b) to initialize the active contour. The active contours approach requires an intensity or gray-scale image to perform the optimization procedure. Therefore, instead of solely using the lightness channel L and, more importantly, to be consistent with the color mathematical morphology approach, we decided to use the weighting function based on the Euclidean distance within the Lab space. This feature is fundamental to obtain a uniform OD region because our approach is based on the segmentation of pixels with similar color properties.

Following the color morphological pre-processing step, we initialized the contour as a circle with the center at the brightest area and with a diameter equivalent to 80% of the ROI diameter. From these initial conditions the active contour iteratively shrank towards the final boundary. The number of iterations for the final contour convergence was determined empirically and set to 450 for all cases. In Fig. 7(a)-(c) we show the hand-labeled ground-truth OD, the initial contour, and the final contour respectively.

In Fig. 7(d) we show the hand-labeled boundary together with the final contour to illustrate the close match achieved. We quantify the accuracy of the boundary localization against the manually labeled ground-truth produced by an expert. We use a simple and effective overlap measure of the match between two regions as:

$$M = \frac{n(R \cap T)}{n(R \cup T)} \times 100 , \quad (7)$$

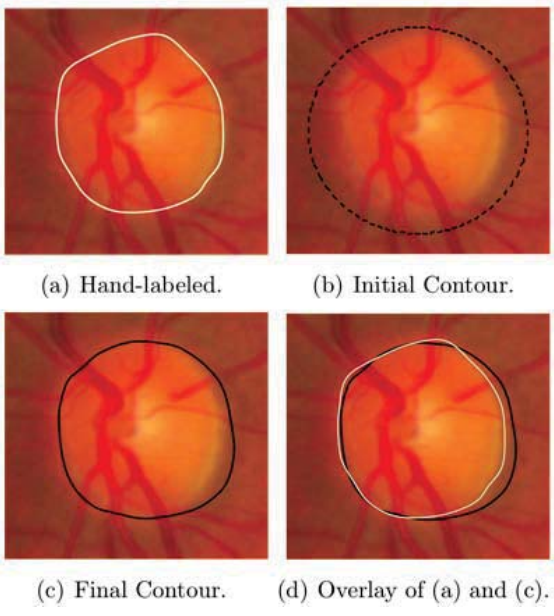
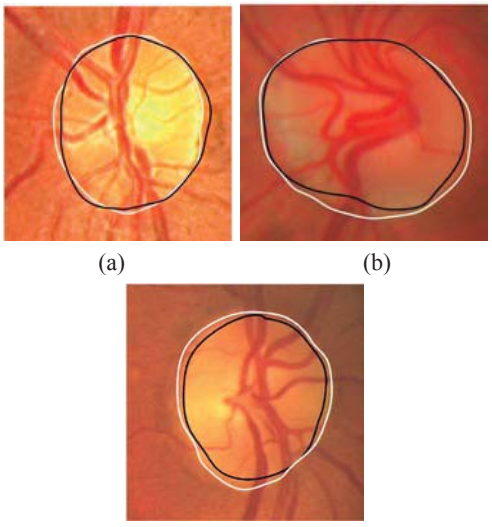


Fig.7. Optic disc boundary extraction results.

where R and T correspond to the ground-truth and the final OD contour region respectively, and $n(.)$ is the number of pixels in a region. In the optimal case, when both contours perfectly match $M = 100$. The measure M represents the accuracy. When compared with the hand-labeled ground-truth information from the expert, our method was able to localize the OD pixels in all test images with an average accuracy of 85.67 % ($\sigma = 7.82$). Additional tests are shown in Fig. 8 for some ODs whose shapes differ significantly from a circle. Notice the excellent agreement in Fig. 8(b) and the improvement achieved in Fig. 8(c) in comparison with the previous segmentation of Fig. 2(b).



(a) (b)

(c)

Fig.8. Other OD segmentation results. Ground-truth in white and algorithm output in black. M values are: (a) 92.61, (b) 90.32, (c) 88.15.

4. Summary and conclusions

In this work we have discussed two different approaches towards OD segmentation. The analysis of the algorithm in [5,6] revealed the need for a more general and robust approach, which would enable the segmentation of OD boundaries that differ considerably from a circular shape. As regards to compression effects in segmentation of the optic nerve head, we determined that degradation introduced by lossy compression plays an important role and cannot be neglected when processing compressed images. Nonetheless, our results showed that JPEG2000 compression might provide a safer ground for retinal image segmentation than classical JPEG. A different strategy for OD localization based on active contours was developed. The pre-processing stage consisted in performing color mathematical morphology. This provided a vessel-free OD region with uniform color distribution and preservation of sharp edge position. The active contours algorithm for OD segmentation yielded a fair approximation to the actual hand-labeled OD. Our method was able to achieve an average accuracy rate in pixel classification of 85.67 % ($\sigma = 7.82$).

Acknowledgements

The authors thank Isabel Martínez (MSc in Optometry) for helpful discussions and for providing some of the retinal images analyzed in the paper. This research has been funded by the Spanish Ministerio de Ciencia e Innovación y Fondos FEDER (project DPI2009-08879). The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

Retinal Image Analysis: Preprocessing and Feature Extraction

Journal of Physics: Conference Series, **274**(1), 012039, (2011).

Journal of Physics: Conference Series **274** (2011) 012039 doi:10.1088/1742-6596/274/1/012039

Retinal image analysis: preprocessing and feature extraction

Andrés G. Marrugo and María S. Millán

Grup d'Òptica Aplicada i Processament d'Imatge, Departament d'Òptica i Optometria
Univesitat Politècnica de Catalunya, Spain

E-mail: andres.marrugo@upc.edu

Abstract. Image processing, analysis and computer vision techniques are found today in all fields of medical science. These techniques are especially relevant to modern ophthalmology, a field heavily dependent on visual data. Retinal images are widely used for diagnostic purposes by ophthalmologists. However, these images often need visual enhancement prior to apply a digital analysis for pathological risk or damage detection. In this work we propose the use of an image enhancement technique for the compensation of non-uniform contrast and luminosity distribution in retinal images. We also explore optic nerve head segmentation by means of color mathematical morphology and the use of active contours.
Keywords – retinal image, medical imaging, eye fundus, optic disc.

1. Introduction

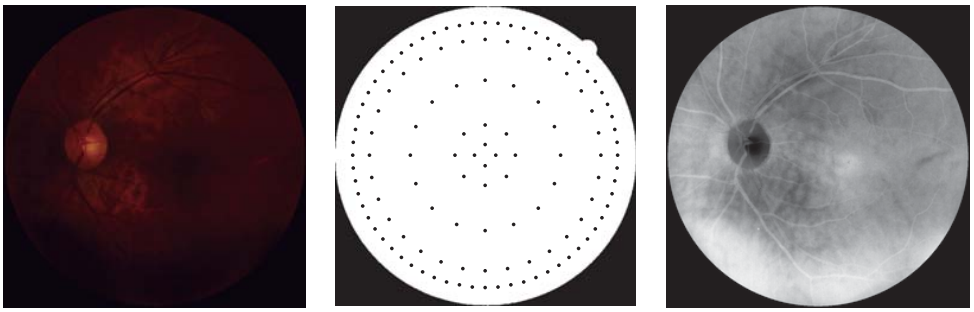
Over the last decade, color digital photography has been recognized as an acceptable modality for documenting retinal appearance as it provides vital information about the health of the sensory part of the visual system [1]. Automatic segmentation and analysis of retinal images can be used to detect pathological risk or damage, and to assist in diagnosis.

Digital image analysis techniques in retinal imaging span from preprocessing techniques for visual enhancement or for further processing and ultimately any sort of feature extraction or segmentation. This work aims to illustrate the relationship of the ensemble of these different techniques. The paper mainly consists of two parts: The first part deals with an enhancement technique to compensate for uneven illumination and poor contrast in retinal images. In the second, we develop a strategy for optic disc (OD) segmentation based on color mathematical morphology as a preprocessing stage.

2. Retinal image enhancement

Retinal images are acquired with a digital fundus camera which captures the illumination reflected from the retinal surface. Despite controlled conditions, many retinal images suffer from non-uniform illumination given by several factors: the curved surface of the retina, pupil dilation (highly variable among patients), or presence of diseases, among others. The curved retinal surface and the geometrical configuration of the light source and camera, lead to a poorly illuminated peripheral part of the retina with respect to the central part (Figure 1(a)).

Several techniques have been used to enhance retinal images. Histogram equalization has been shown to be inappropriate for retinal images [2]. A local normalization of each pixel to



(a) (b) (c)

Figure 1. (a) Retinal image with uneven illumination and contrast, (b) non-uniform sampling grid, and (c) first principal component of (a) from PCA analysis.

zero mean and unit variance aims to compensate lighting variation and enhancing local contrast but also introduces artifacts [2]. Histogram matching between the red and green planes have been used as a preprocessing step for vessel segmentation [3]. This improves the contrast of gross dark features like vessels but reduces the contrast of bright objects and tiny dark objects like micro-aneurysms. While most of the aforementioned methods are motivated by automatic analysis, as a preprocessing stage, they are all formulated for a single color plane or for gray-scale images.

Color retinal image enhancement is required for human visual inspection or for the application of vector processing techniques. The work of Foracchia et al. [4] aimed to introduce a strategy for luminosity and contrast enhancement on each color plane of the RGB color space, independently. This approach tended to produce hue-shifting related artifacts, given by the introduction of new colors to the image. More recently, Joshi et al. [5], proposed a strategy that would avoid the color artifacts by performing the enhancement on single color plane to compensate equally every channel and ultimately perform linear color remapping. Our approach is based on [5, 4], and aims to improve the compensation of luminosity by using principal component analysis (PCA) [6].

2.1. Image enhancement on a single color plane

The main idea is that the image can be enhanced by estimating the background luminosity and contrast distribution in order to compensate for uneven illumination. Thus, the enhanced image $U(x, y)$ is expressed as:

$$U(x, y) = \frac{I(x, y) - L(x, y)}{C(x, y)}, \tag{1}$$

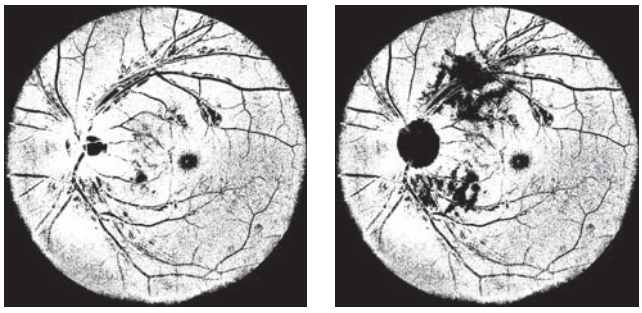
where I is the original image, C and L are the contrast and luminosity drifts, respectively. L and C can also be understood in terms of gain and offset. They have to be estimated by sampling the original image. The sampling approach of [4] divides the whole image into a square sampling grid, whereas in [5] they use a more intuitive sampling scheme based on the knowledge of the illumination distribution that leads to less computational burden. Therefore, we decided to use a similar type of non-uniform sampling grid shown in figure 1(b). The sampling is coarse in the central region and dense in the periphery.

As described in [5], this enhancement is oriented towards compensating the green channel of the RGB retinal image because it is the component with highest contrast. Thus, in order to estimate accurately the background luminosity the first step must be to separate the image into a set of background and foreground pixels. The foreground set consists of pixels belonging to retinal structures, such as the OD and blood vessels, while the remaining pixels are considered to belong to the background. The next step is to estimate the C and L components using

only the background pixels. As pointed out in [5] this strategy is motivated by the fact that the retinal structures can bias the luminosity component. For instance, the OD is a naturally high luminosity zone, while the blood vessels typically exhibit low luminosity. The sampling scheme is as follows: for each sampling point on the grid we take a window of size $w_0 \times w_0$ large enough to include retinal structures and the background. We compute the local mean μ and the standard deviation σ for each point. We perform bi-cubic interpolation to obtain μ_0 and σ_0 for all (x, y) points. To identify background pixels the criteria in [4] states that a pixel is considered to belong to the background if its Mahalanobis distance from μ_0, defined as:

$$D(x, y) = \left| \frac{I(x, y) - \mu(x, y)}{\sigma(x, y)} \right|, \tag{2}$$

is lower that a certain threshold t, which for this work will be taken as 1. This threshold is somewhat critical, because, as pointed out before, any retinal structure that does not belong to the background, especially the OD, can bias the background components. Thus, to ensure that the OD region is not taken into account in this estimation we developed a strategy using PCA. In [6] a PCA based model was used to approximately localize the OD region in an intensity retinal image. More recently, in [7] they developed a model for retinal pigment identification using independent component analysis (ICA) on the RGB retinal image, although no experiments were carried out including the OD. The main drawback of ICA is that it does not prioritize the output components, whereas in PCA this is not an issue. Therefore, here we have used PCA on the three RGB channels to identify the OD region. The first principal component from the image in Figure 1(a) is shown in Figure 1(c). It can clearly be seen that the OD region has different properties than the surrounding retinal regions. Using the first principal component the OD region can be entirely left out from $D(x, y)$ as as shown in Figure 2(b).



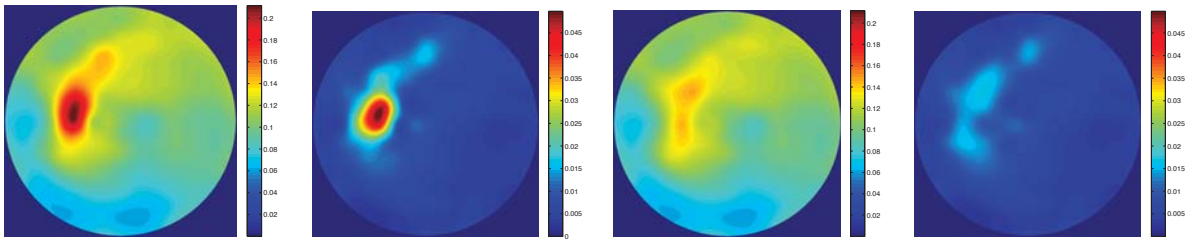
(a) (b)

Figure 2. Background pixel classification using (a) the strategy in [5] and (b) with additional PCA analysis. Notice that the OD region has been left out in order not to bias the estimation of the luminosity component.

The following step involves repeating the sampling scheme but this time including just background pixels and with a smaller window of size $w_1 \times w_1$ to increase the precision in the estimation of μ_1 and σ_1. Bi-cubic interpolation is carried out to obtain μ_1 and σ_1 for all (x, y). From [4] L and C can be approximated as μ_1 and σ_1. The enhanced image is obtained by applying (1). In our experiments we set $w_0 = 125$ and $w_1 = 51$. The estimated C and L components are shown in Figure 3. To illustrate the impact on the single channel enhancement by applying (1) both images are shown in Figure 4.

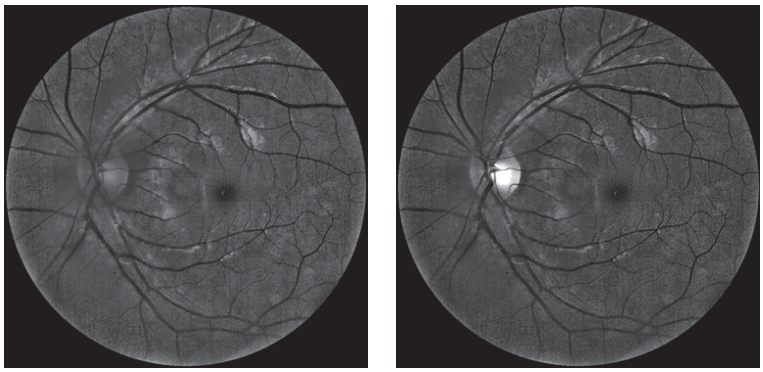
2.2. Color remapping

After the single channel enhancement we perform the following color remapping: given a color image with color components (r, g, b), the single plane enhancement is applied to the g plane



(a) (b) (c) (d)

Figure 3. Estimated L and C components using background pixels (a-b) from Figure 2(a) and (c-d) from Figure 2(b). For the purpose of comparison (a) and (c), as well as (b) and (d) are in the same scale. Notice how the OD region has little influence on the components in (c-d).



(a) (b)

Figure 4. Image enhancement on single channel from (a) the strategy in [5] and (b) with additional PCA analysis. In (b) the illumination in the surrounding area of the OD has not been modified significantly compared to that in (a).

and g_{enh} is obtained. Next, the enhanced color image $(\hat{r}, \hat{g}, \hat{b})$ is computed on pixel basis as:

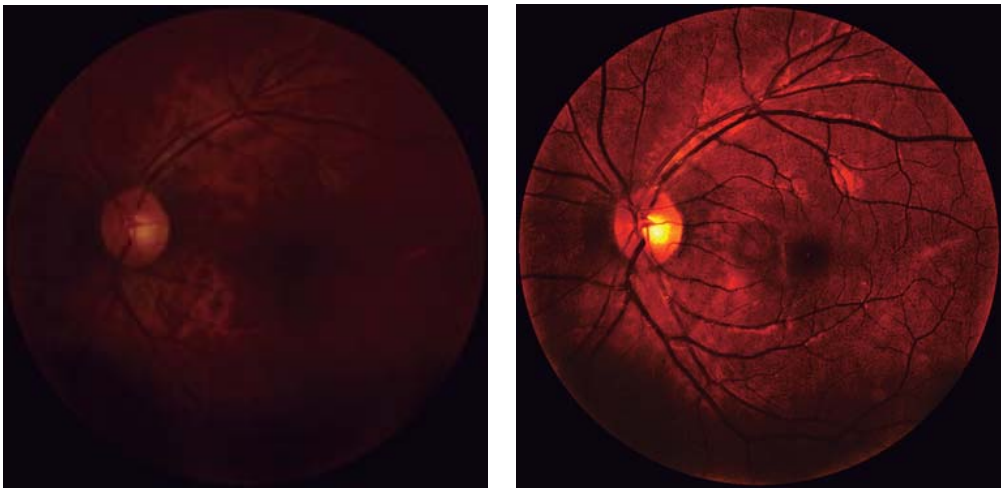
$$\hat{r} = \frac{g_{enh}}{v} \cdot r,\ \hat{g} = \frac{g_{enh}}{v} \cdot g,\ \hat{b} = \frac{g_{enh}}{v} \cdot b, \tag{3}$$

where v is a scalar defined as $v = max[r_{max}, g_{max}, b_{max}]$ to play a normalization role in the enhancement. Thus, the ratio of the original r, g, and b is maintained. Figure 5(b) shows the enhanced color retinal image. Notice that it has good luminosity and different retinal structures are contrasted well against the background.

3. Optic disc segmentation by means of active contours

In this section we develop a strategy for OD boundary extraction in ocular fundus images [10]. The preprocessing stage consists in performing color mathematical morphology to remove the blood vessel regions. Subsequently, an active contours approach is used to determine the OD boundary. An active contour is an energy minimizing spline guided by external constraint forces influenced by image forces that pull it toward features such as lines and edges [8]. Mendels et al. [9] presented a technique to localize the OD based on active contours formulated on gray-level images. In this work we formulate our approach in the Lab color space to take full advantage of the color features available for the preprocessing and feature extraction stages.

The segmentation algorithm is fully automatic. We have processed 20 color 24 bit-depth RGB fundus images of size 768×576 pixels. All images were acquired using a Topcon TRC-NW6S



(a) (b)

Figure 5. (a) Original color retinal image with uneven illumination and (b) resulting enhanced color retinal image.

retinograph and a 3CCD Sony DXC-990P camera. The accuracy of the method is compared to ground-truth manual segmentation produced by an expert. The manual segmentation was carried out by an ophthalmologist with the use of an assisted graphical user interface.

