Universidad de Huelva

Departamento de Ingeniería Electrónica, Sistemas Informáticos y Automática

Fundus digital image processing : automated segmentation of the main retinal anatomical structures

Memoria para optar al grado de doctor
presentada por:

Arturo Aquino Martín

Fecha de lectura: 22 de marzo de 2011

Bajo la dirección de los doctores:

Manuel Emilio Gegúndez Arias y Diego Marín Santos

Huelva, 2011

ISBN: 978-84-15147-57-2
D.L.: H 85 - 2011

PhD Thesis

Fundus Digital Image Processing:

Automated Segmentation of the Main Retinal Anatomical Structures

Procesamiento de Imágenes Digitales de Fondo de Ojo:

Segmentación Automática de las Principales Estructuras Anatómicas de la Retina

Author: Arturo Aquino Martín

Directors: Manuel E. Gegúndez Arias
Diego Marín Santos

February 2011

Universidad de Huelva

Departamento de Ingeniería Electrónica, de Sistemas Informáticos y Automática

Escuela Técnica Superior de Ingeniería

Fundus Digital Image Processing: Automated Segmentation of the Main Retinal Anatomical Structures

Procesamiento de Imágenes Digitales de Fondo de Ojo: Segmentación Automática de las Principales Estructuras Anatómicas de la Retina

Tesis Doctoral de

Arturo Aquino Martín

Dirigida por

Manuel E. Gegúndez Arias y Diego Marín Santos

UNIVERSIDAD DE HUELVA

2010

D. Manuel Emilio Gegúndez Arias, Profesor Titular de Escuela Universitaria de la Escuela Técnica Superior de Ingeniería de la Universidad de Huelva, y D. Diego Marín Santos, Profesor Contratado Doctor de la Escuela Técnica Superior de Ingeniería de la Universidad de Huelva,

CERTIFICAN:

Que D. Arturo Aquino Martín, Ingeniero en Informática, ha realizado bajo nuestra dirección y dentro del programa de doctorado **Investigación, desarrollo e innovación industrial** el trabajo de investigación correspondiente a su Tesis Doctoral titulado:

Fundus Digital Image Processing: Automated Segmentation of the Main Retinal Anatomical Structures

Procesamiento de Imágenes Digitales de Fondo de Ojo: Segmentación Automática de las Principales Estructuras Anatómicas de la Retina

Revisado el presente trabajo, estimamos que puede ser presentado al Tribunal que ha de juzgarlo.

Y para que conste a efectos de lo establecido en el capítulo tercero, artículo vigésimo primero del Real Decreto 56/2005 y por la Normativa Reguladora del título de Doctor de la Universidad de Huelva, autorizo la presentación de este trabajo en la Universidad de Huelva.

Huelva, a 14 de febrero de 2010

Director:	Director:	Doctorando:
Dr. Manuel Emilio Gegúndez Arias	Dr. Diego Marín Santos	Arturo Aquino Martín

Acknowledgement

In first place, I would like to express my deepest thanks to my Ph. D. supervisors, Manuel Emilio Gegúndez and Diego Marín, for the long caring and patient hours been dedicated to this PhD student. The end of this thesis would have not been possible without their constant dedication, and this can only be eternal and fraternally appreciated.

I would also like to make a special mention to my colleagues and friends José Manuel Andújar and José Manuel Bravo. The first one for opening the doors of this University and generously supporting me through the easy and hard times along these years. The second one for being always available for help in either scientific and personal matters, even with no responsibility over me at all.

This doctoral thesis was mostly developed at the University of Huelva, but the contribution of my stay at the Center of Mathematical Morphology of the National High School of Mines in Paris cannot be missed at all. I would like to express my gratitude to the human team I shared five unforgettable months with: Fernand Meyer, Estelle Parra-Denis, Etienne Decenciere, Beatriz Marcotegui, Jesús Angulo and, specially, to honorable and endearing Jean-Claude Klein. To finish my acknowledgments to this French centre, and on a personal level, I would like to affectionately mention Catherine Moysan and Santiago Velasco, two wonderful people I had the pleasure to meet and enjoy.

These acknowledgments also have a very special room for all my friends. To the ones from the Universidad de Huelva, with whom so many lunches and bad times I shared; to the ones

from the Universidad de Sevilla, with whom I lived some of the best years in my universitary stage and, of course, to my friends from Isla Cristina, my lifetime friends, for a whole life of support and love.

Agradecimientos

En primer lugar, querría expresar mi más profundo agradecimiento a mis directores de tesis, Manuel Emilio Gegúndez y Diego Marín, por todas las horas que con cariño y paciencia han dedicado a este doctorando. Sin su constante dedicación a esta tesis no habría sido posible su culminación, y ello no puede ser más que eterna y fraternalmente agradecido.

Mención especial querría hacer también a mis compañeros y amigos José Manuel Andújar y José Manuel Bravo. El primero me abrió las puertas de esta Universidad y me ha prestado generosamente su apoyo en los momentos buenos y de dificultad a lo largo de estos años. El segundo, sin responsabilidad alguna sobre mí, siempre ha estado disponible para ayudarme de forma desinteresada en cuantos asuntos científicos e incluso personales me han acontecido.