3.1. Color mathematical morphology

Active contour methods generally work by locking onto homogeneous regions of a given image. This task is made extremely difficult since the OD region is fragmented into multiple subregions by blood vessels. Furthermore, the blood vessels enter the OD from different directions with a general tendency to concentrate around the nasal side of the OD region. Mathematical morphology can extract important shape characteristics and also remove irrelevant information. It typically probes an image with a small shape or template known as a structuring element. Using gray-level morphology, the operation can be applied to the intensity or lightness channel. Osareh et al. [11] showed that in retinal images color morphology outperforms gray-level morphology, which results in more homogeneous regions and better preservation of the OD edges. They used a definition of color morphology within the CIELAB [12] color space (from now on Lab space) based on a color difference metric. We performed a closing operation, i.e. dilation to first remove the blood vessels and then an erosion to approximately restore the boundaries to their former position.

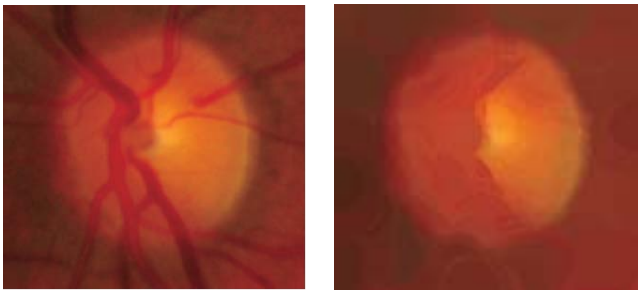
In color morphology, each pixel must be considered as a vector of color components. Definitions of maximum and minimum operations on ordered vectors are necessary to perform basic operations. Hence, for each arbitrary point x in the color space, the definitions for dilation (I_d) and erosion (I_e) by structuring element B are:

$$I_d(x) = \{I(y) : I(y) = \max[I(z)], z \in B_x\} \tag{4}$$

$$I_e(x) = \{I(y) : I(y) = \min[I(z)], z \in B_x\}. \tag{5}$$

A proper lexicographical order in the Lab space was introduced in [11] such that basic morphological operations could be performed. This is a problem-oriented formulation based on the knowledge that the OD region contains contrasting pixels: bright, almost saturated regions crossed by dark blood vessel regions. These color differences will reside in well-separated

regions of the Lab color space. Given that color differences in the Lab space correspond to the metric distance between them, the basic morphological operations of dilation and erosion can be defined using the color difference of all pixels within the structuring mask to a certain reference point. The color difference within the Lab color space can be obtained using the Euclidean norm, and the reference point is established at the origin $(0, 0, 0)$. The dilation is the furthest point from the origin, and the erosion is the point closest to the origin. The closing operation involves a dilation followed by an erosion. An example of closing using this formulation with a disc type-structuring element is shown in Figure 6. It is evident that this approach produces a more homogeneous region while approximately preserving the OD edges.



(a) (b)

Figure 6. Color morphology closing. (a) Original ROI with optic disc inside, (b) Lab closing with a 25x25 disc-type structuring element. The result produces a homogeneous region while approximately preserving the OD edges.

3.2. Optic disc segmentation

The optic disc boundary is determined by fitting a geometric active contour model, namely the Chan-Vese model [8]. The active contour consists of a set of points placed near the contour of interest that are gradually brought close to the exact shape of the desired region in the image. The performance is evaluated by fitting the active contour onto the optic disc and comparing with hand-labeled ground truth. The Chan-Vese model establishes the following energy function in level-set formulation for an image u_0:

$$F(c_1, c_2, \phi) = \int_{\phi>0} |u_0(x,y) - c_1|^2 H(\phi) dxdy + \int_{\phi<0} |u_0(x,y) - c_2|^2 (1 - H(\phi)) dxdy + g(\phi), \quad (6)$$

where ϕ is a signed distance function that is zero exactly at the boundary Q, that increases in absolute value with respect to the distance from Q and that is positive inside and negative outside. c_1 and c_2 represent the average intensity value of u_0 inside and outside the curve, respectively. $H(.)$ is the Heaviside function and $g(.)$ is any function evaluated over the boundary. If we minimize F with respect to ϕ and parameterizing the descent direction by an artificial time $t > 0$ (number of iterations) we obtain:

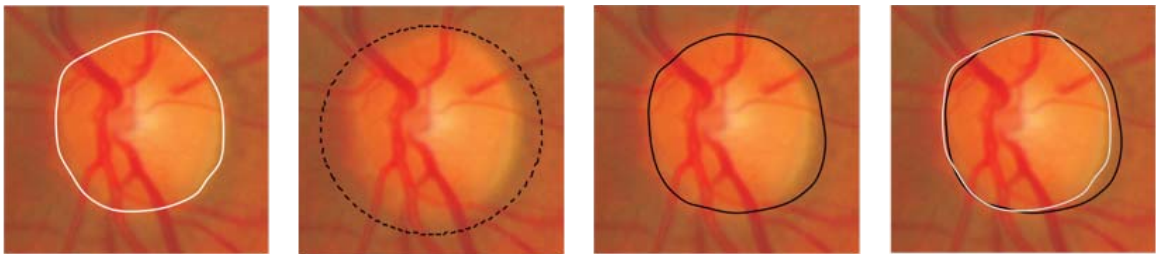
$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - (u_0 - c_1)^2 - (u_0 - c_2)^2 \right], \quad (7)$$

where $\delta(.)$ is the Dirac function. By minimizing the fitting error in (6) the model looks for the best partition of u_0 (see [8] for more details).

3.3. Results

The ROI is selected manually as a window of 150 × 150 pixels, with the whole OD inside. We applied the Lab closing to all images using a symmetrical 25-by-25 pixels disc-structuring element since the blood vessels were determined not to be wider than 20 pixels. The Lab closing allowed to remove the blood vessels cleanly and provided the required uniform OD region to initialize the active contour. The active contours approach requires an intensity or gray-scale image to perform the optimization procedure. Therefore, instead of solely using the lightness channel L and, more importantly, to be consistent with the color mathematical morphology approach, we decided to use the weighting function based on the Euclidean distance within the Lab space. This feature is fundamental to obtain a uniform OD region because our approach is based on the segmentation of pixels with similar color properties.

Following the color morphological preprocessing step, we initialized the contour as a circle with the center at the brightest area and with a diameter equivalent to 80% of the ROI diameter. From these initial conditions the active contour iteratively shrank towards the final boundary. The number of iterations for the final contour convergence was determined empirically and set to 450 for all cases. In Figure 7(a)-(c) we show the hand-labeled ground-truth OD, the initial contour, and the final contour respectively.



(a) (b) (c) (d)

Figure 7. Optic disc segmentation results. (a) hand labeled contour, (b) initial contour, (c) final contour, and (d) overlay of (a) and (c).

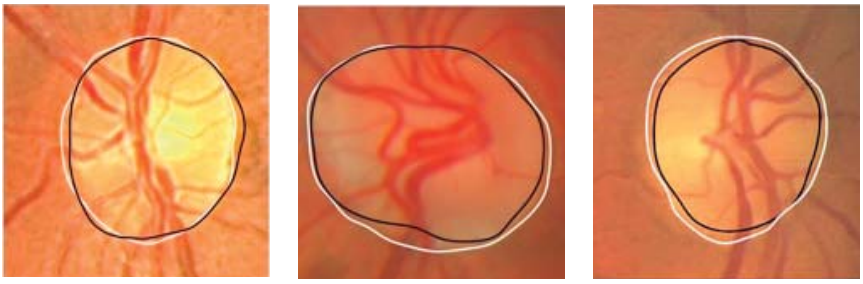
In Figure 7(d) we show the hand-labeled boundary together with the final contour to illustrate the close match achieved. We quantify the accuracy of the boundary localization against the manually labeled ground-truth produced by an expert. We use a simple and effective overlap measure of the match between two regions as:

$$M = \frac{n(R \cap T)}{n(R \cup T)} \times 100, \tag{8}$$

where R and T correspond to the ground-truth and the final OD contour region respectively, and $n(.)$ is the number of pixels in a region. In the optimal case, when both contours perfectly match $M = 100$. The measure M represents the accuracy. When compared with the hand-labeled ground-truth information from the expert, our method was able to localize the OD pixels in all test images with an average accuracy of 85.67% ($\sigma = 7.82$). Additional tests are shown in Figure 8.

4. Summary and conclusions

In this work we have developed a strategy for retinal image enhancement and OD segmentation. We showed that the problem of non-uniform illumination and poor contrast in retinal images may be addressed via an image enhancement technique based on the knowledge of luminosity distribution in the retina. With the use of additional PCA analysis we were able to leave



(a) (b) (c)

Figure 8. Other OD segmentation results. Ground-truth in white and algorithm output in black. M values are: (a) 92.61, (b) 90.32, (c) 88.15.

out the OD region so as to estimate proper luminosity components for the illumination compensation. The resulting enhanced image showed remarkable gain in contrast related to retinal structures against the background. The background exhibited a much more uniform illumination distribution, in spite of an minor decrease in intensity.

As regards to feature extraction, we presented a strategy for optic disc segmentation by means of active contours. The preprocessing stage consisted in performing color mathematical morphology to provide a vessel-free OD region with uniform color distribution and preservation of sharp edges. The active contours algorithm yielded a fair approximation to the actual hand-labeled optic disc.

Acknowledgments

This research has been funded by the Spanish Ministerio de Ciencia e Innovacion y Fondos FEDER (project DPI2009-08879). The first author also thanks the Spanish Ministerio de Educacion for an FPU doctoral scholarship.

References

[1] Patton N, Aslam T, MacGillivray T, Deary I, Dhillon B, Eikelboom R, Yogesan K and Constable I 2006 *Progress in Retinal and Eye Research* **25** 99–127
[2] Feng P, Pan Y, Wei B, Jin W and Mi D 2007 *Pattern Recognition Letters* **28** 516 – 522
[3] Salem N and Nandi A 2007 *Journal of the Franklin Institute* **344** 243–256
[4] Foracchia M, Grisan E and Ruggeri A 2005 *Medical Image Analysis* **9** 179–190
[5] Joshi G and Sivaswamy J 2008 *Computer Vision, Graphics & Image Processing, 6th Indian Conf. on* 591 – 598
[6] Li H and Chutatape O 2003 *Journal of Electronic Imaging* **12** 97–105
[7] Fadzil M, Nugroho H, Venkatachalam P, Nugroho H and Izhar L 2008 *Biomedical Engineering, 4th Kuala Lumpur Int. Conf. on* 555–558
[8] Chan T and Vese L 2001 *Image Processing, IEEE Transactions on* **10** 266–277
[9] Mendels F, Heneghan C, Harper P, Reilly R and Thiran J 1999 *Proc. 1st Joint BMES/EMBS Conf* 1139
[10] Marrugo AG and Millán MS 2010 *Opt. Pura Apl.* **43** 79–86
[11] Osareh A, Mirmehdi M, Thomas B and Markham R 2002 *Pattern Recognition, 16th Int. Conf. on* **1** 743 – 746
[12] Zhang X and Wandell B 1997 *Journal of the Society for Information Display* **5** 61–63

Image Analysis in Modern Ophthalmology: From Acquisition to Computer Assisted Diagnosis and Telemedicine

SPIE Photonics Europe, Proceedings SPIE, **8436**:84360C, (2012).

Invited Paper

Image Analysis in Modern Ophthalmology: From Acquisition to Computer Assisted Diagnosis and Telemedicine

Andrés G. Marrugo[a], María S. Millán[a], Gabriel Cristóbal[b], Salvador Gabarda[b], Michal Šorel[c], and Filip Šroubek[c]

[a]Department of Optics and Optometry, Universitat Politècnica de Catalunya Violinista Vellsolà 37, 08222 Terrassa, Spain.

[b]Instituto de Óptica, Consejo Superior de Investigaciones Científicas Serrano 121, 28006 Madrid, Spain.

[c] Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 18208 Prague 8, Czech Republic.

ABSTRACT

Medical digital imaging has become a key element of modern health care procedures. It provides visual documentation and a permanent record for the patients, and most important the ability to extract information about many diseases. Modern ophthalmology thrives and develops on the advances in digital imaging and computing power. In this work we present an overview of recent image processing techniques proposed by the authors in the area of digital eye fundus photography. Our applications range from retinal image quality assessment to image restoration via blind deconvolution and visualization of structural changes in time between patient visits. All proposed within a framework for improving and assisting the medical practice and the forthcoming scenario of the information chain in telemedicine.

Keywords: Computer-aided diagnosis, medical image, retinal image, telemedicine, ophthalmology.

1. INTRODUCTION

Ophthalmology is no longer a stand-alone branch of medicine conducted exclusively by specialists with the sole purpose of providing medical aid for the visual health of the general population–it is much, much more than that. As it is conceived today, ophthalmology is undisputedly an interdisciplinary field in both research and clinical practice. A field that in the last decade has shown that digital information based systems can be both clinical and cost effective with high levels of patient satisfaction.[1] However, successful deployment is not without great difficulty.

While ophthalmology *per se* involves a great number of sub-specialties and also an ever increasing number of probing techniques, this work focuses on the main ocular fundus imaging modality: color fundus photography. Fundus imaging or fundus photography is basically the process whereby a 2D-representation of the 3D-retinal tissues projected onto the imaging plane is obtained using reflected light.[2] In color fundus photography the image intensities represent the amount of reflected red (R), green (G), and blue (B) wavebands, as determined by the spectral sensitivity of the sensor. Fundus imaging plays a key role in the diagnosis and management of ophthalmologic disorders, such as diabetic retinopathy, glaucoma, and age-related macular degeneration; all of them being the most prevalent causes of blindness in the industrialized world.[2]

In this paper we provide a brief introduction to image analysis in modern ophthalmology that comprises a general view of the field, the presentation of several examples of image analysis techniques, and perspectives on future developments in the field.

Further author information: (Send correspondence to A.G. Marrugo)
A.G.M.: E-mail: andres.marrugo@upc.edu, Telephone: +34 93 739 8678; http://www.goapi.upc.edu/

Optics, Photonics, and Digital Technologies for Multimedia Applications II, edited by Peter Schelkens, Touradj Ebrahimi, Gabriel Cristóbal, Frédéric Truchetet, Pasi Saarikko,
Proc. of SPIE Vol. 8436, 84360C · © 2012 SPIE · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.921299

1.1 A link to the past

The eye fundus has been observed since 1850 with the invention of the ophthalmoscope by the German physician Hermann Von Helmholtz.[3] This was an instrument that enabled the examination of the retina by using a bright light near the eye and shining it into the patient's pupil. However, it was not until the mid 1920s that the Carl Zeiss Company made available the first commercial fundus camera. Many were the limitations in clinical use of fundus photography in the 1930s and 1940s which can be attributed to the difficulty in obtaining good quality images.[3] Significant progress was made in later decades and fundus photography became ubiquitous in the practice of ophthalmology providing a means for recording, storing, and indexing at low cost the images of a patient. This opened many possibilities: specific patients could easily be recorded and catalogued, longitudinal changes could be documented, and later retrieval of images could foster scientific research. Other important imaging modalities appeared only to enhance diagnostic and observational capabilities in ophthalmology such as: fluorescent angiography, modern digital fundus photography, stereo fundus photography, confocal laser opthalmoscopy, and optical coherence tomography. Out of all of these imaging modalities, however, it may be said that fundus photography is the one that provides a more general fundus examination with relatively simple and affordable equipment, and little patient intervention. Interestingly, it has been suggested that information extracted from the eye fundus could be useful in a variety of diseases such as heart disorders, stroke, hypertension, peripheral vascular disease and diabetic retinopathy.[4]

1.2 The new paradigm: computer-aided diagnosis and telemedicine

With the advances of computer technology, various types of computer-aided diagnosis (CAD) systems[5,6] have been developed in recent years. The main idea of CAD is to assist medical staff in interpreting medical images by using dedicated computer systems to provide "second opinions". The final medical decision is made by the physicians. Studies on CAD systems show that it can help to improve diagnostic accuracy, lighten the burden of increasing workload, reduce missed disease detection due to fatigue, overlooked data, and improve inter- and intra-reader variability.[5,7] Meanwhile a number of fundus image CAD systems have also been developed for the diagnosis of various types of ocular diseases such as glaucoma,[8] and diabetic retinopathy.[9,10] These CAD systems have the potential to provide an alternative solution to mass screening programs that need to examine a vast number of fundus images as fast as possible. In a recent work by Sánchez *et al.*,[10] they were able to show that the performance of a CAD system for diabetic retinopathy screening could be comparable with that of human experts. However, while this represents unprecedented performance for CAD systems it does not translate readily to clinical practice. Furthermore, this should not be understood in the sense that CAD systems are to replace specialists, instead they are to ensure that specialists spend more time dealing with the ill by serving the purpose of screening. Nevertheless, further extensive and thorough evaluation is still needed before deployment in clinical practice.

In other regards, CAD systems are also pivotal to the practice of telemedicine. Telemedicine is basically the use of telecommunication and information technologies in order to provide clinical health care at a distance. Traditional examination of the retina requires dilated pupils, a skilled examiner, and a visit to the ophthalmologist, typically in a separate location from the primary care center. Numerous studies in the United States have shown that many diabetes patients fail to seek or receive this important examination on a regular basis.[11] This shortfall has been attributed to a variety of factors but mainly due to socioeconomic and health system barriers. These problems can be overcome by incorporating retinopathy screening into primary care practices, using telemedicine powered by CAD systems for the evaluation of retinal images. The screening process can be simpler and more cost-efficient than sending the patients to the ophthalmologist's office for a live evaluation. A successful and reliable telemedicine retinopathy screening system must be comparable to a live retinal examination in the detection of vision-threatening retinopathy. There are a number of constraints so as to achieve an adequate sensitivity and specificity, and to date there is no general consensus on how to approach them. Meanwhile, a promising strategy rises from the use of license-free web-based software, standard interfaces, and flexible protocols to allow primary care providers to adopt retinopathy screening with minimal effort and resources.[11]

2. PREPROCESSING OF RETINAL IMAGES

The main purpose of preprocessing techniques is to reduce the effect of image variation via normalization of the retinal image against a reference model or data set for subsequent visualization, processing or analysis.

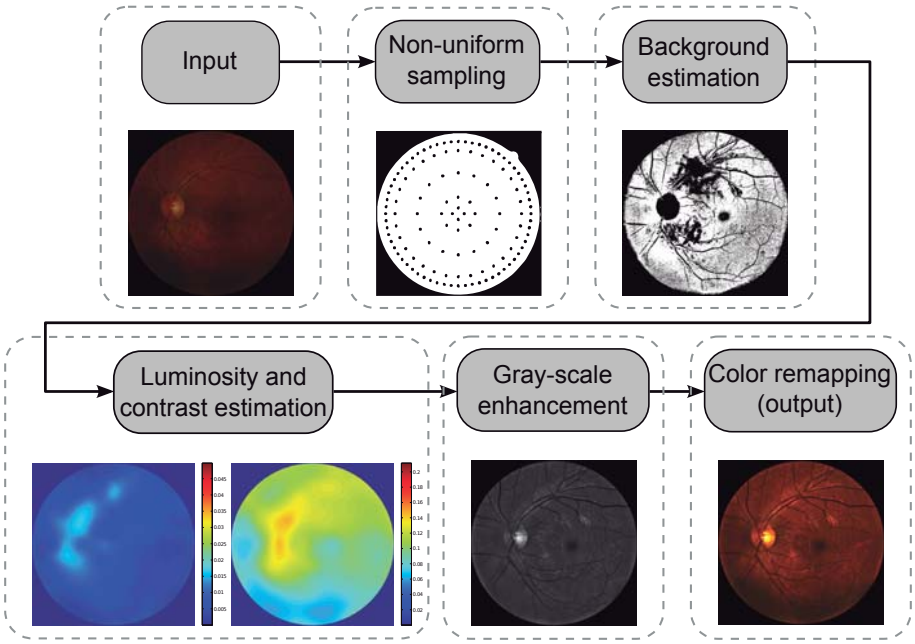


Figure 1. Illumination compensation algorithm

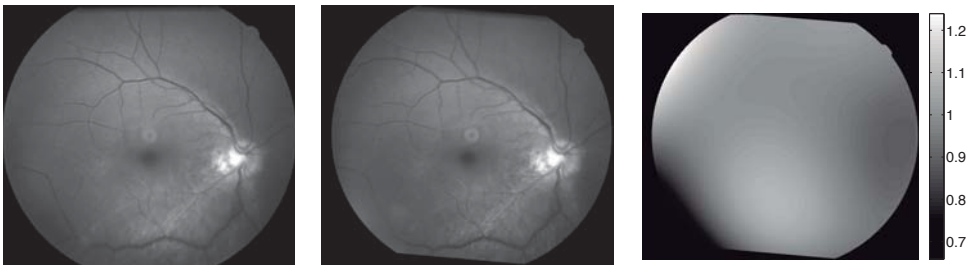
Variations may occur within the same image or between images. The main differences between images are most likely due to differences in cameras, illumination, field of view and retinal pigmentation. Thus, in order to extract meaningful information from an image, it is necessary to compensate for this variability. General preprocessing tasks for both monochromatic and color retinal images may be broadly categorized in terms of the correction for non-uniform illumination, contrast enhancement and color normalization.[12]

Despite controlled conditions, many retinal images suffer from non-uniform illumination. The curved retinal surface and the geometrical configuration of the light source and camera, lead to a poorly illuminated periphery. This problem can be approached in a number of ways. There have been many works on this topic,[12] however they can be loosely classified into single image or multiple image compensation techniques. Several single image techniques have been used to enhance retinal images. Histogram equalization has been shown to be inappropriate for retinal images.[13] A local normalization of each pixel to zero mean and unit variance aims to compensate lighting variation and enhancing local contrast but also introduces artifacts.[13] Histogram matching between the red and green planes has been used as a preprocessing step for vessel segmentation.[14] This improves the contrast of gross dark features like vessels but reduces the contrast of bright objects and tiny dark objects like microaneurysms. While most of the aforementioned methods are motivated by automatic analysis, as a preprocessing stage, they are all formulated for a single color plane or for gray-scale images. Color retinal image enhancement is required for human visual inspection or for the application of vector processing techniques.

An example of a single image illumination compensation algorithm is depicted in Fig. 1. This algorithm was described in Ref. 15 and is based on a simple model of degradation proposed by Foracchia *et al.*[16] The main idea is that the image can be enhanced by estimating the background luminosity and contrast distribution in order to compensate for uneven illumination. Therefore, the enhanced image $U(x, y)$ is expressed as:

$$U(x,y) = \frac{I(x,y) - L(x,y)}{C(x,y)} \quad , \tag{1}$$

where I is the original degraded image, C and L are the contrast and luminosity drifts, respectively. C and L can also be understood in terms of gain and offset. They have to be estimated by sampling the original image. This is achieved by using a non-uniform sampling grid as shown in Fig. 1. The sampling is coarse in the central region and dense in the periphery. This is carried out on the green channel of the RGB retinal image because is the channel with highest contrast. The background pixels are estimated by a procedure described in



(a) (b) (c)

Figure 2. Illumination compensation for two retinal images (a) and (b). (c) Compensation function k.

Ref. 15. From these pixels the background luminosity and contrast components are computed to produce the image enhancement given by Eq. (1). The final color enhanced retinal image is obtained by performing a color remapping on the gray-scale image that preserves the RGB ratios, however it may introduce color modifications.

Multiple image illumination compensation techniques are mainly motivated by the need for a processing strategy that requires two or more images, e.g. image comparison or change detection,[17] multichannel deconvolution,[18] sequential multispectral imaging,[19] etc. In Ref. 20 we described such a technique for compensating the uneven illumination distribution from two images of the same retina acquired at different moments for further processing. The basic idea is a follows: the illumination distribution can be compensated by adjusting the intensity values on one image to approximately match that of the other while satisfying a predetermined illumination model. Because the illumination of the retina is formed by a slowly varying light field over a smooth surface it can be modeled by a low-order parametric surface, in this case a 4th-order polynomial. The compensation is then formulated via a parametric surface fitting equation

$$\arg\min_{k} \|I_1(x,y) - k(x,y) \cdot I_2(x,y)\| \ , \tag{2}$$

where I_1 and I_2 are the two retinal images, k is the illumination compensation function given by $k(x,y) = \alpha_{15}y^4 + \alpha_{14}y^3x + \cdots + \alpha_2 y + \alpha_1$. Eq. (2) is minimized in the least squares sense to estimate the 15 parameters. In Fig. 2 we show an example of two retinal images and the compensation function $k(x,y)$. The different shades of gray indicate the average contrast and intensity difference between the two original images in Figs. 2(a) and (b).

3. DETECTION OF LONGITUDINAL CHANGES IN RETINAL IMAGES

One of the main concerns of ophthalmologists when they visually compare fundus images of the same retina over time is to identify true structural or morphological changes pertaining to possible pathological damage. In the same inspection they must disregard other changes merely caused by variation of illumination or blur. A correct assessment of a patient's state evolution requires sharp images obtained on a regular time basis. However, this is not always guaranteed and is the main motivation for developing preprocessing techniques as the ones described in the previous section. Image registration is another preprocessing technique necessary for image-based longitudinal change assessment.[12]

In this section we briefly describe a strategy for the identification of areas of structural change in time sequences of retinal images.[20] An initial step in order to identify these changes comes from computing the difference from the two registered images with previous illumination compensation,

$$\Delta I(x,y) = I_1(x,y) - I_2(x,y) \ . \tag{3}$$

An example of a difference image is shown in Fig. 3(c) in absolute value for visualization purposes. The structural changes can now be visualized and detected from the difference image $\Delta I(x,y)$ by taking a statistical significance test, as proposed in Ref. 17. First, structural changes are often associated with a group of pixels, thus the change decision at a given pixel j should be based on a small block of pixels in the neighborhood

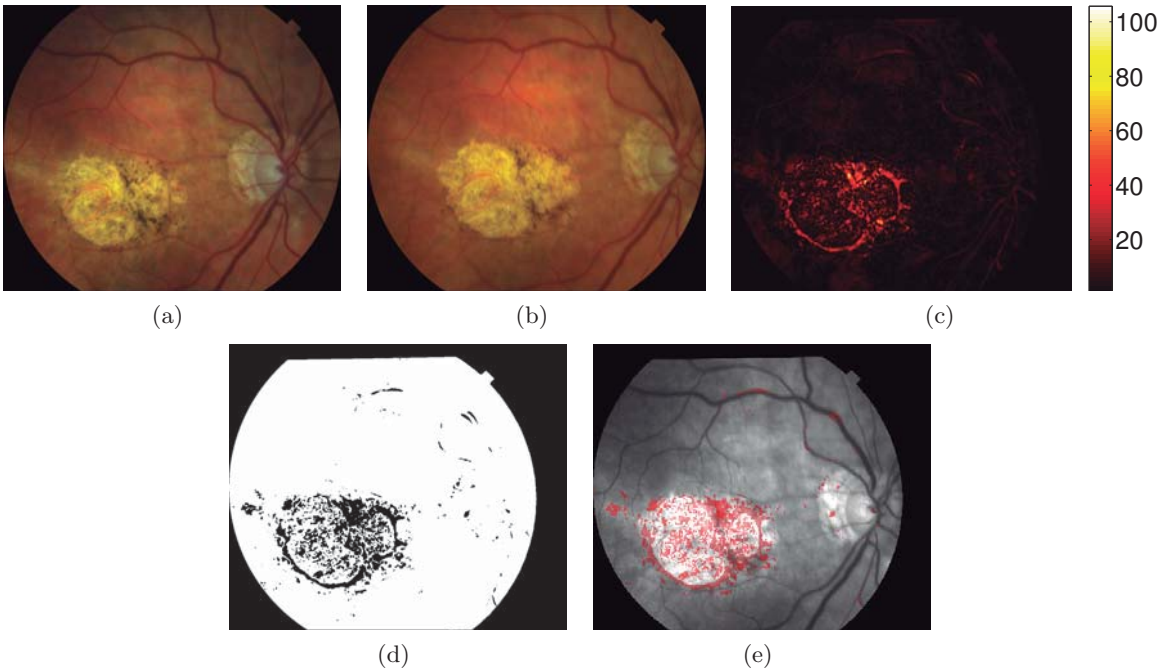


Figure 3. Retinal image change detection: Images (a) I_1 and (b) I_2, (c) Image difference $\Delta I(x, y)$ in absolute value, (d) Image change map, and (d) Image change map in red on top of gray-scale retinal image.

of j denoted as w_j. Second, in the absence of any change, the difference can be assumed to be due to noise alone. Therefore, the decision as to whether or not a change has occurred corresponds to choosing one of two competing hypothesis: the *null hypothesis* $\mathcal{H}_0$ or the *alternative hypothesis* $\mathcal{H}_1$, corresponding to *no-change* and *change* decisions, respectively. Assuming a Gaussian distribution for the difference values, the changes can be identified by comparing the normalized sum square of the differences within the neighborhood w_j to a predetermined threshold τ as described by Aach et al.[21] The test is carried out as below

$$\Omega_j = \frac{1}{\sigma_n^2} \sum_{(x,y)\in w_j} \Delta I(x,y)^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau \ , \tag{4}$$

where σ_n is the noise standard deviation of the difference in the no-change regions. The threshold τ is derived from the fact that Ω_j follows a χ^2 distribution with N degrees of freedom, where N is the number of pixels in the window w_j. It can be obtained for a particular false positive rate α from the χ^2 tables. The image change map resulting from the change detection test with an $\alpha = 0.05$ is shown in Fig 3(d). Notice that the central whitish region (pathological area) is the main cause of structural changes. To better understand this result, in Fig. 3(e) we show one of the retinal images in gray-scale where the pixels related to structural changes are highlighted in red. For further details the reader is referred to Ref. 20.

4. RETINAL IMAGE RESTORATION

In addition to uneven illumination fundus images often suffer from blurring. This hinders diagnosis and the evolution assessment of a disease. In this section we describe a method for fundus image deblurring by means of multichannel blind deconvolution. It consists of a series of preprocessing steps to adjust the images so they comply with the considered degradation model, followed by the estimation of the point spread function, and image deconvolution.

Blind deconvolution consists in the recovery of the original scene from a single or set of blurred images in the presence of a poorly determined or unknown point spread function (PSF).[22] Here we consider multichannel blind deconvolution because it is better posed, as opposed to single-channel, and the PSF is estimated directly

from the degraded images. The restoration strategy is given in Ref. 20. As a regularization term it includes the total variation of the image, which provides good quality of restoration. To properly restore the images the degradation should be adequately modeled.

We assume two registered input images, I_1 and I_2, both originating from an ideal sharp image U

$$\begin{aligned} I_1 &= U * h_1 + n_1 \\ I_2 &= \left(Uk^{-1}\right) * h_2 + n_2 \ , \end{aligned} \tag{5}$$

where $*$ is the standard convolution, h_i are called convolution kernels or PSFs and k is a function accounting for relative local illumination change between images. For pixels where no illumination changes occur $k \approx 1$. The noise n_i is assumed Gaussian additive with zero mean in both images. Despite the fact that we consider the PSFs to vary in time between the two image acquisitions, we assume them to be spatially invariant within each image.

The PSF estimation and image deconvolution algorithm can be viewed as a Bayesian maximum a posteriori estimation of the most probable sharp image and blur kernels. The algorithm is basically the minimization of the functional

$$\arg \min_{U,h_1,h_2} \frac{1}{2}||U * h_1 - I_1||^2 + \frac{1}{2}||U * h_2 - kI_2||^2 + \lambda_u \int |\nabla U| + \lambda_h ||I_1 * h_2 - kI_2 * h_1||^2, \tag{6}$$
$$h_1, h_2 \geq 0 \ ,$$

with respect to the latent image U and blur kernels h_1 and h_2. The first and second terms measure the difference between the input blurred images and the searched image U blurred by kernels h_1 and h_2. The size of this difference is measured by L_2 norm $||.||$ and should be small for the correct solution. Ideally, it should correspond to the noise variance in the given image. Function k compensates for uneven illumination. The two remaining terms are regularization terms with positive weighting constants λ_u and λ_h. The third term is the total variation of U. It improves stability of the minimization and from the statistical viewpoint incorporates prior knowledge about the solution. The last term is a condition linking the PSFs of both images, which also improves the numerical stability of the minimization. For this procedure we set $\lambda_u = 1000$ and $\lambda_h = 10$. The functional is alternately minimized in the subspaces corresponding to the image and the PSFs. The minimization in the PSF subspace is equivalent to the solution of a system of linear equations in the least squares sense with the non-negativity constraint. In the same minimization procedure both the PSFs and the restored image are estimated. If I_1 and I_2 were acquired in a lapse of time, it would be necessary to introduce the structural change detection strategy (Section 3) in both the model of Eq. (5) and the functional given by Eq. (6) (Ref. 20).

An example of a restored retinal image is shown in Fig. 4. In this example the PSF was estimated by Eq. (6), but we have performed deconvolution (restoration) with a single image and with both images to demonstrate the advantages of using multiple images in the restoration as well. From the profile of the original image not much detail can be properly resolved. In contrast there is a noticeable enhancement in both restored images in such a way that much more details are properly resolved. The multichannel deconvolution overcomes the limitations of single-channel deconvolution due to information redundancy. The improvement in resolution is evidenced by gain in contrast and steeper slopes. Notice the small c-shaped blood vessel within the optic disc, and how they are much sharper and properly resolved in the multichannel restoration in Fig. 4 in comparison with the single-channel restoration and the original images. For a detailed examination of this topic see Refs. 18, 20.

5. RETINAL IMAGE QUALITY ASSESSMENT

Image quality evaluation is a limiting factor for automated retinopathy detection.[2] The imaging procedure is typically carried out in two separate stages: image acquisition and diagnostic interpretation. Image quality is subjectively evaluated by the person capturing the images and they can sometimes mistakenly accept a low quality image. Accurate image quality assessment algorithms can allow operators to avoid poor images. Furthermore, a quality metric would permit the automatic submission of only the best images if many were available.

In this section we provide a short description of the work contained in Ref. 23, where we studied the performance of several state-of-the-art no-reference image quality metrics for retinal imaging. Two examples are

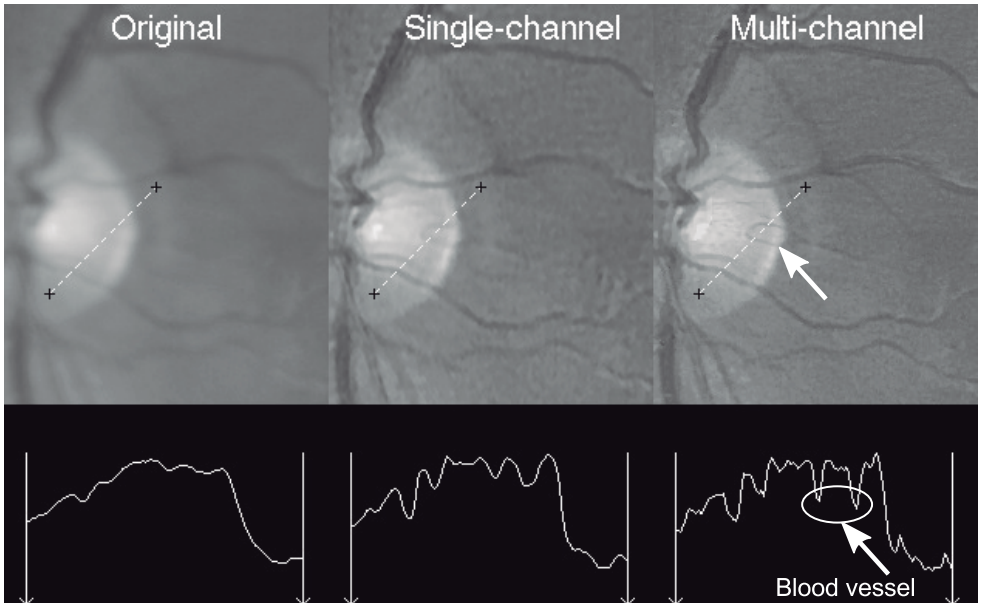


Figure 4. (Top, from left to right) Detail from the original degraded retinal image, the restored version using single-channel deconvolution, and multi-channel deconvolution. (Bottom) Intensity profiles corresponding to the dashed segment. Note how the profiles depict the level of detail in each image.

their application for image quality sorting or image sharpness assessment for focusing. In fact, most no-reference quality assessment methods were initially proposed in the context of focusing applications.[24] The most interesting finding relates to the importance of directional properties with image quality. In other words, the measure of anisotropy as a quality metric. This concept of anisotropy was proposed by two co-authors of this paper (Gabarda and Cristóbal) in Ref. 25 and it represents an important step forward in the area of no-reference quality metrics.

The considered image quality metrics are the following. The first metric Q_1 was proposed by Gabarda and Cristóbal[25] and is based on measuring the variance of the expected entropy of a given image upon a set of predefined directions. The entropy is computed on a local basis using the generalized Rényi entropy and the normalized pseudo-Wigner distribution as an approximation for the probability density function. The authors were able to show that this measure provides a good estimate for the assessment of fidelity and quality in natural images, because their degradations may be seen as a decrease in their directional properties. The second metric Q_2 was proposed by Zhu and Milanfar[26] and it seeks to provide a quantitative measure of –what they call– "true image content". It is correlated with the noise level, sharpness, and intensity contrast manifested in visually salient geometric features such as edges. Q_2 is based upon singular value decomposition of local image gradient matrix. Its value generally drops if the variance of noise rises, and/or if the image content becomes blurry. To avoid regions without edges this algorithm divides the image into small patches and only processes anisotropic ones (non-homogeneous), thus local information is embedded into the final result. The third metric Q_3 was proposed by Ferzli and Karam.[24] It is a sharpness metric designed to be able to predict the relative amount of blurriness in images regardless of their content. Q_3 is conceived on the notion that the human visual system is able to mask blurriness around an edge up to a certain threshold, called the"just noticeable blur" (JNB). It is an edge-based sharpness metric based on a human visual system model that makes use of probability summation over space. JNB can be defined as the minimum amount of perceived blurriness given a contrast higher than the "Just Noticeable Difference". Finally, for the sake of completeness we include the image variance as metric Q_4. This measure has been proven to be monotonic and has a straight-forward relation with image quality for autoregulative illumination intensity algorithms.[27]

In Figs. 5(a)-(b) we show an example of a sharp retinal image and a close-up region for depicting the details of the image. The idea behind this experiment is to determine which quality metric describes better the decreasing image quality. The sharp image was artificially blurred with a 15×15 Gaussian kernel with a varying standard

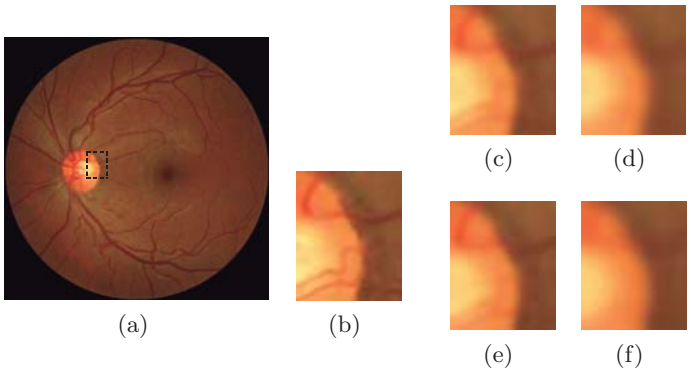


Figure 5. (a) Original sharp retinal image and (b) detail. (c)-(d) details from artificially blurred images with $\sigma = 1.5$ and $\sigma = 3$, respectively. (e)-(f) detail from images with different degrees of focus 3 and 6, respectively (See Fig. 6).