Esta tesis doctoral se ha desarrollado en su mayor parte en la Universidad de Huelva, pero en absoluto puede ser pasada por alto la aportación que en ella ha tenido mi estancia en el Centro de Morfología Matemática de la Escuela Nacional Superior de Minas de París. Me gustaría agradecer muy cariñosamente a todo el grupo humano con el que compartí cinco meses inolvidables: Fernand Meyer, Estelle Parra-Denis, Etienne Decenciere, Beatriz Marcotegui, Jesús Angulo y, muy especialmente, al honorable y entrañable Jean-Claude Klein. Para cerrar estos agradecimientos al centro francés, en el plano estrictamente personal, me gustaría mencionar muy afectuosamente a Catherine Moysan y Santigo Velasco, dos personas maravillosas que tuve el placer de conocer y disfrutar.

Estos agradecimientos también guardan un hueco muy especial a todos mis amigos. A los de la Universidad de Huelva, con quienes tantos almuerzos y buenos y malos ratos he compartido, a mis amigos de la Universidad de Sevilla, con quienes viví algunos de mis mejores años en mi etapa universitaria, y cómo no, a mis amigos de toda la vida, a mis amigos de Isla Cristina, a quienes les agradezco toda una vida de apoyo y cariño.

To my dearly beloved parents Arturo and María Teresa and to my adored sister María José, without whose love and faith in me this dream would have been unreachable. This thesis is also dedicated to my dear and forever missed friend Miguel

A mis queridísimos padres Arturo y María Teresa y a mi adorada hermana María José, sin cuyo amor y fe en mí habría sido imposible alcanzar este sueño. Esta tesis también está dedicada a mi por siempre querido y añorado amigo Miguel

Resumen de la Tesis (thesis summary in spanish)

Esta tesis no fue concebida como un trabajo aislado, sino como parte de los trabajos que se están llevando a cabo en el grupo de investigación *Control y Robótica* (código TEP-192) de la Universidad de Huelva. El trabajo que se presenta comenzó en 2007 con el nacimiento del proyecto de investigación *Sistema Experto para la Detección Precoz Automatizada de la Retinopatía Diabética mediante Análisis de Imágenes Digitales de la Retina*, financiado por la Consejería de Salud de la Junta de Andalucía (Andalucía, España). Este proyecto, enmarcado dentro del plan integral para la prevención de la diabetes puesto en marcha por dicho gobierno, tiene como objetivo final la obtención de un sistema para la detección precoz automatizada de la retinopatía diabética mediante el análisis de imágenes digitales de fondo de ojo. Concretamente, el sistema está proyectado que sea un software para ser usado, como herramienta de apoyo, por médicos no especialistas, con en el fin de discriminar y filtrar con mayor confianza casos diagnósticos de retinopatía diabética. Esto es, el futuro sistema filtrará aquellos casos que requieran visita al oftalmólogo para una evaluación más precisa y tratamiento específico de aquellos otros que no, reduciendo de esta manera la carga de trabajo de estos especialistas y, por tanto, decreciendo costes.

En el marco de trabajo descrito anteriormente, se realizó una estancia en el Centro de Morfología Matemática de la Escuela Nacional Superior de Minas de París (Francia). Los principales objetivos de dicha estancia fueron el estudio de la morfología matemática y su aplicación al procesamiento de imágenes digitales de la retina. Para ello, la estancia se dividió

en dos fases: una primera fase de formación en morfología matemática y una segunda fase de inmersión en el trabajo del grupo de investigación de imagen retinal del mencionado centro.

Este resumen está compuesto por siete secciones. La *primera sección* realiza una introducción desde el punto de vista médico de los aspectos más relevantes que ayudan a enmarcar y comprender las investigaciones que se presentan a continuación. Las secciones *segunda, tercera y cuarta* resumen las nuevas metodologías descritas en esta tesis para la segmentación del disco óptico y árbol vascular y para la identificación de la zona macular en imágenes digitales de la retina. La *quinta sección* propone una nueva función de evaluación de calidad de segmentaciones del árbol vascular. La *sexta* sección presenta las principales conclusiones y trabajo futuro de la presente tesis. Por último, en la *séptima* sección se detallan los principales méritos e hitos conseguidos con el presente trabajo.

1 Introducción y motivación

La retinopatía diabética (RD) es una enfermedad crónica que actualmente constituye la primera causa de ceguera en la población en edad laboral de los países desarrollados. Esta patología se caracteriza por un conjunto de daños retinales causados por complicaciones de la diabetes mellitus, una enfermedad que provoca una elevación anormal de la concentración de glucosa en la sangre. Este exceso de glucosa afecta al sistema micro y macro vascular dañando el endotelio de los vasos, lo cual incrementa la permeabilidad de los mismos. Usualmente el ojo, concretamente la retina, es uno de los lugares donde ésto es más evidente. Cuando el sistema microvascular del ojo se daña progresivamente, puede desarrollarse pérdida de visión y, en última instancia, ceguera.