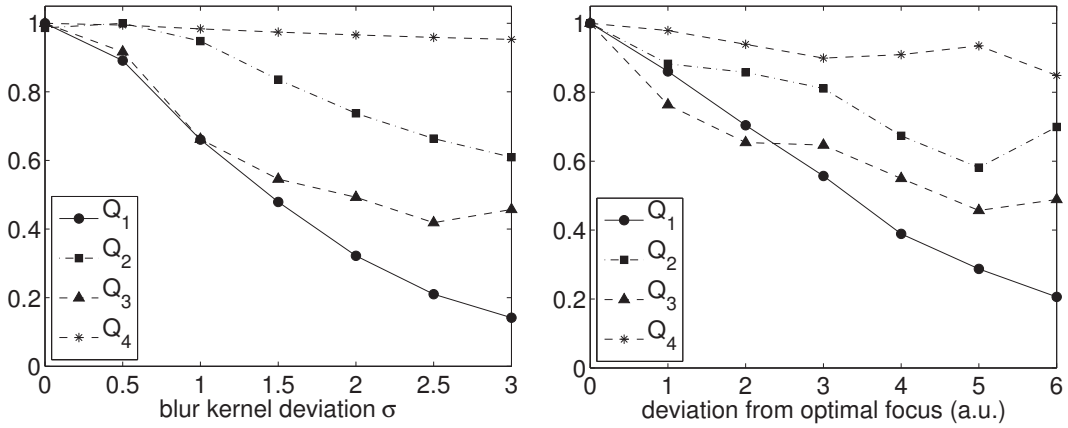


Figure 6. No-reference metrics for assessing image sharpness in relative value. (a) Fundus image artificially blurred with a 15×15 gaussian kernel with varying σ. (b) Fundus images corresponding to the same eye in (a) but with different degrees of fine focus acquired with the retinal camera.

deviation σ. Figs. 5(c)-(d) show the close-up region for the blurred images with $\sigma = 1.5$ and $\sigma = 3$, respectively. The increase in blurriness hinders the resolution of fine structures, hence the medical use is affected as well. In Figs. 5(e)-(f) we show the experimental close-up regions from degraded out of focus images acquired from the same eye fundus.

In Fig. 6(a) we show the resulting quality metrics for the artificial blurring. The figure clearly reveals the overall monotonic nature of all metrics, however Q_1 is the only metric that rapidly decreases with respect to increase in blurriness. The results from the experimental images are shown in Fig. 6(b). Notice how Q_1 also behaves in a consistent way with respect to the deviation from optimal focus. The other metrics seem to be reliable for a considerable amount of blurriness. One possible explanation for the discrepancy between the artificial and real blur for the metrics Q_{2-4} is that the overall illumination distribution cannot be exactly the same, moreover it is also non-uniform. If the metric is not conceived for variations in illumination –even if they were small– it is prone to produce an unreliable result, whereas in Q_1 the intensity normalization provides a certain level of robustness to this type of variation.

In Ref. 23 these experiments led us to conclude that, even though all metrics proved to decrease with the increase in blurriness, strict monotonic decrease was only appreciable for Q_1. This lends strong support for the design of image sharpness metrics based on a directional measure of image content.

6. CONCLUSIONS AND PERSPECTIVES

As we have seen the application of digital image processing techniques for medical image analysis, in this case retinal images, is not only extremely beneficial but can also prove to be effective and cost-efficient for disease management, diagnosis, screening, etc. The increasing need for early detection and screening, along with the ever increasing costs of health care, are likely to be the driving force for the rapid adoption and translation of research findings into clinical practice. The direction of progress, in the short and mid term in this field, is generally conceived within two scenarios: 1) when the amount of data to be analyzed by the medical specialist is excessively large and 2) when the analysis is complex and requires quantification, as opposed to the more qualitative nature of the human expert.

ACKNOWLEDGMENT

This research has been partly funded by the Spanish Ministerio de Ciencia e Innovación y Fondos FEDER (project DPI2009-08879) and projects TEC2010-09834-E and TEC2010-20307. Financial support was also provided by the Czech Ministry of Education under the project 1M0572 (Research Center DAR). The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

REFERENCES

[1] de Mul, M., de Bont, A., and Berg, M., "IT-supported skill-mix change and standardisation in integrated eye-care: lessons from two screening projects in The Netherlands," *International journal of integrated care* **7**(2) (2007).

[2] Abramoff, M. D., Garvin, M., and Sonka, M., "Retinal imaging and image analysis," *IEEE Rev Biomed Eng* **3**, 169–208 (2010).

[3] Saine, P. and Tyler, M., [*Ophthalmic photography: retinal photography, angiography, and electronic imaging*], Butterworth-Heinemann, 2nd ed. (2004).

[4] Bernardes, R., Serranho, P., and Lobo, C., "Digital ocular fundus imaging: A review," *Ophthalmologica* **226**, 161–181 (2011).

[5] Fujita, H., Uchiyama, Y., Nakagawa, T., Fukuoka, D., Hatanaka, Y., Hara, T., Lee, G., Hayashi, Y., Ikedo, Y., Gao, X., and Zhou, X., "Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs," *Computer Methods and Programs in Biomedicine* **92**(3), 238–248 (2008).

[6] Doi, K., "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," *Comput Med Imag Grap* **31**(4-5), 198–211 (2007).

[7] Lu, S., Liu, J., Lim, J., Zhang, Z., and Meng, T., "Automatic fundus image classification for computer-aided diagonsis.," *31st Annual Int. Conf. of the IEEE EMBS* (2009).

[8] Muramatsu, C., Hayashi, Y., Sawada, A., Hatanaka, Y., Hara, T., Yamamoto, T., and Fujita, H., "Detection of retinal nerve fiber layer defects on retinal fundus images for early diagnosis of glaucoma," *J. Biomed. Opt.* **15**(1), 016021 (2010).

[9] Hatanaka, Y., Nakagawa, T., Hayashi, Y., Mizukusa, Y., Fujita, A., Kakogawa, M., Kawase, K., Hara, T., and Fujita, H., "CAD scheme for detection of hemorrhages and exudates in ocular fundus images," *Medical Imaging 2007: Computer-Aided Diagnosis* **6514**(1), 65142M–8 (2007).

[10] Sánchez, C. I., Niemeijer, M., Dumitrescu, A. V., Suttorp-Schulten, M. S. A., Abramoff, M. D., and van Ginneken, B., "Evaluation of a computer-aided diagnosis system for diabetic retinopathy screening on public data," *Investigative Ophthalmology & Visual Science* **52**(7), 4866–4871 (2011).

[11] Cuadros, J. and Bresnick, G., "EyePACS: an adaptable telemedicine system for diabetic retinopathy screening," *Journal of diabetes science and technology* **3**(3), 509–516 (2009).

[12] Winder, R., Morrow, P., McRitchie, I., Bailie, J., and Hart, P., "Algorithms for digital image processing in diabetic retinopathy," *Comput Med Imag Grap* **33**(8), 608–622 (2009).

[13] Feng, P., Pan, Y., Wei, B., Jin, W., and Mi, D., "Enhancing retinal image by the contourlet transform," *Pattern Recognition Letters* **28**(4), 516 – 522 (2007).

[14] Salem, N. and Nandi, A., "Novel and adaptive contribution of the red channel in pre-processing of colour fundus images," *Journal of the Franklin Institute* **344**(3-4), 243–256 (2007).

[15] Marrugo, A. G. and Millán, M. S., "Retinal image analysis: preprocessing and feature extraction," *Journal of Physics: Conf Series* **274**(1), 012039 (2011).

[16] Foracchia, M., Grisan, E., and Ruggeri, A., "Luminosity and contrast normalization in retinal images," *Med Image Anal* **9**(3), 179–190 (2005).

[17] 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 Trans Biomed Eng* **53**(6), 1084–1098 (2006).

[18] Marrugo, A. G., Sroubek, F., Sorel, M., and Millán, M. S., "Multichannel blind deconvolution in eye fundus imaging," in [*ISABEL '11-Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*], 7:1–7:5, New York, NY, USA (2011).

[19] 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), 093706–093709 (2010).

[20] Marrugo, A. G., Sroubek, F., Sorel, M., and Millán, M. S., "Retinal image restoration by means of blind deconvolution," *J. Biomed. Opt.* **16**(11), 116016 (2011).

[21] Aach, T. and Kaup, A., "Bayesian algorithms for change detection in image sequences using markov random fields," *Signal Processing: Image Communication* **7**(2), 147–160 (1995).

[22] Levin, A., Weiss, Y., Durand, F., and Freeman, W., "Understanding Blind Deconvolution Algorithms," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **33**(12), 2354–2367 (2011).

[23] Marrugo, A. G., Millán, M. S., Cristóbal, G., Gabarda, S., and Abril, H. C., "No-reference quality metrics for eye fundus imaging," *CAIP 2011, Lecture Notes in Computer Science* **6854**, 486–493 (2011).

[24] Ferzli, R. and Karam, L. J., "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans Image Process* **18**(4), 717–28 (2009).

[25] Gabarda, S. and Cristóbal, G., "Blind image quality assessment through anisotropy," *J. Opt. Soc. Am. A* **24**(12), B42–51 (2007).

[26] Zhu, X. and Milanfar, P., "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans Image Process* **19**(12), 3116–3132 (2010).

[27] Qu, Y., Pu, Z., Zhao, H., and Zhao, Y., "Comparison of different quality assessment functions in autoregulative illumination intensity algorithms," *Optical Engineering* **45**, 117201 (2006).

Image Analysis and Optics in Ophthalmology

Lecture Notes of the Iternational Centre of Byocybernetics, Polish Academy of Sciences, Warsaw, (2009).

Image Analysis and Optics in Ophthalmology

María S. Millán and Andrés G. Marrugo

Grup d'Òptica Aplicada i Processament d'Imatge, Departament d'Òptica i Optometria
Univesitat Politècnica de Catalunya, Spain

E-mail: andres.marrugo@upc.edu

Abstract. Digital and optical image processing techniques are found today in all fields of medical science. These techniques are especially relevant to modern ophthalmology, a field heavily dependent on visual data. Retinal images are widely used for diagnostic purposes by ophthalmologists. Here, we briefly review some of the existing techniques for retinal image analysis for pathology and secure identification. A global categorization of current algorithms is discussed with especial emphasis on optic disc segmentation. Finally, we mention the global perspectives in the field.
Keywords – retinal image, medical imaging, eye fundus, optic disc.

1. Image analysis in retinal pathology

When alluding to the eye, the posterior segment, or all regions behind the iris, including the vitreous humor, a clear gel that fills the globe and the retina, is referred to as the fundus. The retina is made up several layers including the photoreceptors, which are responsible for the eye's ability to distinguish light and dark, color and fine details. The retina lies at the rear surface of the globe, and in ophthalmology, is an area of primary concern if visual loss is an issue not traceable to refractive error or problems in the cornea. The normal features of fundus images include the optic disc, the fovea, and the blood vessels (see Fig. 1).

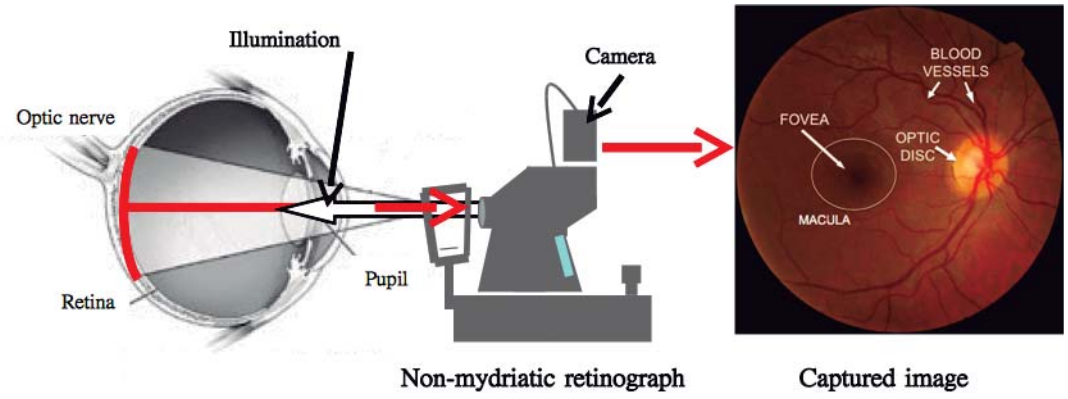


Figure 1. Retinal image.

1.1. Image Acquisition

The eye fundus has been observed by physicians as early as 1850 with the invention of the ophthalmoscope by the German physician Hermann Von Helmholtz [1]. This was an instrument that enabled the examination of the retina by using a bright light near the eye and shining it into the patient's pupil. However, it is not until the mid 1920s that the Carl Zeiss Company made available the first commercial fundus camera. Many were the limitations in clinical use of fundus photography in the 1930s and 1940s which can be attributed to the difficulty in obtaining quality images [1]. From then on, many advances came forth both in quality and specific-purpose developments of techniques for imaging the eye fundus such as: Fluorescent angiography, modern electronic fundus photography, stereo fundus photography, confocal laser opthalmoscopy, among others.

All of the modern systems, including the aforementioned and others like optical coherence tomography (OCT) and scanning laser polarimetry, are mostly oriented towards diagnosis of specific diseases related with certain fundus structures best imaged by a determined technique. In other words, most of these techniques are capable to image the eye fundus under a specific modality, and therefore usually work in a complimentary way. Although, out of all of them it may be said that fundus photography is the one that provides a more general fundus examination with little patient intervention.

To photograph the retina, it is usually required that the iris be dilated with pharmaceuticals and photographed with a mydriatic fundus camera (mydriatic means the iris must be dilated to form an image). A fundus camera (Fig. 1) is an integrated camera system with a circular, axial flash lighting setup integrated directly into the lens. In the last decades, non-mydriatic retinal cameras have been developed. These cameras use near infrared light to focus an a white light arc lamp to image the retina with no need of artificial pupil dilation.The fundus images are captured using a conventional digital camera, attached to the retinal camera body designed to image the eye fundus in association with the optical system of the eye. The digital format provides a permanent record of the appearance of the retina at any point in time [2].

1.2. Algorithms

All of the algorithms that encompass digital image processing techniques provide a feasible way towards developing automated diagnostic systems for conditions such as, diabetic retinopathy, age-related macular degeneration, glaucoma, among others. These diagnostic systems offer the potential to be used in large scale screening programs, with the possibility for resource savings, as well as being free from observer bias and fatigue. In addition, advances in electronic media transmission increase the relevance of using image processing in teleophthalmology as an aid in clinical decision making [3].

Grading retinal images for the presence of pathological features is largely a pattern recognition task [2]. The typical features of diabetic retinopathy are micro-aneurysms, small intra retinal dot hemorrhages, larger blot hemorrhages, all of which are red lesions, and whitish lesions for example lipid exudates, and cotton wool spots which are nerve fiber layer micro-infarcts. The process of detecting these features within a retinal image is made up of a series of operations or steps, with low-level image processing operations providing a basis for higher level analysis [2]. Digital retinal images are usually processed in an algorithmic sequence, with the output of one stage forming the input to the next. This sequence may span from preprocessing techniques for visual enhancement or for further processing and ultimately any sort of feature extraction or segmentation. According to Winder et al. [2] these algorithms can be classified into 5 basic categories:

Preprocessing The main purpose of preprocessing techniques is to reduce the effect of image variation via normalization of the retinal image against a reference model or data set for

subsequent visualization, processing or analysis. Variations may occur within the same image or between images. The main differences between images are most likely due to differences in cameras, illumination, field of view and retinal pigmentation. Thus, in order to extract meaningful information from an image, it is necessary to compensate for this variability. General preprocessing tasks for both monochromatic and color retinal images may be broadly categorized in terms of the correction for non-uniform illumination, contrast enhancement and color normalization [2].

Localization and segmentation of the optic disc The location of the optic disc (OD) is important in retinal image analysis, to locate anatomical structures, for vessel tracking, as a reference length for measuring distances, and for registering changes within the optic disc region due to disease [3]. The segmentation of the optic disc, that is determining the contour of the disc, is a non-trivial problem. However, its shape is an important indicator of various ophthalmic pathologies [3], such as glaucoma [4].

Segmentation of the retinal vasculature The segmentation and measurement of the retinal blood vessels is a key role in diagnosis and treatment of several ophthalmic conditions. Additionally, the segmentation of the vessels is also useful for image registration or spatial alignment of retinal images. Most techniques take advantage of the existing contrast between the vessels and the surrounding background, the cross-sectional intensity-level profile of a typical vessel conforms to a Gaussian shape, the vasculature may be represented by a series of connected line segments that have their origin from the same point (the optic disc). The main techniques used for the segmentation of the vasculature are: matched filtering, morphological processing, vessel tracking, and pixel-based classification.

Localization of the macula and fovea At a certain angular distance from the optic disc location we find the macula, a region that appears darker in color and has no blood vessels present in the center. The fovea is the central part of the macula and is the part of the retina used for sharp vision. Any type of pathological damage in this area is associated with a high risk of visual loss. Due to the poor contrast, and the absence of prominent features, the detection of the macula and fovea is mainly carried out by estimating the position in relation to other retinal structures [5].

Localization and segmentation of retinopathy A great variety of pathological signs are of concern to the specialist who will take them into consideration depending on their relevance as symptoms of a given illness. Thus, for instance diabetic retinopathy is a common complication of diabetes and it implies damage to the microvasculature of the retina. The clinical features are typical and easily recognized, as they consiste mainly of microaneurysms, hemorrhages and lipid exudates. The main task consists in the detection and differentiation of all of these features in the retina. This enables both early detection and the control of the disease progression. Most approaches for retinopathy detection are mainly based on morphological transformations [6], neural networks [7], or other types of machine learning algorithms [2].

2. Examples of retinal image analysis

In this subsection we briefly show two examples of retinal image analysis algorithms with the purpose of further illustrating the aforementioned techniques.

2.1. Optic disc segmentation

Here we show a strategy for OD boundary extraction in ocular fundus images. The interested reader is referred to [8] for a detailed discussion about this method. The preprocessing stage consists in performing color mathematical morphology to remove the blood vessel regions. Subsequently, an active contours approach is used to determine the OD boundary. An active contour is an energy minimizing spline guided by external constraint forces that pull it toward features such as lines and edges [9]. The method is formulated within the Lab color space to take full advantage of the color features available for the preprocessing and feature extraction stages.

Color mathematical morphology Active contour methods generally work by locking onto homogeneous regions of a given image. This task is made extremely difficult since the OD region is fragmented into multiple subregions by blood vessels. Mathematical morphology can extract important shape characteristics and also remove irrelevant information. Here, a closing operation in the Lab color space is carried out, to first remove the blood vessels (dilation) and then approximately restore the boundaries to their former position (erosion).

For each arbitrary point x in the color space, the definitions for dilation (I_d) and erosion (I_e) by structuring element B are:

$$I_d(x) = \{I(y) : I(y) = \max[I(z)], z \in B_x\} \tag{1}$$

$$I_e(x) = \{I(y) : I(y) = \min[I(z)], z \in B_x\}. \tag{2}$$

Optic disc segmentation The optic disc boundary is determined by fitting a geometric active contour model, namely the Chan-Vese model [9]. The performance is evaluated by fitting the active contour onto the optic disc image after closing, then the so obtained contour is superposed on the original image and compared with hand-labeled ground truth. The Chan-Vese model establishes the following energy function in level-set formulation for an image u_0:

$$F(c_1, c_2, \phi) = \int_{\phi>0} |u_0(x,y) - c_1|^2 H(\phi) dxdy + \int_{\phi<0} |u_0(x,y) - c_2|^2 (1 - H(\phi)) dxdy + g(\phi), \tag{3}$$

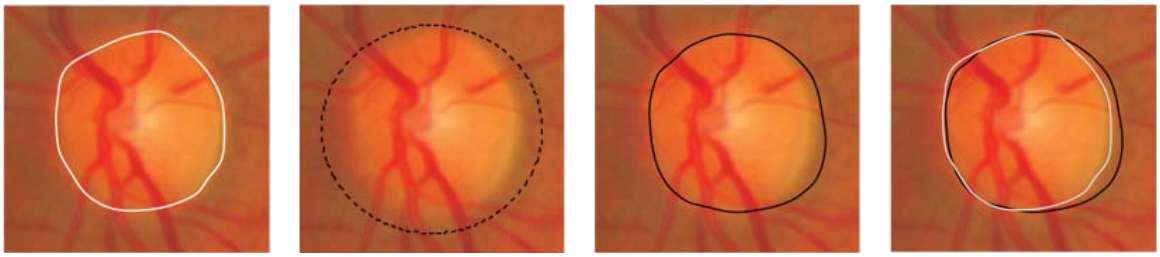
where ϕ is a signed distance function that is zero exactly at the boundary C, that increases in absolute value with respect to the distance from C and that is positive inside and negative outside. c_1 and c_2 represent the average intensity value of u_0 inside and outside the curve, respectively. If we minimize F with respect to ϕ and parameterizing the descent direction by an artificial time $t > 0$ (number of iterations) we obtain:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - (u_0 - c_1)^2 - (u_0 - c_2)^2 \right], \tag{4}$$

where $\delta(.)$ is the Dirac function (see [9] for more details). In Figure 2(a)-(d) we show the hand-labeled ground-truth OD, the initial contour, the final contour, and the overlay of the ground-truth with the final contour respectively.

2.2. Secure Identification

There are many situations with high demands for security and identity authentication, such as access to restricted areas. The blood vessel distribution is a unique pattern in each individual, and highly stable in time. It is quite difficult to forge and most common diseases do not change the pattern in a way that its topology is affected. Therefore, much effort has been put into the development of security systems based on the blood vessel distribution as a biometric signal for authentication purposes [10, 11].



(a) (b) (c) (d)

Figure 2. Optic disc segmentation results. (a) hand labeled ground-truth contour, (b) initial contour, (c) final contour, and (d) overlay of (a) and (c).

The authentication via optical security systems based on the blood vessel distribution provides a fast and effective way for secure identification. The acquisition of the eye fundus is carried out in situ and is compared with an encrypted ID tag. The verification is performed in real time by means of an optical system. For more details the reader is referred to [10, 12].

3. Trends and perspectives

Image analysis and optics in ophthalmology is an ever-growing field that has taken full advantage of modern digital systems. In the last two decades new modalities for imaging the eye fundus have appeared and further improvements have been incorporated to the existing techniques. As regards to digital image analysis, new and effective tools for carrying out massive screening programs are highly desirable. Efficient systems can routinely analyze fundus images and eliminate those that do not have any signs of pathology. This will enable the reduction off the workload of ophthalmologists, the early detection or risk of suffering a disease, and to advance in Tele-care and e-Health programs.

It is somewhat early to have fully automated systems for diagnosis. However, in spite of the increasing availability of digital databases, there still exists the need for gold-standards or ground truth information. Finally, one of the most important trends, lies in the work of multi-disciplinary groups that enable to comprehend and address appropriately the problems in the field.

References

[1] Saine P and Tyler M 2004 *Ophthalmic photography: retinal photography, angiography, and electronic imaging* 2nd ed (Butterworth-Heinemann)

[2] Winder R, Morrow P, McRitchie I, Bailie J and Hart P 2009 *Computerized Medical Imaging and Graphics* **33** 608–622

[3] Patton N, Aslam T, MacGillivray T, Deary I, Dhillon B, Eikelboom R, Yogesan K and Constable I 2006 *Progress in Retinal and Eye Research* **25** 99–127

[4] Valencia E, Millán M S and Kotynski R 2006 *AIP Conf. Proc.* **CP860** 290–299

[5] Tobin K, Chaum E, Govindasamy V and Karnowski T 2007 *Medical Imaging, IEEE Transactions on* **26** 1729 –1739

[6] Spencer T, Olson J A, McHardy K C, Sharp P F and Forrester J V 1996 *Computers and Biomedical Research* **29** 284 – 302

[7] Usher D, Dumskyj M, Himaga M, Williamson T H, Nussey S and Boyce J 2004 *Diabet Med* **21** 84–90

[8] Marrugo A G and Millán M S 2010 *Opt. Pura Apl.* **43** 79–86

[9] Chan T and Vese L 2001 *Image Processing, IEEE Transactions on* **10** 266–277

[10] Millán M, Pérez-Cabré E and Javidi B 2006 *Opt. Lett.* **31** 721–723

[11] Ortega M, Penedo M G, Rouco J, Barreira N and Carreira M J 2009 *EURASIP J. Adv. Signal Process.* Article ID 235746

[12] Pérez-Cabré E, Millán M S and Javidi B 2007 *Opt. Express* **15** 15615–15627

Multichannel Blind Deconvolution in Eye Fundus Imaging

ISABEL'11-Proc. of the 4th Int. Symp. on Applied Sciences in Biomedical and Communication Technologies, **7**:1–5, (2011).

Multichannel Blind Deconvolution in Eye Fundus Imaging

Andrés G. Marrugo
Dept. of Optics and Optometry
Universitat Politècnica de Catalunya, Spain
andres.marrugo@upc.edu

Michal Šorel
UTIA Academy of Sciences of the Czech Republic
Prague, Czech Republic
sorel@utia.cas.cz

Filip Šroubek
UTIA Academy of Sciences of the Czech Republic
Prague, Czech Republic
sroubekf@utia.cas.cz

María S. Millán
Dept. of Optics and Optometry
Universitat Politècnica de Catalunya, Spain
millan@oo.upc.edu

ABSTRACT

Eye fundus imaging is vital for modern ophthalmology. Due to the acquisition process, fundus images often suffer from blurring and uneven illumination. This hinders diagnosis and the evolution assessment of a disease. We present a method for fundus image deblurring by means of multichannel blind deconvolution. It consists of a series of preprocessing steps to adjust the images so they comply with the considered degradation model, followed by the estimation of the point spread function, and image deconvolution. Results show that our approach is capable of significant resolution improvement in degraded retinal images.

Categories and Subject Descriptors

I.4 [**Image processing and computer vision**]: Restoration, Enhancement, Applications; J.3 [**Life and medical sciences**]: Health, Medical information systems

Keywords

Blind deconvolution, fundus image, medical image.

1. INTRODUCTION

Digital imaging of the eye fundus is widely used to diagnose and manage ophthalmologic disorders such as diabetic retinopathy and age-related macular degeneration [1]. Fundus images are captured using a digital camera, attached to a retinal camera body designed to image the eye fundus in association with the optical system of the eye. Major source of fundus image quality degradation are eye aberrations, imperfections in the camera optics, and improper camera adjustment. The imaging procedure is usually carried in two separate steps: Acquisition and diagnostic interpretation. Image quality is subjectively evaluated by the person capturing the images and they can sometimes mistakenly accept a low quality image. A recent study in [2] using an automated system for detection of diabetic retinopathy found that from 10000 exams 23% had insufficient image quality.

In this paper we develop a practical strategy for retinal image deblurring. The core of our proposal is based on an image processing technique for restoration called blind deconvolution (BD) [3]. The goal of BD is to recover the original scene from a single or set of blurred images in the presence of a poorly determined or unknown point spread function (PSF). The main assumption is that blur can be described by a convolution of a sharp image with the unknown PSF. There are basically two groups of BD algorithms; one group that uses a single input image (single-image blind deconvolution SBD) and the other multiple images of the same object blurred in a different way (multichannel blind deconvolution MBD). There are many reliable SBD algorithms [4], however most of them require that the blurred image be governed by relatively strong edges, which is not case here. Therefore we have chosen MBD as a suitable strategy for the restoration of blurred retinal images.

The paper is organized in the following way. Initially we describe a general model for the image degradation which encompasses blurring and uneven illumination, subsequently we detail our approach which consists in an extensive preprocessing stage followed by the PSF estimation and image deconvolution. Finally we discuss a result of the image enhancement strategy and compare with a recent state-of-the-art SBD method [5].

2. MATHEMATICAL MODEL OF IMAGE DEGRADATION

We assume two registered input images, z_1 and z_2, both originating from an ideal sharp image u

$$\begin{aligned} z_1 &= u * h_1 + n_1 \\ z_2 &= \left(uk^{-1}\right) * h_2 + n_2 \ , \end{aligned} \tag{1}$$

where $*$ is the standard convolution, h_i are called convolution kernels or point-spread functions (PSFs) and k is a function accounting for relative local illumination change between images z_1 and z_2. For pixels where no illumination changes occur $k \approx 1$. The noise n_i is assumed Gaussian ad-

ISABEL '11, October 26-29, Barcelona, Spain

ditive with zero mean in both images. In our case, the PSFs comprise all radiometric degradations described above except structural changes in the eye. Despite the fact that we consider the PSFs to vary in time between the two image acquisitions, we assume them to be spatially invariant within each image.

3. ALGORITHM

In the input, the algorithm accepts two eye fundus images. The images are processed in the following four steps: 1. Image registration, 2. Compensation of uneven illumination, 3. Segmentation of areas with structural changes, and 4. PSF estimation and Image deconvolution. The individual steps are detailed in the following subsections. For illustration purposes we consider two color fundus images acquired from a patient that suffered from age-related macular degeneration, which we denote hereafter by z_1 and z_2 and their enhanced versions as u_1 and u_2. The images were captured seven months apart from each other and are shown in Fig. 1. They are color 24 bit-depth fundus images of size 1500×1230 digitized in TIFF format. This is a general example were both images do not correspond exactly to the same object field, the illumination distribution across both images is not exactly the same, and there are some structural differences between them given by the pathological development in the macula (centered yellowish region).

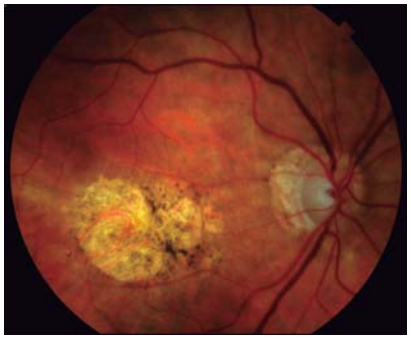
3.1 Image registration

Image registration consists in the spatial alignment of two or more images. Image registration techniques are usually divided into two groups: intensity-based and feature-based methods [6]. Intensity based methods have the drawback of poor performance under varying illumination conditions. Feature based methods are robust to such effects but rely on accurate and repeatable extraction of the features. The retinal vasculature is known to provide a stable set of features for registration for the conditions of interest.

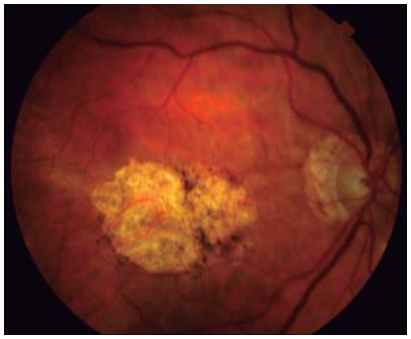
For registering the images we use the robust dual-bootstrap iterative closest point algorithm [7]. The vessel branching and crossover points are used as landmarks to register the images.The registration algorithm starts from initial low-order estimates that are accurate only in small image regions called bootstrap regions. The transformation is then refined using constraints in the region, and the bootstrap region is expanded iteratively. The algorithm stops when the boot-strap region expands to cover the overlap between the images, and uses a 12-dimensional quadratic mapping that accounts for the curvature of the retina. This registration algorithm is very robust to local changes and low overlap between images as demonstrated by its high success rate on test images with at least one common landmark point and overlaps even as low as 35% [7]. The pair of images after registration are shown in Fig. 1(c) in checkerboard representation. For the following subsection, we will consider only the overlapping area of both registered fundus images as a region of interest (ROI).

3.2 Compensation of uneven illumination

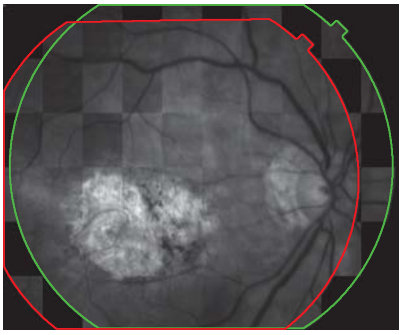
Despite controlled conditions in retinal image acquisition, there are many patient-dependent aspects that are difficult to control and mainly affect the illumination component with gradual non-uniform spatial variations [1]. Some of the contributing factors are: (a) the curved surface of the retina. As a consequence, all regions cannot be illuminated uniformly; (b) Imaging requires either a naturally or an artificially dilated pupil. However, the degree of dilation is highly variable across patients; (c) Unexpected movements of the patients eye and (d) presence of diseases. This effect hinders both quantitative image analysis and the reliable operation of subsequent global operators.



(a) z_1



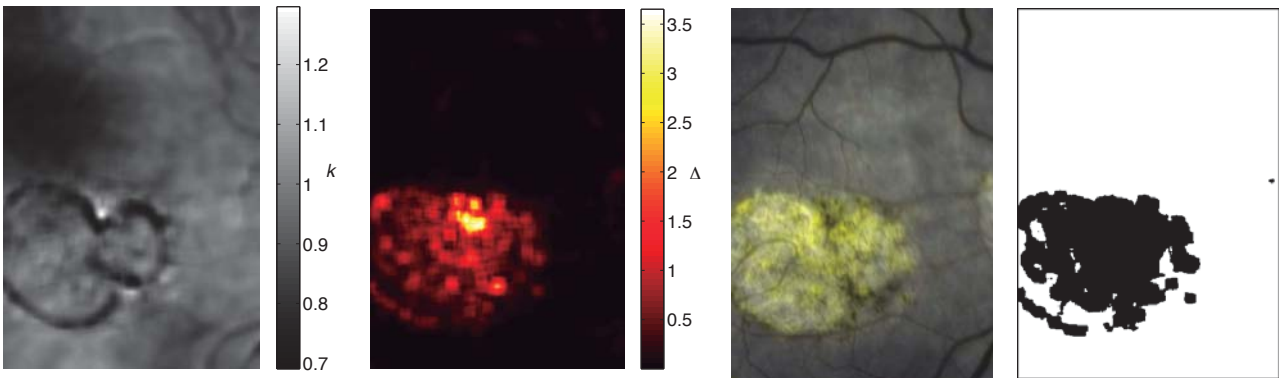
(b) z_2



(c)

Figure 1: Color eye fundus images affected by age-related macular degeneration. (a) Image z_1 was captured seven months prior to (b) image z_2. (c) Registration of z_1 and z_2 in checkerboard representation.

In our model, described by (1), the relative changes in intensity between two fundus images of the same eye cannot be accounted exclusively by convolution with different PSFs and must be compensated. A number of general-purpose techniques have been investigated to attenuate the variation of illumination. However, most techniques, ranging from simple histogram operations to more elaborate models, are oriented toward single-image compensation [8, 9]. Thus, no consistency between a pair of images is guaranteed. This



(a) (b) (c) (d)

Figure 2: Intermediate outputs from the algorithm: (a) illumination compensation function k, (b) local quadratic difference Δ, (c) areas of structural change, and (d) mask for avoiding areas with structural changes.

uneven illumination can be compensated by adjusting the intensity values on one image to approximately match that of the other; i.e. we simply compensate the relative changes of illumination between the images so they meet the requirements from the model. This can be carried out if the blurring is not too large and the illumination changes smoothly, which is usually the case for eye fundus images.

For the illumination compensation we use a sliding window procedure. This provides a simple and effective way to estimate the smooth variation of illumination [8]. To get the compensation function k of our model (1), in the square neighborhood of fixed width W^j of each pixel j we minimize

$$k^j = \arg\min_{\kappa} \sum_{i \in W^j} \left(z_1^i - \kappa z_2^i \right)^2 \quad , \tag{2}$$

where z_1 and z_2 are the registered input images. By differentiating (2) with respect to κ we obtain the expression for k^j:

$$k^j = \frac{\sum_{i \in W^j} z_1^i z_2^i}{\sum_{i \in W^j} {z_2^i}^2} \quad . \tag{3}$$

The interpretation of k from (2) is straightforward. If the registered images z_1 and z_2 had neither illumination changes nor structural changes, then $k \approx 1$ throughout the common object field. The reason why the computation by the sliding window works is that illumination changes in fundus images are smooth and can be considered locally constant, together with the fact that convolution does not change the total image energy. Since the sums in (3) can be computed in an incremental way, the computation of k is very fast, with time complexity proportional to the number of pixels.

The resulting k computed with a window 17 pixels wide is shown in Fig. 2(a). The different shades of gray indicate the average contrast and intensity difference between the two images. From the image it can be seen that most areas have similar intensity values except for the upper left part (dark region), where image z_2 is up to 30 percent darker. This is also noticeable for the bottom-left quadrant (the pathological damaged region) where the differences are not due to illumination variation, but to morphological or structural changes, thus the values have no meaning in terms of illumination compensation.

3.3 Segmentation of structural changes

The pathological region is actually a structural change and cannot be taken as illumination variation. To identify these changes we computed the local quadratic difference from the two images including illumination compensation as:

$$\Delta^j = \sum_{i \in W^j} \left| z_1^i - k z_2^i \right|^2 \quad . \tag{4}$$

We highlight the fact that structural changes can be located with this approach because the illumination varies smoothly, whereas structural changes are often local and appear abruptly. The interpretation of (4) is rather simple, the output is close to zero where the illumination has been adequately compensated and the opposite in areas where the images differ structurally. The result, shown in Fig. 2(b), clearly reveals the existence of regions that after illumination compensation still differ significantly. If these regions appear clustered in the image, one suspects they are probably caused by a source of variation that is not related to illumination variation.These are most likely related to pathological damage which is significant from the clinical viewpoint. To better understand this result, in Fig. 2(c) we show one of the retinal images in gray-scale where the pixels related to structural changes are represented in pseudo-color. This image constitutes an important output of our algorithm. On the other hand, since these changes do not fulfill our convolution model, they should be masked out in order to correctly estimate the PSFs from both images.

In our experiments we applied Otsu's thresholding method [10] to automatically generate a mask from the difference image Δ (Fig. 2(b)). In this way the regions that have structurally changed are are not included in the minimization routine. This is not critical because in practice there is enough information in the remaining pixels to adequately estimate the PSFs. The obtained mask is shown in Fig. 2(d).