La retina es una lámina sensible a la luz que se encuentra en la parte interna trasera del ojo, cubriendo aproximadamente el 65% de su superficie interior. Las células fotosensibles de la retina, llamadas bastones y conos, convierten la luz incidente en señales, las cuales son transportadas al cerebro para ser procesadas, produciéndose de esta manera el sentido de la visión. Los principales componentes anatómicos de la retina son: el disco óptico, el árbol vascular y la mácula (ver la Figura 1 como complemento a la descripción que se ofrece a continuación). El disco óptico o papila representa la localización donde las fibras nerviosas de la retina se unen y salen del ojo en dirección al cerebro para transportar las señales nerviosas. Las células retinales son nutridas por los vasos capilares retinales. Estos vasos

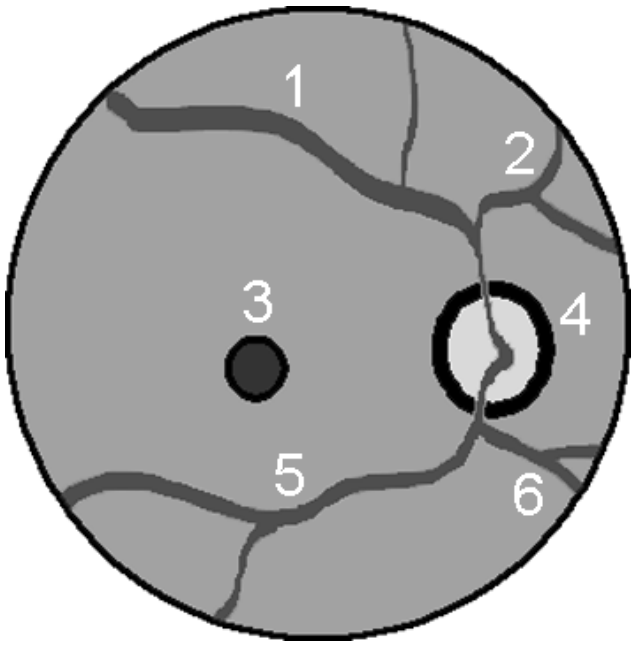


Figura 1: *Diagrama de la superficie retinal: (4) Disco óptico. (3) Centro de la mácula (fóvea). (1), (2), (5) y (6) Árbol vascular retinal: (1) vasos superiores temporales, (2) vasos superiores nasales, (5) vasos inferiores temporales y (6) vasos inferiores nasales.*

forman una estructura vascular en forma de árbol que nace en el disco óptico con algunos vasos gruesos principales. Estos vasos principales se dividen de forma progresiva en otros más pequeños para cubrir la retina. Principalmente, en la estructura vascular se pueden distinguir cuatro tipos de vasos según su curso y ubicación: vasos superiores temporales, vasos superiores nasales, vasos inferiores temporales y vasos inferiores nasales. Por otro lado, las células fotosensibles no se encuentran igualmente distribuidas a lo largo de la retina. La zona periférica de ésta contiene principalmente bastones, lo cual le hace responsable de la visión periférica. El área central de la retina posee una mayor concentración de conos, lo que le confiere la capacidad de la visión central. La mácula es una zona redondeada en la región central de la retina, y presenta una pequeña depresión en su centro llamada fóvea. Ésta se corresponde con la región de la retina con mayor concentración de células fotosensibles. La Figura 2 presenta una imagen digital de la superficie retinal en la que se muestran el disco óptico, la mácula y el árbol vascular.

Las primeras manifestaciones de la RD son diminutas dilataciones capilares conocidas como microaneurismas. Con la progresión de la enfermedad, estos microaneurismas pueden estallar dando lugar a hemorragias. Además del derrame de sangre, los vasos también pueden ocasionar derrames de lípidos y proteínas, causando de esta manera la aparición de manchas