3.4 PSF estimation and image deconvolution

For PSF estimation and image deconvolution we have chosen one of the best working MBD methods [11]. Matlab implementation of this method is available on the web of the authors[1]. The algorithm can be viewed as a Bayesian

[1]http://zoi.utia.cas.cz/download

maximum a posteriori estimation of the most probable sharp image and blur kernels. For our purposes, we used a modification of the original method that ignores regions affected by structural changes, which improves stability and precision of the computation. This is similar to the solution proposed in [12] within the super-resolution context. Without this modification, represented by the mask m in (5), the algorithm does not work reliably. The algorithm can be described as a minimization of the functional:

$$\arg\min_{u,h_1,h_2} \frac{1}{2}||u*h_1 - z_1||^2 + \frac{1}{2}||m(u*h_2 - kz_2)||^2 + \lambda_u \int |\nabla u| + \lambda_h ||m(z_1*h_2 - kz_2*h_1)||^2, \quad (5)$$
$$h_1, h_2 \geq 0 \ ,$$

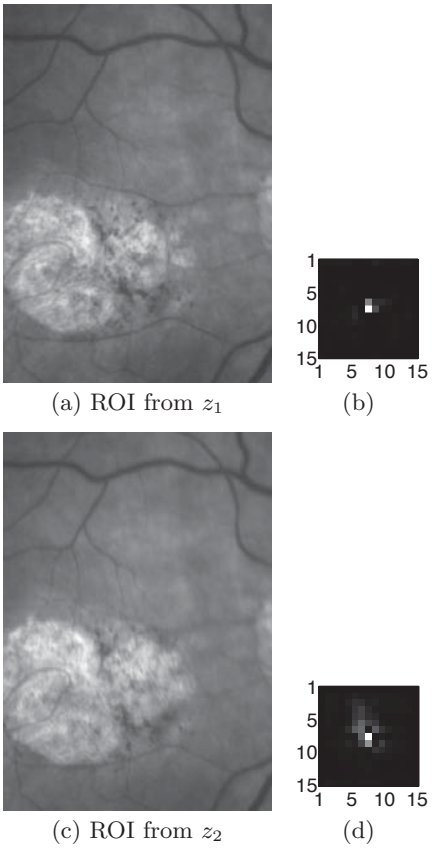
with respect to the latent image u and blur kernels h_1 and h_2. The first and second terms measure the difference between the input blurred images and the searched image u blurred by kernels h_1 and h_2. The size of this difference is measured by L_2 norm $||.||$ and should be small for the correct solution. Ideally, it should correspond to the noise variance in the given image. Function k compensates for uneven illumination. The value of the masking function m is one in the valid pixels (white in Fig. 2(d)) and zero in pixels where the eye fundus is significantly different; in these pixels, we consider only the information from image z_1. The two remaining terms are regularization terms with positive weighting constants λ_u and λ_h. The third term is the total variation of u. It improves stability of the minimization and from the statistical viewpoint incorporates prior knowledge about the solution. The last term is a condition linking the PSFs of both images, which also improves the numerical stability of the minimization. For this procedure we set $\lambda_u = 1000$ and $\lambda_h = 10$.

The functional is alternately minimized in the subspaces corresponding to the image and the PSFs. The minimization in the PSF subspace is equivalent to the solution of a system of linear equations in the least squares sense with the non-negativity constraint (in our implementation solved by Matlab fmincon function). The deconvolution realized by the minimization in the image subspace, is solved by half-quadratic iterative scheme [13], replacing the total variation by $\int \sqrt{|\nabla u|^2 + \epsilon^2}$ in order to achieve smoothness of the functional for zero gradients.

The image deconvolution, i.e. the procedure to obtain the enhanced versions of z_1 and z_2 consists in the following. The minimization of (5) yields u_1, h_1, and h_2. However, u_2 (the enhanced version of z_2) is obtained by minimizing (5) again with fixed PSFs and masking z_1 instead of z_2. The enhanced RGB fundus image is obtained by estimating the PSF from the green channel (being the one with highest contrast) and subsequently deconvolving each channel independently.

4. RESULTS

The ROIs for each image z_1 and z_2, which consist of pathological area plus the unaltered surrounding area and blood vessels, are shown in Fig. 3. Both images are slightly blurred which hinders the possibility to properly resolve details. The PSFs estimated from the minimization of (5) are shown in figures 3(b) and 3(d). The difference in size of the PSFs indicates that one image is more blurred than the other. In Fig. 4(b) we show the enhanced version of z_2 by our method.



(a) ROI from z_1 (b)

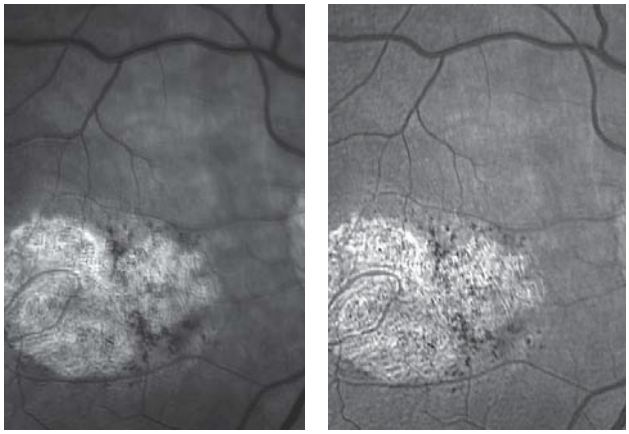
(c) ROI from z_2 (d)

Figure 3: ROIs of retinal images (a) z_1 and (c) z_2. Estimated PSFs in (b) and (d), respectively. The enhanced version of z_2 is shown in Fig. 4(b).

There is a significant improvement in resolution evidenced by the clear definition of the thinner blood vessels and the small pathological structures in the center. To corroborate our assumption that MBD methods are better suited for this type of images, we tried to restore z_2 with a recent SBD method proposed in [5]. The result is shown in Fig. 4(a) and reveals that this method does not follow the true nature of the blurring, thus it is prone to produce a poor restoration. None of the aforementioned fine structures are properly resolved in this image. Finally, we show the color restored versions of z_1 and z_2 in Fig. 5. The enhancement is much more noticeable in color.

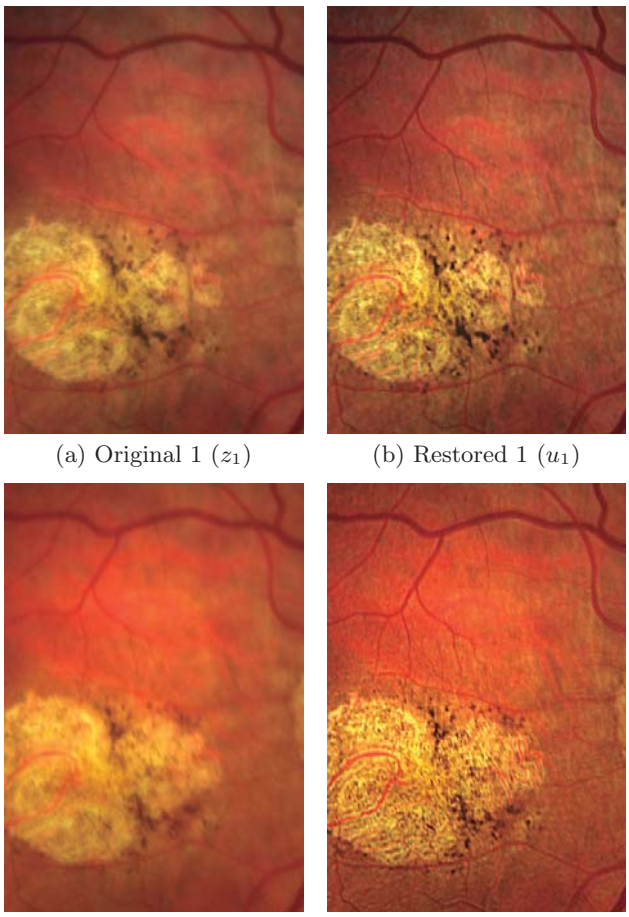
5. CONCLUSIONS

In this paper we have presented a new approach for eye fundus image deblurring based on MBD. We have verified, that a pair of retinal images, belonging to the same eye, contain enough common information to be restored with the proposed method. The method consists of a series of preprocessing steps to adjust the images so they comply with the convolutional model, followed by the final stages of PSF estimation and deconvolution. This approach leads to the improvement in resolution of degraded retinal images, which we have also compared with one of the most sophisticated SBD algorithms [5]. While the initial findings are promising, further tests are necessary.



(a) (b)

Figure 4: Enhanced version of z_2 with (a) SBD method of [5] and (b) proposed method. There is a significant resolution improvement in (b), compare with the original in Fig. 3(c).



(a) Original 1 (z_1) (b) Restored 1 (u_1)

(c) Original 2 (z_2) (d) Restored 2 (u_2)

Figure 5: Original and restored color retinal images. The enhancement is much more noticeable in color.

6. ACKNOWLEDGMENT

This research has been partly funded by the Spanish Ministerio de Ciencia e Innovación y Fondos FEDER (project DPI2009-08879). Financial support was also provided by the Czech Ministry of Education under the project 1M0572 (Research Center DAR). Authors are grateful to the ophthalmologist Jordi Monés (MD) from the Institut de la Màcula i la Retina (Barcelona) for providing us the images. The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

7. REFERENCES

[1] R. Winder, P. Morrow, I. McRitchie, J. Bailie, and P. Hart, "Algorithms for digital image processing in diabetic retinopathy," *Comput Med Imag Grap*, vol. 33, no. 8, pp. 608–622, 2009.

[2] M. D. Abràmoff, M. Niemeijer, M. S. A. Suttorp-Schulten, M. A. Viergever, S. R. Russell, and B. van Ginneken, "Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes," *Diabetes Care*, vol. 31, no. 2, pp. 193–198, 2008.

[3] E. P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*. CRC Press, 1996.

[4] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1964 –1971, 2009.

[5] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Computer Vision – ECCV 2010*, ser. LNCS, K. Daniilidis, P. Maragos, and N. Paragios, Eds. Springer Berlin - Heidelberg, vol. 6311, pp. 157–170, 2010,.

[6] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image Vision Comput*, vol. 11, no. 21, pp. 977–1000, 2003.

[7] C. Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE T Med Imaging*, vol. 22, no. 11, pp. 1379–1394, 2003.

[8] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med Image Anal*, vol. 9, no. 3, pp. 179–190, 2005.

[9] A.G. Marrugo and M.S. Millán, "Retinal image analysis: preprocessing and feature extraction," *Journal of Physics: Conf. Series*, vol. 274, pp. 012039, 2011.

[10] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE T Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[11] F. Šroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," *IEEE T Image Process*, vol. 14, no. 7, pp. 874–883, 2005.

[12] F. Šroubek, J. Flusser, and M. Šorel, "Superresolution and blind deconvolution of video," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1–4, 2008.

[13] A. Chambolle and P. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.*, vol. 76, no. 2, pp. 167–188, 1997.

No-reference Quality Metrics for Eye Fundus Imaging

CAIP'11: Proc. 14th Int. Conf. on Computer Analysis of Images and Patterns, Lecture Notes in Computer Science, **6854**, 486–493, (2011).

No-reference Quality Metrics for Eye Fundus Imaging

Andrés G. Marrugo[1], María S. Millán[1], Gabriel Cristóbal[2], Salvador Gabarda[2], and Héctor C. Abril[1]

[1] Group of Applied Optics and Image Processing, Department of Optics and Optometry, Universitat Politècnica de Catalunya, Terrassa, Spain.
{andres.marrugo,hector.abril}@upc.edu, millan@oo.upc.edu
[2] Instituto de Óptica "Daza de Valdés" (CSIC), Serrano 121, Madrid 28006, Spain.
{gabriel,salvador}@optica.csic.es

Abstract. This paper presents a comparative study on the use of no-reference quality metrics for eye fundus imaging. We center on auto-focusing and quality assessment as key applications for the correct operation of a fundus imaging system. Four state-of-the-art no-reference metrics were selected for the study. From these, a metric based of Rényi anisotropy yielded the best performance in both auto-focusing and quality assessment.

Keywords: No-reference metrics, fundus image, image quality.

1 Introduction

Eye fundus imaging is an integral part of modern ophthalmology, and as so it can truly benefit from emerging methods for image content estimation and quality assessment. In this paper we present a preliminary study on the use of no-reference measures of image content in fundus imaging. We have chosen four state-of-the-art no-reference metrics that have been recently introduced. In the following sections we discuss the applicability of these metrics in two different aspects of fundus imaging: auto-focusing and image quality assessment. In fact, most of the no-reference quality assessment methods were initially proposed in the context of autofocusing applications [1]. These two aspects play a crucial role in the correct operation of a fundus imaging system, which at present day are still chiefly performed by human operation.

2 No-reference metrics

No-reference assessment of image content is, perhaps, one of the most difficult – yet conceptually simple– problems in the field of image analysis [2]. It is only until recently that several authors have proposed no-reference metrics in an attempt to shed some light on this uncertain problem. We have considered four metrics to apply them in fundus imaging. The first metric Q_1 was proposed by Gabarda

and Cristóbal [3] and is based on measuring the variance of the expected entropy of a given image upon a set of predefined directions. The entropy is computed on a local basis using the generalized Rényi entropy and the normalized pseudo-Wigner distribution as an approximation for the probability density function. Therefore, a pixel-by-pixel entropy can be computed, and histograms as well. The Rényi entropy associated to a pixel n in an image can be computed as:

$$R[n] = -\frac{1}{2}\log_2\left(\sum_{k=1}^{N}\breve{P}_n^3[k]\right) \quad , \tag{1}$$

where N is the size of the spatial window used, and $\breve{P}$ is the normalized probability distribution. We can now compute an entropy value for any given orientation θ_i to obtain $R[n,\theta_i]$. The expected value for the whole image is calculated as:

$$\bar{R}[\theta_i] = \sum_n R[n,\theta_i]/M \quad , \tag{2}$$

where M is the image size. And finally the standard deviation from the expected entropy for K orientations –the metric itself– is computed as:

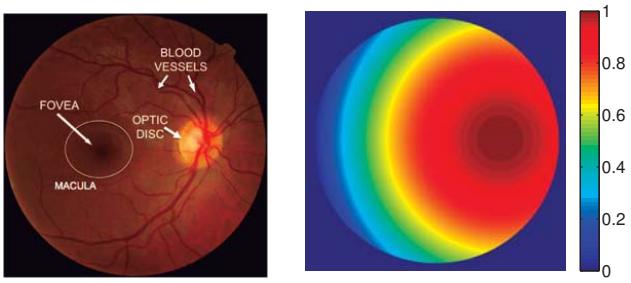
$$Q_1 = \left(\sum_{i=1}^{K}\left(\mu - \bar{R}[\theta_i]\right)^2/K\right)^{1/2} \quad , \tag{3}$$

where μ is the mean of $\bar{R}[\theta_s]$. Q_1 is a good indicator of anisotropy and the authors were able to show that this measure provides a good estimate for the assessment of fidelity and quality in natural images, because their degradations may be seen as a decrease in their directional properties. This directional dependency is also true for fundus images, especially due to blurring or uneven illumination

A drawback of Q_1 is that it requires uniform degradation across the whole image. However, here we show that the use of domain knowledge for retinal imaging provides a means to adjust the metric so as to meet local quality requirements. For this case it would imply to multiply every $R[n,\theta_s]$ by a weighting function $w[n] \in [0,1]$ such that some specific areas are given more importance,

$$\bar{R}[\theta_i] = \sum_n R[n,\theta_i]w[n]/M \quad . \tag{4}$$

This yields a modified Q_1'. Considering two of the most relevant features of a fundus image, the optic disc (OD) and the blood vessels, we have designed a weighting function that takes this fact into account. It is known that in order to assess image sharpness, specialists fixate on the surroundings of the OD to visualize the small blood vessels [4]. The weighting function used is an elliptic paraboloid centered at the OD with values ranging from one exactly at the position of the OD to approximately zero very near the periphery. This function has also been used to model the illumination distribution in fundus images. The approximate position of the OD is determined via template matching [5]. The spatial distribution of the weighting function is shown in Fig. 1(b).



(a) (b)

Fig. 1. (a) Normal fundus image. (b) Weighting function $w[n]$ described by an elliptic paraboloid centered at the OD.

The second metric Q_2 was recently proposed by Zhu and Milanfar [6] and it seeks to provide a quantitative measure of –what they call– "true image content". It is correlated with the noise level, sharpness, and intensity contrast manifested in visually salient geometric features such as edges. Q_2 is based upon singular value decomposition of local image gradient matrix. Its value generally drops if the variance of noise rises, and/or if the image content becomes blurry. To avoid regions without edges this algorithm divides the image into small patches and only processes anisotropic ones (non-homogeneous), thus local information is embedded into the final result.

The third metric Q_3 was proposed by Ferzli and Karam [1]. It is a sharpness metric designed to be able to predict the relative amount of blurriness in images regardless of their content. Q_3 is conceived based on the notion that the human visual system is able to mask blurriness around an edge up to a certain threshold, called the"just noticeable blur" (JNB). It is an edge-based sharpness metric based on a human visual system model that makes use of probability summation over space. JNB can be defined as the minimum amount of perceived blurriness given a contrast higher than the "Just Noticeable Difference". The probability of blur detection (P_{blur}) at an edge given a contrast C can be modeled as a psychometric function given by:

$$Q_3 = P_{blur} = P(e_i) = 1 - \exp\left(-\left|w(e_i)/w_{JNB}(e_i)\right|^{\beta}\right) \quad , \tag{5}$$

where $w_{JNB}(e_i)$ is the JNB edge width which depends on the local contrast C, $w(e_i)$ is the measured width of the edge e_i inside a small patch of the image and β is a fitting constant with a median value of 3.6. Finally, for the sake of completeness we include the image variance as metric Q_4 defined as:

$$Q_4 = \sum_n \left(I[n] - \bar{g}\right)^2 \quad , \tag{6}$$

where $I[n]$ indicates the gray level of pixel n, and $\bar{g}$ the gray mean of the image. This measure has been proven to be monotonic and has a straight-forward relation with image quality for autoregulative illumination intensity algorithms [7].

3 Experimental details

All images were acquired using a digital fundus camera system (TRC-NW6S, Topcon, Tokyo Japan) with a Fuji FinePix S2 Pro camera, with an image resolution of 1152×768. The images were digitized in color RGB of 24 bit-depth in TIFF format without compression. In all figures the images are shown in color, however all metrics were computed using only the luminance channel (Y) of the YUV color space as usual in image quality assessment. From Fig. 1(a) it is evident that the region of interest of the image is that of an approximately circular shaped area that corresponds to the captured object field. The remaining black pixels are not of interest, thus all metrics have been modified to solely include pixels within the circular region of interest in the calculation. The neighboring pixels of the sharp black edge are also left aside from all calculations.

4 Fundus auto-focusing

In fundus photography, the task of fine focusing the image is demanding and lack of focus is quite often the cause of suboptimal photographs [4]. Autofocus algorithms have arisen from the possibility that digital technology offers to continuously assess the sharpness of an image an indicate when the best focus has been achieved. Any given focus measure should be in principle monotonic with respect to blur and robust to noise.

The first experiment we carried out was to observe the behavior of the considered metrics with artificially blurred fundus images (Fig. 2(c)-(d)). Notice the detail from the sharp image and how the fine structures are properly resolved. The increase in blurriness hinders this level of detail, thus the medical use as well. In Fig. 3(a) the original sharp image (Fig. 2(a)) was convolved with a 15×15 Gaussian kernel with a varying standard deviation σ. All metrics are in relative value. The figure clearly reveals the overall monotonic nature of all metrics, however Q_1 is the only metric that rapidly decreases with respect to increase in blurriness.

To validate experimentally this result we captured a series of fundus images from an optimal position of focus to the end of the fine focus capability of the retinal camera (Fig. 2(e)-(f)). The fine focus knob of the retinal camera is operated manually and is able to compensate over a range of $-13 \sim +12D$. Fig. 3(b) shows the relative values for all metrics for seven images with increasing levels of blurriness. Notice how Q_1 also behaves in a consistent way with respect to the deviation from optimal focus. The other metrics seem to be reliable for a small amount of blurriness. One possible explanation for the discrepancy between the artificial and real blur for the metrics Q_{2-4} is that the overall illumination distribution cannot be exactly the same, moreover it is also non-uniform. If the metric is not conceived for variations in illumination –even if they are small– it might be prone to produce an unreliable measure. The algorithm Q_1 is based on a normalized space-frequency representation of the image and not in the image-levels statistics, hence it is robust against illumination changes. In addition, we

have adjusted the metric to meet the local quality requirements by means of a spatially-variant weighting function defined after the geometry of the problem. Similar results have been obtained for other fundus images (10>), but are not reported here for a matter of space.