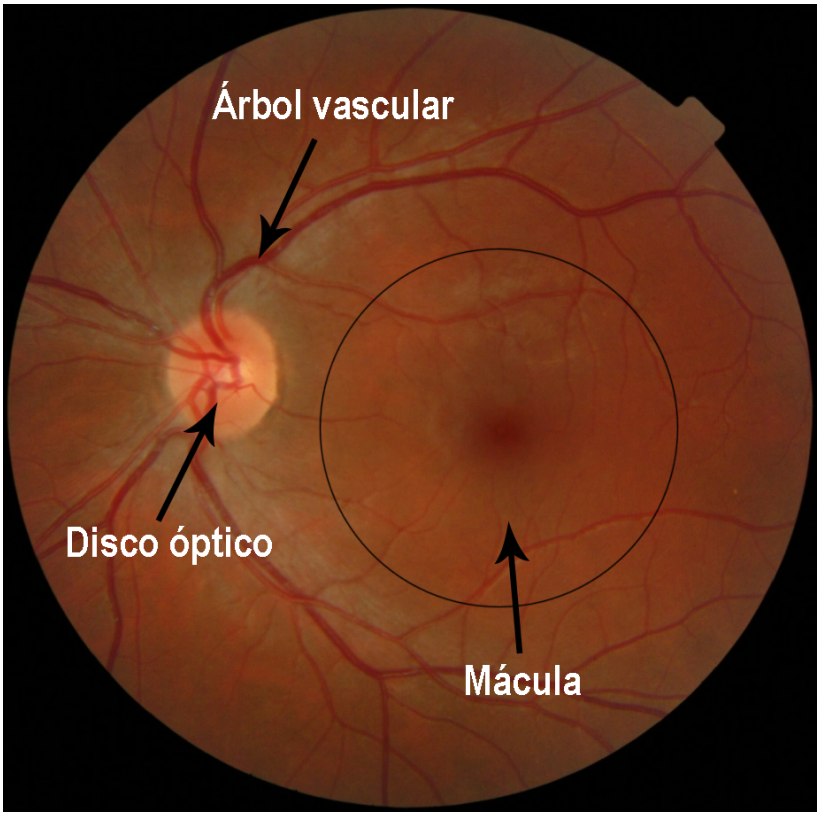


Figura 2: *Imagen digital de la superficie de la retina mostrando el disco óptico, la mácula y el árbol vascular.*

brillantes llamadas exudados (en la Figura 3 se muestra una imagen de la retina donde se ilustra la aparición de microaneurismas, hemorragias y exudados). El derrame de fluido de los vasos puede producir grandes áreas exudativas en la mácula que pueden provocar inflamación de la retina. Esta complicación se denomina edema macular retiniano. Todo este daño puede producir que partes de la retina se vuelvan isquémicas, privadas de sangre. Estas áreas isquémicas son visibles sobre la superficie retinal como gotas de aspecto esponjoso denominadas nubes de algodón (véase la Figura 3). Como respuesta a la aparición de estas áreas isquémicas, comienzan a proliferar en la retina nuevos vasos para suministrar más oxígeno. Éstos, llamados neovascularizaciones, tienen mayor riesgo de ruptura y de provocar hemorragias que los vasos normales. En las últimas etapas de la RD, como consecuencia del gran daño causado a la retina, puede producirse su desprendimiento.

Aunque la RD no es una enfermedad curable, la fotocoagulación láser puede prevenir en gran medida la pérdida de visión si la enfermedad es detectada en sus fases iniciales. Sin embargo, los pacientes con RD no perciben síntomas hasta que la pérdida de visión ya se ha desarrollado, usualmente en las últimas etapas de la enfermedad, cuando el tratamiento

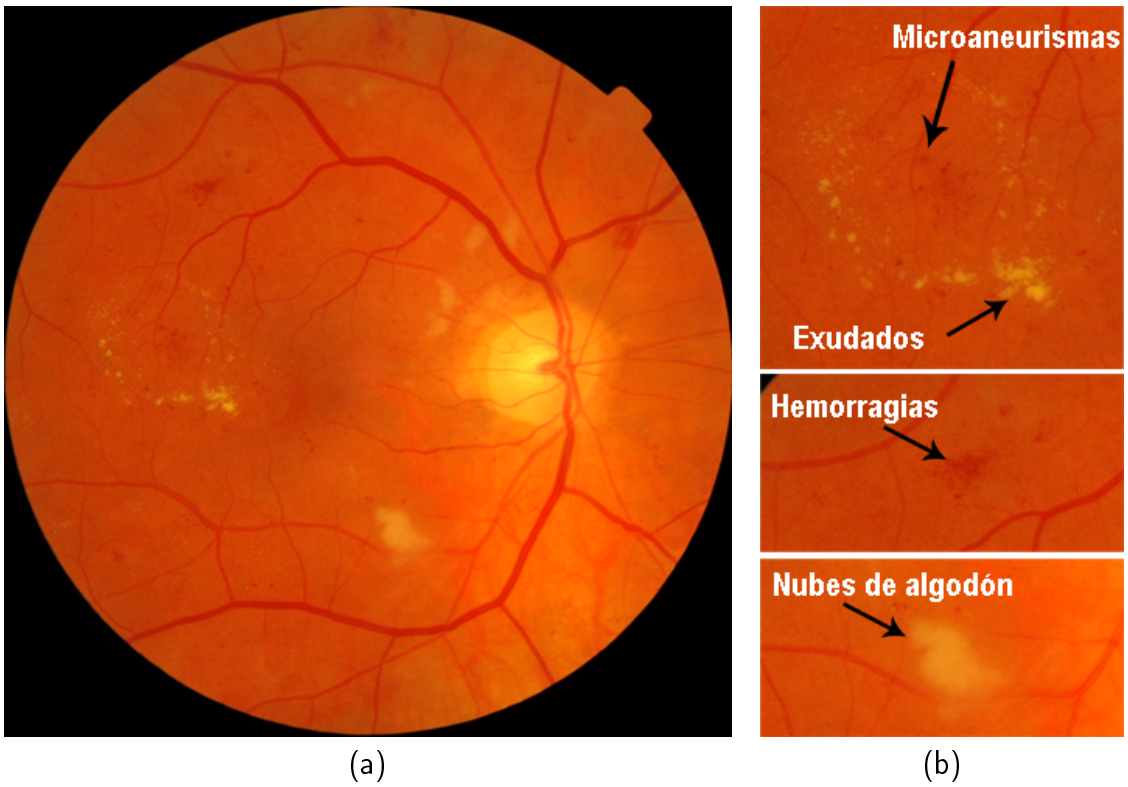


Figura 3: *Ejemplos de las lesiones más comunes producidas por la RD en la retina: (a) Retinografía. (b) Detalles de la imagen (a) mostrando ejemplos de microaneurismas, hemorragias, exudados y nubes de algodón.*