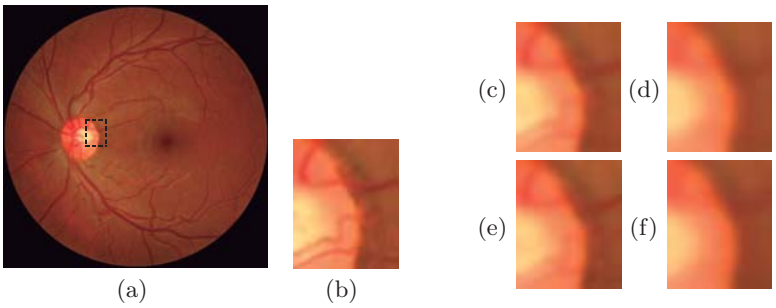


Fig. 2. (a) Original sharp fundus image and (b) detail. (c)-(d) details from artificially blurred images with σ of 1.5 and 3, respectively. (e)-(f) detail from images with different degrees of focus 3 and 6, respectively (See Fig. 3).

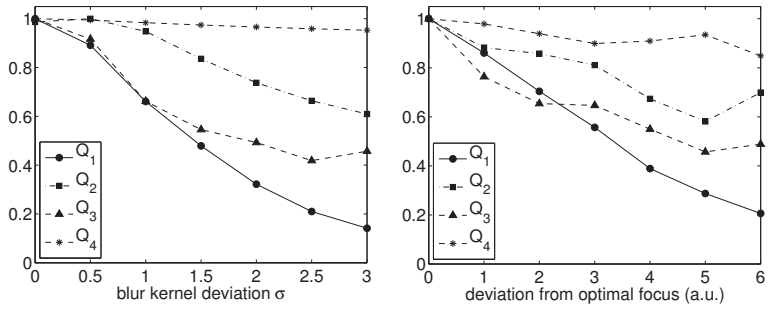


Fig. 3. No-reference metrics for assessing optimal focus in relative value. (a) Fundus image artificially blurred with a 15×15 gaussian kernel with varying σ. (b) Fundus images corresponding to the same eye in (a) but with different degrees of fine focus acquired with the retinal camera.

5 Fundus image quality assessment

Initial image quality is a limiting factor for automated retinopathy detection [8]. The imaging procedure is usually carried out in two separate steps: image acquisition and diagnostic interpretation. Image quality is subjectively evaluated by the person capturing the images and they can sometimes mistakenly accept a low quality image [9]. A recent study by Abràmoff et al. [10] using an automated system for detection of diabetic retinopathy found that from 10 000 exams 23%

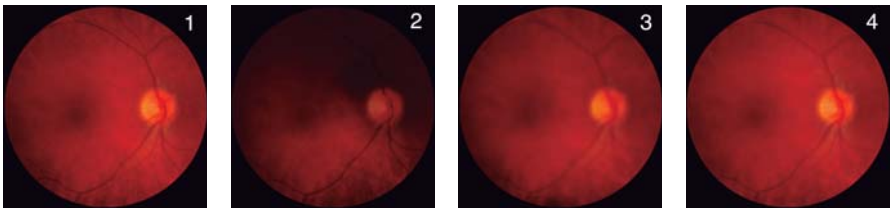


Fig. 4. Fundus images with varying degree of quality corresponding to the same eye.

had insufficient image quality. Accurate quality assessment algorithms can allow operators to avoid poor images. Furthermore, a quality metric would permit the automatic submission of only the best images if many are available. It is from this point of view on that no-reference metrics can be truly useful.

It is often the case that for a given patient several fundus images are acquired. A multilevel quality estimation algorithm at the first few levels has to determine if the images correspond to fundus images, if they are properly illuminated, etc; in other words, if they meet some minimum quality and content requirements. This is in some way what the operator does, he acquires the image and then decides to accept it or not by rapidly visualizing a downscaled version of the image. Once several images of acceptable quality pass this first filter (human or machine), the system would need a final no-reference metric to decide which image to save or to send for further diagnostic interpretation. This metric should in principle yield the sharpest image, with less noise and with the most uniform illumination as possible.

Here we seek to elucidate the possible use of the no-reference metrics for fundus image quality assessment. For this experiment we have analyzed a set of 20 fundus images divided in 5 subsets of 4 images corresponding to the same eye and acquired within the same session. All images within each subset have a varying degree of quality similar to the first subset shown in Fig. 4. Our purpose is to attempt to organize this set from the best image down to the worse. The relative values from all the metrics applied to this set are shown in Table 1. Notice the value Q_1' for image 2. This image is in focus, however it suffers from uneven illumination. Q_1' puts more emphasis on the retinal structures, which are well defined in spite of the illumination, hence the increase with respect to Q_1. Illumination problems are less difficult to compensate as opposed to blurring [11]. This is in line with the specialist's evaluation of the images.

To validate the results two optometrists were recruited as readers. They were familiarized with fundus images and were asked to examine the whole set of images (4 per subject). They evaluated each subset and organized the images from the best to the worse in terms of sharpness and visibility of retinal structures. The relative scores of the metrics are converted to sorting or permutation indexes so as to compare with the quality sorting carried out by the readers (Table 2). Note that in this case only Q_1 and Q_1' agree entirely with the readers. To quantify the agreement we devised a similarity score based on the Spearman's footrule [12]. It is basically the l_1-norm of the difference between the reference permutation π_r (from the reader) and the metric Q permutation π_q. Given a set U of m elements

Image	Q_1	Q_1'	Q_2	Q_3	Q_4
1	1.00	1.00	1.00	0.91	1.00
2	**0.67**	**0.90**	0.40	1.00	0.81
3	0.10	0.12	0.54	0.81	0.85
4	0.38	0.38	0.79	0.70	0.96

Table 1. Relative values for all the metrics applied to the set of images in Fig. 4.

A	B	Q_1	Q_1'	Q_2	Q_3	Q_4
1	**1**	1	1	1	2	1
2	**2**	2	2	4	1	4
4	**4**	4	4	3	3	3
3	**3**	3	3	2	4	2

Table 2. Reader A and B vs. metric sorting of images from Fig. 4 in accordance to quality. Top to bottom: best to worse.

(images), a permutation π of this set is defined as a set of indexes mapping to U to produce a particular order of the elements, $\pi : \{1, \cdots, m\} \rightarrow \{1, \cdots, m\}$. The similarity score S of two permutations π_r and π_q is defined as:

$$S = 1 - \frac{\sum_{i=1}^{m} |\pi_r(i) - \pi_q(i)|}{p_{\max}} \ , \tag{7}$$

where $p_{\max}$ is the maximum value of the numerator. It occurs when the permutations are reversed and it can be shown that $p_{\max}$ is equal to $m^2/2$ when m is even and $(m^2 - 1)/2$ when m is odd. Perfect agreement means $S = 1$, and the opposite $S = 0$. The inter-reader agreement for the whole set of 20 images yielded a S score of 0.90. The S scores for the first 4 image subset and the whole set of images are shown in Table 3. The difference in the overall scores for both readers is practically negligible. It is also clear that Q_1 outperforms the other metrics in this experiment with agreement scores of 0.8 and 0.9. The most probable reason is the computation of the metric from normalized space-frequency representation of the image.

6 Conclusions

We have considered four state-of-the-art metrics and their applicability for eye fundus imaging. For fundus auto-focusing, all metrics proved to decrease with respect to the deviation from optimal focus, however strict monotonic decrease was only appreciable for the metric Q_1, based on a directional measure of Rényi entropy. This is most likely due to its robustness to illumination variation. As far as image quality assessment is concerned, we showed that from the considered metrics Q_1 and its modified version Q_1' are the most reliable in terms of agreement with expert assessment, evidenced by average similarity scores of 0.8 and 0.9 with readers A and B, respectively. The results lend strong support to the development of a no-reference metric for fundus imaging based on Rényi entropy.

Acknowledgment

This research has been partly funded by the Spanish Ministerio de Ciencia e Innovación and Fondos FEDER (project DPI2009-08879) and projects TEC2010-

	Q_1	Q_1'	Q_2	Q_3	Q_4
S_A 1st subset	1.00	1.00	0.50	0.50	0.50
S_B 1st subset	1.00	1.00	0.50	0.50	0.50
S_A all images	0.80	0.80	0.55	0.55	0.40
S_B all images	0.90	0.90	0.60	0.65	0.45

Table 3. Evaluation of the no-reference metrics w.r.t. reader grading with the use of the similarity score S in (7). The subindex in S indicates reader A or B.

09834-E and TEC2010-20307. The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

References

1. Ferzli, R., Karam, L.J.: A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). IEEE Trans Image Process **18** (2009) 717–28
2. Wang, Z., Bovik, A.: Modern image quality assessment. Synthesis Lectures on Image, Video, and Multimedia Processing **2** (2006) 1–156
3. Gabarda, S., Cristóbal, G.: Blind image quality assessment through anisotropy. J Opt Soc Am A Opt Image Sci Vis **24** (2007) B42–51
4. Moscaritolo, M., Jampel, H., Knezevich, F., Zeimer, R.: An image based autofocusing algorithm for digital fundus photography. IEEE Trans Med Imag **28** (2009) 1703–1707
5. Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E., Kennedy, L.: Optic nerve head segmentation. IEEE Trans Med Imag **23** (2004) 256–264
6. Zhu, X., Milanfar, P.: Automatic parameter selection for denoising algorithms using a no-reference measure of image content. IEEE Trans Image Process **19** (2010) 3116–3132
7. Qu, Y., Pu, Z., Zhao, H., Zhao, Y.: Comparison of different quality assessment functions in autoregulative illumination intensity algorithms. Optical Engineering **45** (2006) 117201
8. Abràmoff, M.D., Garvin, M., Sonka, M.: Retinal imaging and image analysis. IEEE Rev Biomed Eng **3** (2010) 169–208
9. Bartling, H., Wanger, P., Martin, L.: Automated quality evaluation of digital fundus photographs. Acta Ophthalmologica **87** (2009) 643–647
10. Abràmoff, M.D., Niemeijer, M., Suttorp-Schulten, M.S.A., Viergever, M.A., Russell, S.R., van Ginneken, B.: Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes. Diabetes Care **31** (2008) 193–198
11. Marrugo, A.G., Millán, M.S.: Retinal image analysis: preprocessing and feature extraction. Journal of Physics: Conf Series **274** (2011) 012039
12. Fagin, R., Kumar, R., Sivakumar, D.: Comparing top k lists. SIAM J. Discrete Math. **17** (2003) 134–160

Efectos de compresión en imágenes de la retina para la evaluación del riesgo glaucomatoso

IX Reunión Nacional de Óptica, 140, (2009).

Efectos de Compresión en Imágenes de la Retina Para la Evaluación del Riesgo Glaucomatoso

Andrés G. Marrugo[1] * y María S. Millán[1]

[1] *Departamento de Óptica y Optometría. Grupo de Óptica Aplicada y Procesado de Imagen. Universidad Politécnica de Cataluña, Terrassa (Barcelona), España*

http://www.goapi.upc.edu

1. Introducción

Un sistema integrado de telemedicina comprende una infraestructura que se utiliza para capturar, transmitir y almacenar datos e imágenes médicas. Para reducir el coste de dicha infraestructura es necesaria la compresión de la imagen [1]. En los últimos años las dos técnicas de compresión más utilizadas en imágenes médicas son la transformada discreta del coseno (DCT), i.e. Joint Photographics Experts Group (JPEG), y la transformada discreta wavelet, i.e. JPEG2000 [2]. La compresión digital se divide en compresión con pérdidas y sin pérdidas. La desventaja de las técnicas sin pérdidas es que sólo pueden reducir el tamaño de la imagen en un factor de 2 a 3 [2]. Por otra parte, las técnicas con pérdidas proveen niveles de compresión significativamente mayores con cierto nivel de degradación de la calidad. Hasta el momento se han establecido estándares para radiología y patología, mientras que en oftalmología aún no se han establecido los niveles de resolución mínimos requeridos para un diagnóstico fiable [3]. Actualmente existen distintas métricas, objetivas y subjetivas, para evaluar el desempeño de los estándares de compresión frente a un nivel de degradación. Las más importantes son: la calificación media de opinión (MOS), el error cuadrático medio (MSE) y el pico señal a ruido (PSNR). Sin embargo, existen pocos estudios que evaluan el desempeño de algoritmos de extracción de características ante la degradación por compresión [4]. En el área de la telemedicina, existe un gran interés en que el especialista pueda hacer un diagnóstico acertado a pesar de la compresión. Conrath et al. [3] establecieron relaciones de compresión en oftalmología utilizando los estándares JPEG y JPEG2000 evaluando el efecto de la compresión en la detección de la retinopatía diabética.

El glaucoma es una enfermedad caracterizada por la pérdida progresiva de visión periférica. La disminución de fibras nerviosas causada por la hipertensión ocular puede visualizarse en la base del nervio óptico (disco óptico o papila). Dos parámetros comúnmente utilizados para estimar el riesgo glaucomatoso son la relación copa-disco (CDR), en el cual se estima el área ocupada por la excavación (copa) en el disco óptico, y la regla ISNT [5]. Ampliamente utilizada en la práctica clínica, la regla ISNT indica el orden en que normalmente se distribuye el grosor del anillo neuro-retiniano (Inferior≥Superior≥Nasal≥Temporal). En los últimos años ha aumentado el interés por determinar de manera automática o semiautomática el riesgo glaucomatoso a partir de imágenes de la retina [1]. La principal característica de éstos algoritmos consiste en determinar la CDR a partir de la extracción de los contornos del disco óptico y la copa por distintas estrategias, como el analisis por componentes o la transformada watershed [1]. Este trabajo pretende evaluar el desempeño de un algoritmo de inspección para la detección o el seguimiento del riesgo glaucomatoso, desarrollado previamente en nuestro grupo de investigación [6], bajo los efectos de la degradación por distintos niveles de compresión utilizando los estándares JPEG y JPEG2000.

2. Procesado de imagen y extracción de características

Algoritmo. El algoritmo utilizado para la segementación y extracción de características se describe en [6]. El preprocesado de las imágenes consiste en un suavizado del ruido y el realce de contornos

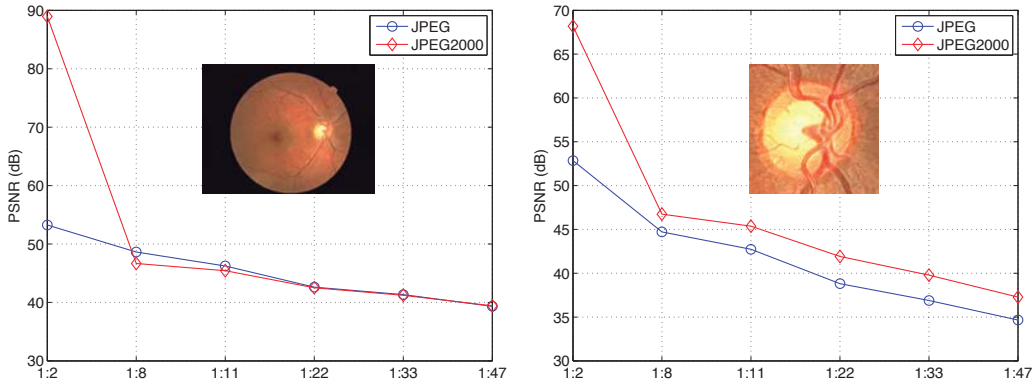
* *e-mail:* andres.marrugo@upc.edu

mediante el operador LoG-vision [7]. Posteriormente se selecciona la región de interés (ROI) alrededor del disco óptico y se realiza una transformación a coordenadas polares para la extracción del contorno del disco óptico. Se calculan diferencias de color ΔE_{00} en todas las direcciones radiales entre píxeles vecinos. Seguidamente se buscan los píxeles con la mayor diferencia de color dentro de una banda de donde se presume se encuentra el contorno del disco óptico. El contorno se extrae a partir de la conversión a coordenadas cartesianas. Para la segmentación de la copa se busca una semilla representativa de la región con coordenadas $(\text{máx}(L), \text{mín}(C), \text{máx}(h))$ dentro del espacio de representación CIELAB. Se realiza una umbralización y se extrae el contorno por un proceso similar al anterior por conversión a coordenadas polares. Por último se calculan la CDR y las relaciones ISNT normalizadas a la sección temporal.

Imágenes de prueba y formatos de compresión. En este trabajo partimos de imágenes digitales del fondo de ojo capturadas con cámara no midriática. Se ha utilizado un retinógrafo Topcon TRC-NW6S y una cámara 3CCD Sony DXC-990P. Las imágenes correspondientes a un campo visual de 30° son digitalizadas a un tamaño de 768x576 píxeles en RGB en profundidad de 24 bits. La región de interés (ROI) ocupa un área aproximada de 100x100 píxeles. Se tomó un conjunto de 6 imágenes y se comprimieron con razones de compresión de 1:2, 1:8, 1:11, 1:22, 1:31 y 1:47, bajo los formatos JPEG clásico y JPEG2000. En total, un conjunto de 72 imágenes. Se utilizaron niveles bajos de compresión (1:2 y 1:8) como referencia y los niveles altos corresponden a los utilizados en un estudio de valoración subjetiva de la retinopatía diabética [3]. Se calculó el PSNR para tener un referente en cuanto a la degradación de la calidad de las imágenes. Se analizaron todas las imágenes y se extrajeron los parámetros CDR e ISNT.

Medida de semejanza. La medida de semejanza se derivó a partir de la l_2-norma de los parámetros CDR e ISNT. Esto es, para cada imagen analizada se obtuvo un conjunto de parámetros que se ordenan en un vector 1-dimensional y que se expresa como $\mathbf{S}^j = \{S_i^j\}$, $j = A, B$, $i = 1, ..., 4$; $S_1^j = \text{CDR}$, $S_2^j = \text{I}$, $S_3^j = \text{S}$, $S_4^j = \text{N}$; donde A y B corresponden a la imagen original y la imagen comprimida, respectivamente. En otras palabras, los parámetros de cada vector corresponden al CDR y los parámetros I, S y N, que se encuentran normalizados a T. La medida de distancia entre estos dos vectores en el espacio de parámetros puede ser calculada como

$$d = |\mathbf{S}^A - \mathbf{S}^B|^2 = \sqrt{\sum_{i=1}^{n} |S_i^A - S_i^B|^2} \tag{1}$$



(a) Imagen completa (fondo de ojo). (b) Región de interés (disco óptico).

Figura 1: Comparación del nivel de distorsión bajo los estándares de compresión JPEG y JPEG2000.

3. Resultados y discusión

En la figura 1(a) se muestra la comparación de los niveles promedio de distorsión para los estándares JPEG y JPEG2000 aplicados a todas las imágenes. El valor más alto de PSNR lo presenta el JPEG2000. Se encontró que para bajos niveles de compresión el JPEG2000 supera al JPEG clásico, mientras que para niveles medios y altos de compresión, en promedio, los dos estándares arrojan valores PSNR muy similares. Adicionalmente, se calcularon niveles promedio de distorsión en PSNR para la ROI de todas las imágenes y se muestra en la figura 1(b). Se observa un comportamiento similar al del análisis general, aunque se destaca una ligera diferencia entre los dos estándares siendo el JPEG2000 el que mantiene valores superiores.

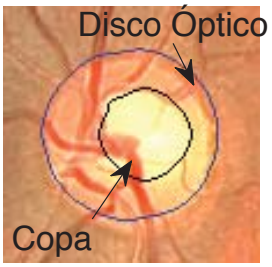


Figura 2: Segementación Copa y Disco, imagen original sin compresión.

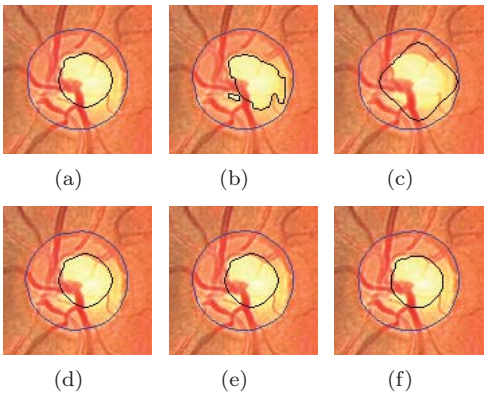


Figura 3: Contornos con JPEG razones de compresión a) 1:2, b) 1:8, c) 1:11 y JPEG-2000 d) 1:2, e) 1:8 y f) 1:11.

En lo que respecta a la extracción de los contornos del disco óptico y la copa, en la figura 2 se muestra un ejemplo de su extracción bajo condiciones normales para una de las imágenes de la retina. En todos los casos se extrajeron los parámetros CDR e ISNT en las imágenes sin comprimir para tomar como referencia. En la figura 3 se muestran los resultados del algoritmo para tres niveles de compresión distintos en los estándares JPEG y JPEG2000. En este caso se observa que para la compresión JPEG los contornos extraídos difieren rápidamente de la referencia al aumentar el nivel de compresión. El efecto es más notorio en la segmentación de la copa. Por lo contrario, los contornos extraídos en imágenes bajo compresión con JPEG2000 parecen mucho menos distorsionados. Las medidas de semejanza para este caso y de valores promedio se observan en las figuras 4(a) y 4(b), respectivamente. De la figura 4 se denota claramente que los parámetros extraídos de las imágenes comprimidas con JPEG2000 son mucho más fiables. Sin embargo, la fiabilidad de las mediciones en relación con la referencia pueden considerarse como óptimas hasta razones de compresión de nivel medio, particularmente entre 1:11 y 1:22 bajo el estándar JPEG2000. Por otra parte, los parámetros extraídos de las imágenes bajo compresión JPEG parecen diferir sustancialmente de la referencia debido a la sensibilidad de un algoritmo guiado por análisis de color. Es conocido que la principal característica de la degradación por compresión JPEG se puede manifestar como mayor riqueza en la textura de una imagen [2]. De igual manera, la introducción de elementos indeseados debido a la partición en bloques de la imagen pueden manifestarse como distorsiones locales que afectan de manera significativa la extracción de contornos.

4. Conclusiones y perspectivas

Se llevó a cabo un estudio sobre el procesado de imágenes de la retina bajo compresión y la extracción de características de relevancia clínica en la detección del riesgo glaucomatoso. Se evaluaron dos tipos de compresión, JPEG y JPEG2000, siendo el segundo el que mostró menor introducción de distorsiones para la extracción de contornos mediante un análisis de color. En ambos casos, la degradación de las imágenes por compresión medida como PSNR mostró estar dentro de los

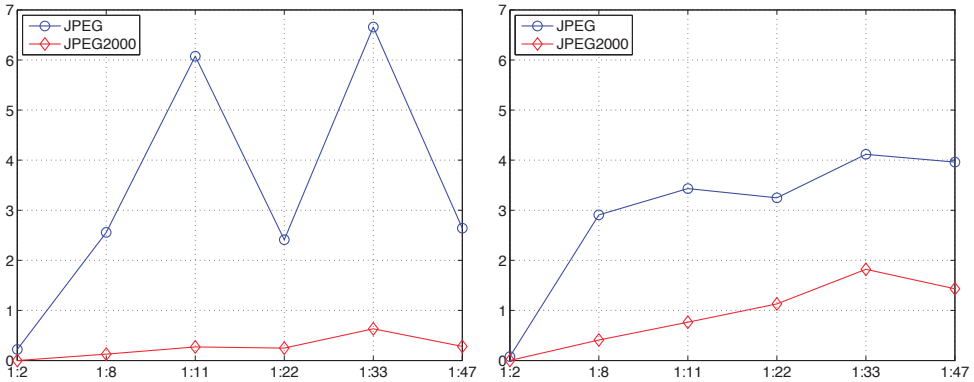


Figura 4: Medida de semejanza en el espacio de los parámetros CDR e ISNT entre la referencia y las imágens comprimidas bajo JPEG y JPEG2000. a) Un caso particular y b) Curva promedio.

valores citados en la literatura como óptimos [2]. Se muestra el JPEG2000 como mejor opción que el JPEG para el procesado de imágenes de la retina bajo compresión y se estima que puede arrojar resultados fiables hasta niveles medios de compresión, entre 1:11 y 1:22. Este trabajo forma parte de un estudio preliminar sobre el procesado de imágenes del fondo de ojo y se extenderá sobre un mayor número de muestras. Adicionalmente, se buscarán posibles modificaciones al algoritmo de segmentación para mejorar su desempeño ante el procesado de imágenes comprimidas.

Agradecimientos: Los autores agradecen la colaboración de E. Valencia y el apoyo económico del Ministerio de Ciencia e Innovación así como de los fondos FEDER al proyecto DPI2006-05479. Andrés Marrugo agradece a la Universidad Politécnica de Cataluña una beca predoctoral.

Bibliografía

[1] N. Patton, et al., "Retinal image analysis: Concepts, applications and potential", Progress in Retinal and Eye Research, **25**, p. 99–127 (2006).

[2] Y. H. Shiao, T. Chen, K. Chuang, C. Lin, C. Chuang, "Quality of Compressed Medical Images" Journal of Digital Imaging, **20**(2): p. 149–159 (2007).

[3] J. Conrath, et al., "Evaluation of the effect of JPEG and JPEG2000 image compression on the detection of diabetic retinopathy", Eye, **21**(4), p. 487–493 (2006).

[4] S. Matschitsch, M. Tschinder, A. Uhl, "Comparison of Compression Algorithms Impact on Iris Recognition Accuracy", Advances in Biometrics, p232–241 (2007).

[5] A. S., Khouri, B. C. Szirth, K. S. Shahid, R. D. Fechtner, "Software-assisted optic nerve assessment for glaucoma tele-screening" Telemed J. E. Health, **14**(3), p. 261–265 (2008).

[6] E. Valencia, and M.S. Millán. "Color image analysis of the optic disc to assist diagnosis of glaucoma risk and evolution" *Proceedings of the 3rd IS&T European Conference on Colour in Graphics, Imaging, and Vision (CGIV '06)* Leeds, UK (2006).

[7] M.S. Millán, E. Valencia, "Color image sharpening inspired by human vision models", Applied Optics, **45** p. 7684-7697 (2006).

Implementation of an Image Based Focusing Algorithm for Retinal Imaging

X Reunión Nacional de Óptica, 40–43, (2012).

Implementation of an Image Based Focusing Algorithm for Retinal Imaging

Andrés G. Marrugo[1]*, María S. Millán[1], and Héctor C. Abril[1]

[1]Departamento de Óptica y Optometría, Universidad Politécnica de Cataluña, Violinista Vellsolà, 37, 08222 Terrassa (España).

andres.marrugo@upc.edu

Abstract: Retinal photography is important for the assessment of eye diseases. The task of fine focusing the image is demanding and lack of focus is often the cause of suboptimal photographs. The advent of digital cameras has provided the opportunity to automate the focusing process. In this work, we propose an auto-focus system for non-mydriatic retinal imaging. The core of the system is based on a robust image-based focus measure. The measure is basically a quantification of image anisotropy computed by means of the normalized discrete cosine transform. Additionally, we optimize the autofocusing method by evaluating different focus search strategies. Encouraging experimental results reveal that the method is able to identify the best focus reliably with optimal speed.

1. Introduction

A fundus imaging device or retinal camera is a specialized low-power microscope with an attached camera designed to photograph the retina in association with the optical system of the eye. Retinal imaging is an important tool for both detection and monitoring the progression of diseases affecting the eye [1]. Retinal cameras can be mydriatic or non-mydriatic. Mydriatic cameras require pharmacological dilation, while non-mydriatic cameras use a near infrared viewing system to exploit the patient's natural dilation in a dark room. Infrared light is used to preview the retina on a video monitor. Once the monitor's image is focused and aligned, a flash of visible light from a Xenon arc lamp is fired and the image is captured. Non-mydriatic retinal cameras are equipped with a focusing mechanism that consists in displacing a compensation lens. It is basically an aspheric objective lens design that, when combined with the optics of the eye, matches the image plane to the eye fundus. The focus control of the fundus camera is used to compensate for refractive errors in the subject's eye. Until recently [2], these cameras were entirely operated manually with the focusing mechanism assisted by a split line visual aid. Manual focusing is error prone especially in the presence of inexperienced photographers and may lead to images that require additional restoration or enhancement [3]. The autofocus feature offered in new retinal cameras is a significant advance that ultimately leads to a more robust imaging system, especially for medical screening purposes. On the other hand, the autofocus feature still relies on the split line mechanism, whereas in this work we propose to implement a passive auto-focus (AF) measure completely based on image analysis. For further details on fundus imaging the reader is referred to [1].

This research covers the practical issues in implementing and optimizing an image-based AF algorithm for non-mydriatic retinal imaging. Each of the parameters used for focusing, such as the focus measure, the searching algorithm, or the focusing window, may be customized for optimal performance. In the following sections we present the image-based focus measure along with several focus search algorithms. In Section 4 we discuss on the performance and practical issues in implementing the focus measure. And finally, we present the results and conclusions.

2. Focus measure

The focus measure computes the sharpness or degree of focus on a region of an image, typically called the focusing window. The degree of focus changes as the lens is moved through different positions. Auto-focusing means automatically moving the lens such that

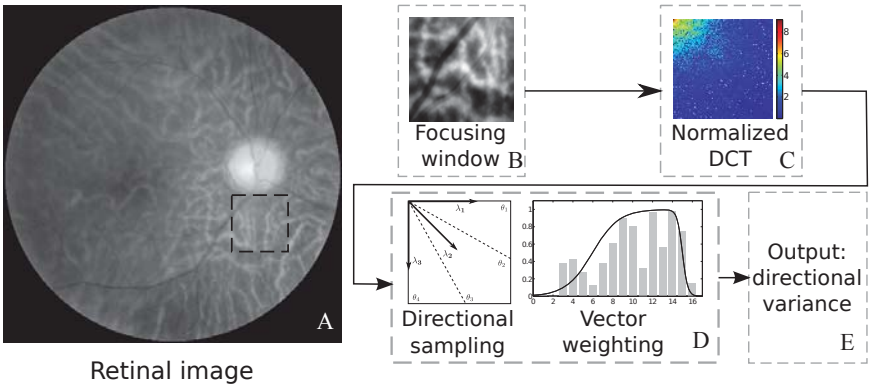


Figure 1: Focus measure algorithm flowchart: A the infra-red retinal image is acquired; B a small focusing window is selected that includes retinal structures; C the normalized DCT is computed; D the DCT is sampled in predefined directions and these coefficient distributions are weighted accordingly; E the focus measure is the variance from these directions.

the sharpness is maximized, i.e. the image is in the optimal focus. In [4] we proposed a new robust focus measure for non-mydriatic retinal imaging. A detailed description of the algorithm is given in[4] and is outlined below.

A description of the algorithm is shown in Figure 1. The focus measure is based on the directional variance of the normalized discrete cosine transform (DCT). First, the focusing window is selected such that there are retinal structures within. Failing to provide this requirement leads to unreliable measurements. Second, the normalized DCT is computed from the focusing window. The normalized DCT is described in [4] and it guarantees invariance to contrast. The third stage consists in calculating a weighted directional sampling on the normalized DCT and finally the focus measure is the variance from all considered directions. This measure is basically an anisotropy measure with robustness given by the weighting scheme designed specifically to achieve invariance to illumination (low coefficients) and noise (high coefficients).

3. Search Algorithms

The focus search strategies suggested in [5] were evaluated. In this section we describe the global search, Fibonacci search, and hill-climbing search methods. The global search strategy goes through every possible lens position in a unidirectional manner. The lens position with the maximum focus value is identified as the optimal focus lens position. It guarantees to find the optimal focus position, however it is the slowest strategy because it has to calculate the focus value for all possible lens positions. A two-step search is a fast version of the global search strategy. During the first search, it performs a global search with a coarse step size. In the second search, the neighborhood around the optimal focus position found in the first search is searched with a fine step size. Although it improves the search speed, it does not guarantee focus accuracy.

The Fibonacci search strategy is based on continuously narrowing the search region by subdividing it according to the Fibonacci sequence. It is basically a divide and conquer process. The required number of iterations for this search is given by $n = \max\{x : S_x < U\}$, where S_x is the x^{th} number on the Fibonacci sequence and U is the number of lens positions. For 256 lens positions, the Fibonacci number just below is $S_{13} = 233$ therefore the search will require at most 13 steps to find the focused image.

The Hill climbing search strategy determines the direction of the next lens movement by the gradient of the previous two consecutive focus values. The movement step is determined by a parameter L. When the direction of the lens movement is reversed, L is reduced to one half of its original value for a finer search. This can be done several times until the optimal focus position is obtained. Although the hill-climbing search strategy can start at

any lens position, it it sensitive to image noise, and its performance depends on the chosen parameter values.

4. Performance issues and experimental details

A proper AF procedure should be accurate and fast. The speed of the focusing procedure depends on the speed of the camera's focusing motor, the speed of the focus measure, and the number of steps required by the search strategy. For this study we consider the speeds of the motor and the focus measure as fixed parameters. Therefore, we evaluate the search strategies described in Section 3 to determine the best performing strategy for the AF procedure. This was done by considering the following three metrics: 1) the number of AF iterations (or lens movements) required for the camera to reach the optimal focus position; 2) the number of motor steps the lens has traveled during the AF process; and 3) the difference (or error) between the final lens position at the end of the AF process and the real optimal focus position. From the three metrics, the first two determine the speed of the AF procedure. Fewer AF iterations and fewer motor steps lead to a faster AF. Fewer motor steps are also indication of a smaller chance of having back-and-forth lens movements. As regards the last metric, a smaller error leads to a sharper image. A good AF method should perform well against the three metrics.

The experimental set-up consisted mainly of an AF mechanism attached to a digital fundus camera system (TRC-NW6S, Topcon, Tokyo Japan). The AF apparatus consisted of in-house assembled stepper motor mechanism for the displacement of the compensation lens controlled via RS232 with a PC. This mechanism was coupled to the fundus camera. The image acquisition and processing, along with the motor control, were carried out in MATLAB. The images were acquired from the video output of the infra-red focusing system with a resolution of 640×480. The fundus camera focusing system enables a compensation range of $-13D : 12D$ in normal operation. For strong myopia or hyperopia two additional compensation lenses are available to compensate the ranges $-12D : -33D$ and $+9D : +40D$, respectively. The image sequences analyzed here were acquired for the normal operation range.

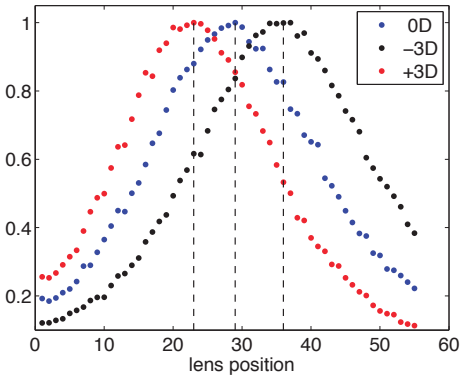


Figure 2: Focus measure validation with a test target. The red and black curves are almost equally shifted versions of the $0D \cdot$ curve, which indicates a successful identification of the optical power needed for compensation.

5. Results

In this section we show the results for the focus measure validation and the evaluated search strategies for the AF procedure on real eye fundi. The curves shown in Figure 2 were computed using a fixed target located at infinity with respect to the fundus camera lens system. To validate the focus measure two lenses of $+3D$ and $-3D$ were placed in front of the fundus camera. The shifts with respect to the $0D \cdot$ curve are nearly the same. This means that the focus measure is able to identify the optimal focus position for the corresponding optical power compensation. In Figure 3 we show a focus curve obtained by focus-

ing the right eye fundus of a 27-year-old subject for testing several focus search algorithm. The curve is monotonic with respect to blur and has a unique global maximum. As described in Section 3, we tested four search algorithms. In Figure 3(a)-(b) we show the best performing algorithms: Fibonacci and Hill-climb, respectively. From our tests the Hill-climb method typically requires a reduced number of AF iterations with short motor step movements, and is accurate enough for the considered application. The Fibonacci method is sufficiently accurate but is generally slower because of long motor step movements.

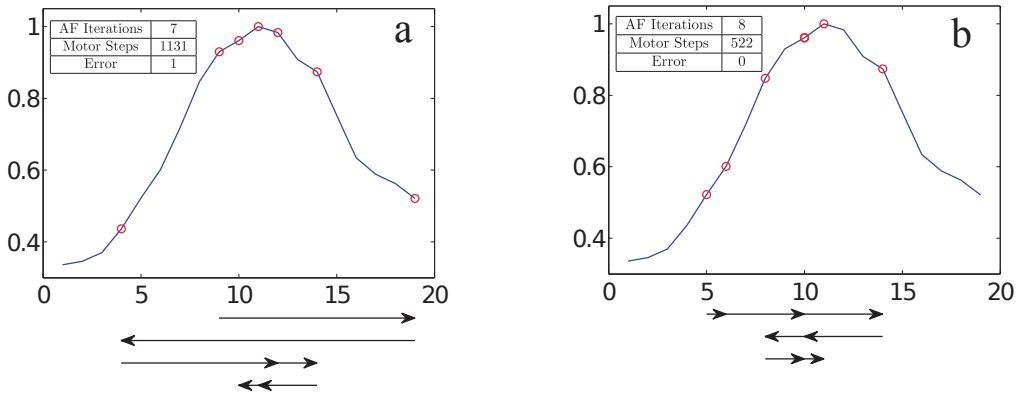


Figure 3: Focus search algorithms: a) Fibonacci and b) Hill-climb. (Arrows) Lens movements, (red circles) lens positions.

6. Conclusions

In this work we have presented the implementation of a new image-based focus measure for non-mydriatic retinal imaging. We have tested the focus measure for correct optical power compensation. The measure faithfully reproduces defocus with the proper displacement of the curves maxima. In addition, we tested several focus search strategies and found the Hill-climb method suitable for this application. Not only is it accurate enough, but it performs sufficiently fast, something truly required for an AF algorithm.

Acknowledgment

This research has been partly funded by the Spanish Ministerio de Ciencia e Innovación and Fondos FEDER (project DPI2009-08879). The first author also thanks the Spanish Ministerio de Educación for an FPU doctoral scholarship.

References

[1] R. Bernardes, P. Serranho, and C. Lobo, "Digital Ocular Fundus Imaging: A Review," Ophthalmologica, vol. 226, pp. 161–181, 2011.

[2] M. Moscaritolo, H. Jampel, F. Knezevich, and R. Zeimer, "An Image Based Auto-Focusing Algorithm for Digital Fundus Photography," Medical Imaging, IEEE Transactions on, vol. 28, no. 11, pp. 1703–1707, 2009.

[3] A. G. Marrugo, M. Sorel, F. Sroubek, and M. S. Millán, "Retinal image restoration by means of blind deconvolution," J. Biomed. Opt., vol. 16, no. 11, p. 116016, 2011.

[4] A. G. Marrugo, M. S. Millán, G. Cristóbal, S. Gabarda, and H. C. Abril, "Anisotropy Based Robust Focus Measure For Non-Mydriatic Retinal Imaging," J. Biomed. Opt. (Submitted), 2012.

[5] Dong-Chen Tsai and H. Chen, "Reciprocal Focus Profile," IEEE Trans. on Image Processing, vol. 21, no. 2, pp. 459–468, 2012.