es menos efectivo. Por tanto, para asegurar el tratamiento precoz de la RD, los pacientes diabéticos requerirían de un examen anual de su fondo de ojo. Esta acción preventiva supone, sin embargo, un gran reto para los servicios de salud debido a la ingente cantidad de pacientes que necesitan estas revisiones oftalmológicas.

La retina es la única parte del cuerpo humano donde los vasos sanguíneos pueden ser inspeccionados directamente de una manera no invasiva por medio de los retinógrafos digitales (ver un ejemplo de retinógrafo en la Figura 4). Estos dispositivos son capaces de capturar fotos digitales de la superficie de la retina denominadas retinografías. El análisis estas retinografías se utiliza para diagnosticar la RD con un alto nivel de confianza y confort para el paciente. Este uso de imágenes digitales podría aprovecharse para la detección precoz computerizada de la RD. Un sistema robusto que pudiera ser usado por médicos no expertos para filtrar casos diagnósticos, reduciría la carga de trabajo de los especialistas e incrementaría la efectividad de los protocolos preventivos y los tratamientos terapéuticos. Ello también se traduciría en importantes ahorros económicos para los sistemas de salud

Figura 4: *Retinógrafo no midriático manufacturado por Topcon Corporation.*

públicos. Debido a estas razones, en los últimos años se ha invertido mucho esfuerzo a nivel internacional en el desarrollo de sistemas de tales características. Fruto de ese esfuerzo son algunos prototipos publicados y disponibles en la literatura. Sin embargo, hasta la fecha, ninguno de ellos ha demostrado los niveles de robustez y precisión necesarios para su uso en la práctica clínica. Esta conclusión es corroborada por los autores de un prototipo de un sistema de detección automatizada de la RD presentado recientemente en 2008 en la revista *Diabetes Care*. Michael D. Abràmoff (doctor en medicina del Servicio de Retina del Departamento de Oftalmología y Ciencias Visuales de la Universidad de Iowa Hospitales y Clínicas, Iowa, U.S.A), el autor principal de dicho trabajo, concluyó a partir de sus resultados lo siguiente: "La detección automatizada de la RD usando algoritmos publicados no puede ser aún recomendada para la clínica práctica. Sin embargo, el rendimiento es tal que la evaluación sobre bases de datos disponibles debe llevarse a cabo. Si los algoritmos pudieran ser mejorados, un sistema de tales características podría en el futuro mejorar la prevención de la ceguera y de la pérdida de visión en pacientes con diabetes".

Centrando ahora la atención en un sistema experto para el diagnóstico automatizado de la RD mediante el análisis de retinografías, el objetivo final es la obtención de un sistema capaz de detectar lesiones producidas por la enfermedad para tener así la capacidad de diagnosticarla automáticamente. Estas lesiones deben ser detectadas sobre una superficie retinal cuya

apariencia varía sustancialmente de individuo a individuo. Este hecho ha motivado el uso de un enfoque estructurado que consiste en la segmentación de los tres principales componentes anatómicos de la retina para usarlos como puntos de referencia. Así, sus características anatómicas conocidas y las relaciones entre éstos se usan para incrementar la confianza en la clasificación de cada objeto. A este respecto, la segmentación e identificación del disco óptico es valiosa para localizar la mácula ya que se conoce que existe una distancia constante entre ellos. Además, ya que el árbol vascular tiene su origen en el disco óptico, el conocimiento de la localización de este último puede también ayudar a identificar el nacimiento de los vasos. Otro ejemplo del uso del disco óptico en el enfoque estructurado mencionado es que, ya que las características de color de éste pueden ser similares a las de los exudados retinales, su segmentación puede ser explotada para reducir los falsos positivos en la detección de estas lesiones. En cuanto a los vasos retinales, su segmentación también juega un importante rol en el desarrollo de sistemas automatizados para el diagnóstico de la RD. Por ejemplo, el origen de la vasculatura puede emplearse para localizar el disco óptico. Por otro lado, ya que el centro de la mácula, la fóvea, es una zona avascular, el conocimiento acerca de la distribución de los vasos puede ayudar a identificar esta zona retinal. Es más, debido a la similitud en color entre los vasos y las lesiones rojas (hemorragias y microaneurismas), tener información acerca de los primeros ayuda a reducir falsos positivos en la detección de las segundas. Finalmente, la avascularidad de la zona central macular puede ser usada en sentido contrario, es decir, para ayudar a reducir falsos positivos en la detección de los vasos.

La identificación de los componentes anatómicos retinales no es sólo necesaria debido a lo expuesto arriba, ya que éstos también son utilizados como claves diagnósticas en muchas clasificaciones internacionales de la RD. En estas clasificaciones se usan reglas basadas en la localización de anomalías con respecto a estos componentes para clasificar la enfermedad. Un claro ejemplo de ésto es el siguiente: ya que la mácula es responsable de la visión central, la distancia a la que las lesiones están localizadas con respecto a ésta, tiene influencia en la relevancia de las mismas y, por tanto, se usa para calibrar el grado de severidad de afectación de la enfermedad.

Teniendo en cuenta todas estas consideraciones, se puede proponer que la construcción de un sistema para el diagnóstico automático de la RD sea dividida estructuralmente en cuatro fases definidas como sigue (ver la Figura 5):

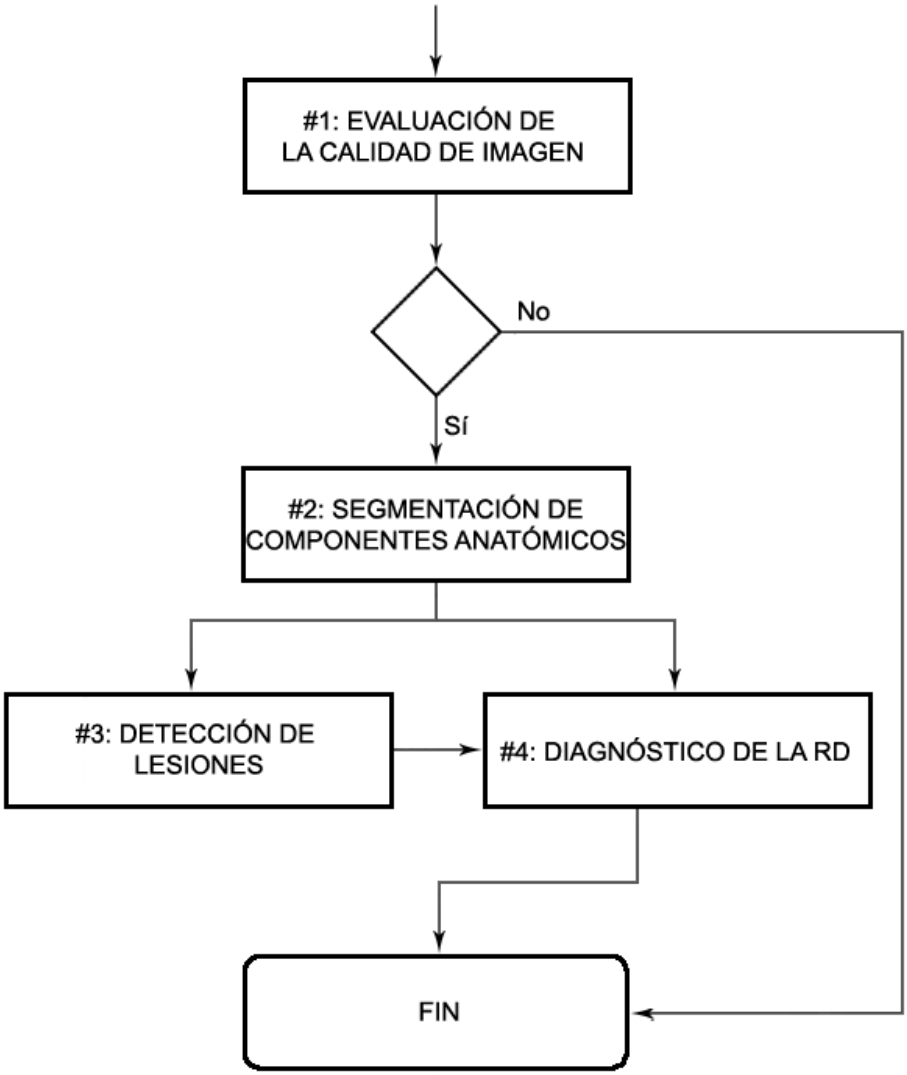


Figura 5: *Diagrama de flujo de un sistema genérico para la detección automatizada de la RD.*

1. Evaluación de la calidad de la imagen: en esta fase se implementan algoritmos para chequear si una imagen tiene suficiente calidad como para ser procesada.

2. Segmentación de los principales componentes anatómicos de la retina: aquí se desarrollan algoritmos para la segmentación automatizada del disco óptico, árbol vascular y mácula.

3. Detección de lesiones producidas por la RD: esta fase trata del desarrollo de algoritmos para la detección automatizada de las principales lesiones producidas por la RD: microaneurismas, hemorragias, exudados y nubes de algodón.

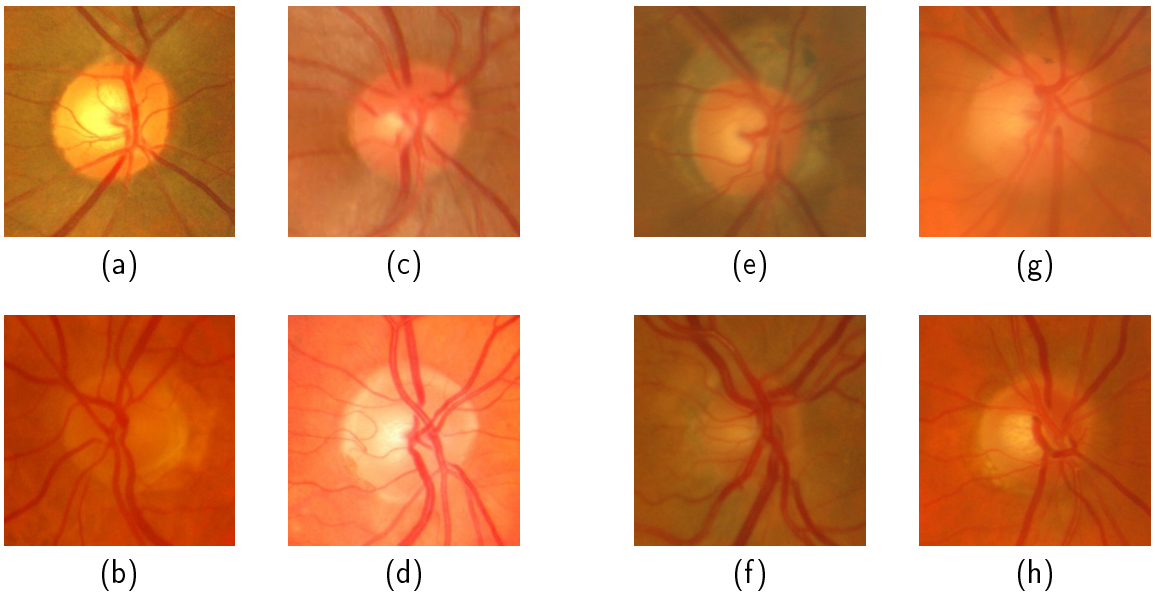
4. Implementación de un sistema experto: esta fase consiste en la implementación de un sistema experto para la diagnosis de la RD. Este sistema, tomando los resultados obtenidos en las fases previas, realiza el diagnóstico basándose en clasificaciones internacionales de la enfermedad.

Esta tesis se centra en la segunda de las fases definidas. Aquí se proponen y describen nuevas metodologías para la segmentación del disco óptico, árbol vascular y mácula que han mostrado un alto grado de robustez y confianza. Además, en el contexto de segmentación automática del árbol vascular, también se describe una nueva función de evaluación de la calidad global de estas segmentaciones. Todas las propuestas de segmentación mejoran la precisión de las técnicas más representativas presentes en la literatura. El trabajo que se presenta, por tanto, constituye la consecución exitosa de una fase del desarrollo de un sistema integral para la detección precoz automatizada de la RD. A continuación, se describen brevemente las metodologías de segmentación mencionadas y la nueva función de evaluación de la calidad de segmentaciones de vasos retinales.

2 Segmentación del disco óptico

El disco óptico (DO) puede distinguirse en imágenes de fondo de ojo como una forma más o menos elíptica. Su tamaño varía significativamente de una persona a otra, ocupando entre un décimo y un quinto de la superficie retinal visible. Con respecto a su color, el DO usualmente tiene una apariencia amarillenta, aunque esta característica también puede experimentar variaciones significativas. Además de estas variaciones en color, forma y tamaño, hay otras circunstancias adicionales que añaden dificultad a la segmentación del DO, como son el contraste irregular y a veces insuficiente de su contorno debido a la no adecuada toma de la retinografía, o la presencia de patologías, especialmente de atrofia peripapilar (ver la Figura 6).

La metodología de segmentación del DO que se propone proporciona una herramienta eficaz para obtener una aproximación circular de su contorno en imágenes digitales de fondo de ojo (ver ejemplos de resultados en la Figura 7). La metodología se basa en técnicas morfológicas y de detección de bordes y usa la Transformada Circular de Hough para obtener la mencionada aproximación. El algoritmo de segmentación parte de una subimagen que contiene el DO, generada a partir de un pixel localizado dentro de éste. Para la obtención de



Variabilidad de color del DO: (a) DO amarillento. (b) DO amarronado. (c) DO rojizo. (d) DO blanquecino.

Distractores para la segmentación del DO: (e) DO con atrofia peripapilar. (f) DO pobremente iluminado. (g) DO ligeramente borroso. (h) Contornos del DO parcialmente oscurecidos.

Figura 6: *Ejemplos que ilustran la variabilidad de apariencia del DO.*

este pixel, también se propone un método de localización del DO (ver ejemplos de resultados en la Figura 8).

La metodología de segmentación fue evaluada subjetiva y objetivamente sobre la base de datos MESSIDOR, la cual contiene 1200 imágenes diagnosticadas de RD. La evaluación subjetiva, realizada por expertos médicos, concluyó que la metodología obtuvo una buena aproximación al contorno del DO en el 94.58% de los casos, lo cual significa éxito en 1135 de 1200 imágenes. La evaluación objetiva fue llevada a cabo midiendo el área común de solapamiento entre las segmentaciones automatizadas y las segmentaciones pertenecientes a un conjunto *gold standard*. Este conjunto fue realizado por expertos médicos mediante el trazado a mano del contorno del DO en las 1200 imágenes. La metodología proporcionó un solapamiento medio de áreas del 86% medido sobre el total de las imágenes. Este porcentaje supera a otras propuestas comparadas y presentes en la bibliografía. Atendiendo a los resultados evaluados de forma separada, según los distintos casos diagnósticos disponibles, las conclusiones extraídas son similares tanto para la evaluación subjetiva como para la objetiva que se realizó: el rendimiento del algoritmo presentado no es sensible a la presencia, ausencia

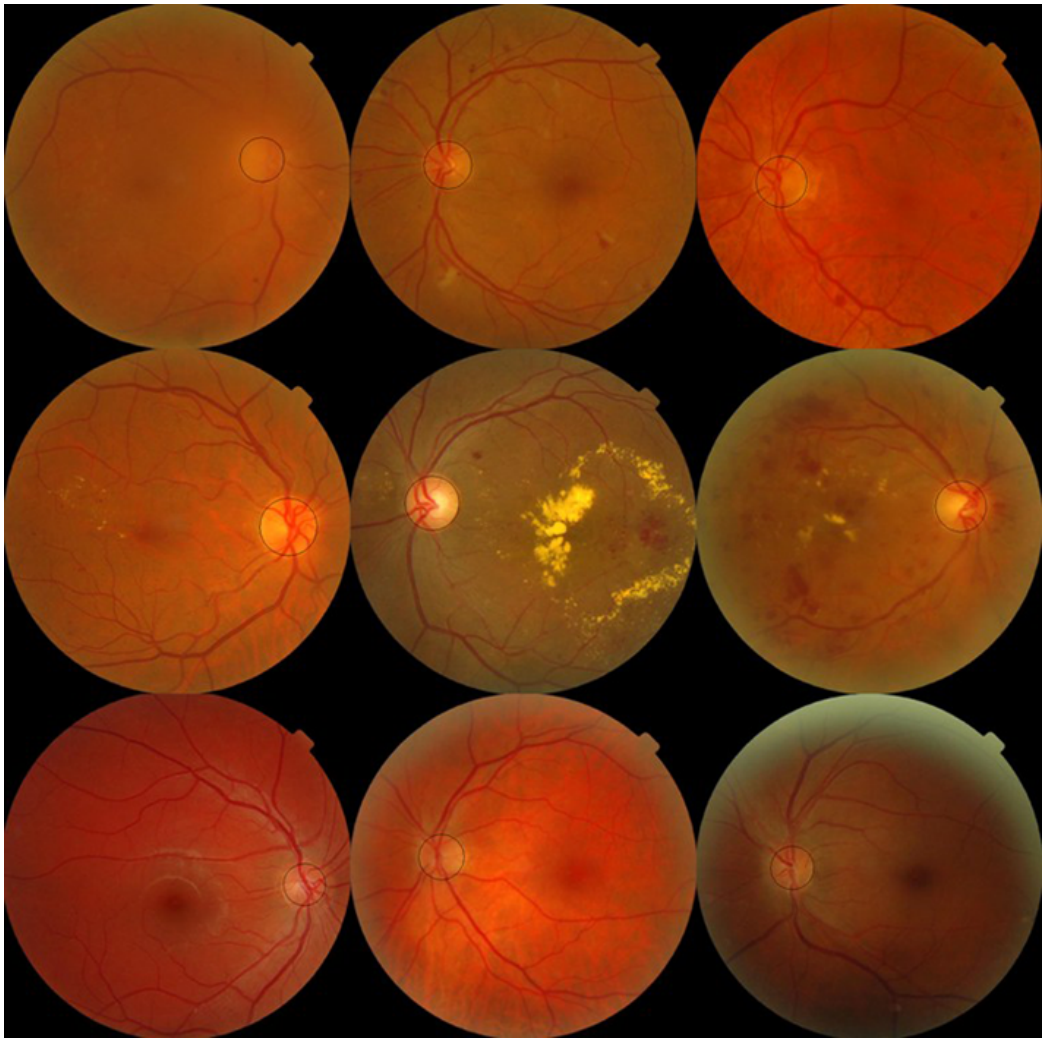


Figura 7: *Ejemplos de resultados obtenidos con la metodología de segmentación del DO. La primera fila muestra tres retinografías de pacientes afectados por RD sin riego de edema macular, la segunda fila muestra tres retinografías de pacientes con RD y riesgo de edema macular, y la última fila contiene retinografías de pacientes sanos.*

o gravedad de afectación de la enfermedad. Con respecto a la eficiencia computacional de la metodología, ésta también constituye una propuesta rápida de segmentación, ya que para un total de 1200 ejecuciones del algoritmo, la media computacional medida fue de 5.69 seg. con una desviación estándar de 0.54 sec.

3 Segmentación del árbol vascular

El árbol vascular retiniano puede ser reconocido en imágenes de fondo de ojo como una red de filamentos conectados rojizos o anaranjados que nace en el disco óptico (ver la Figura 9).

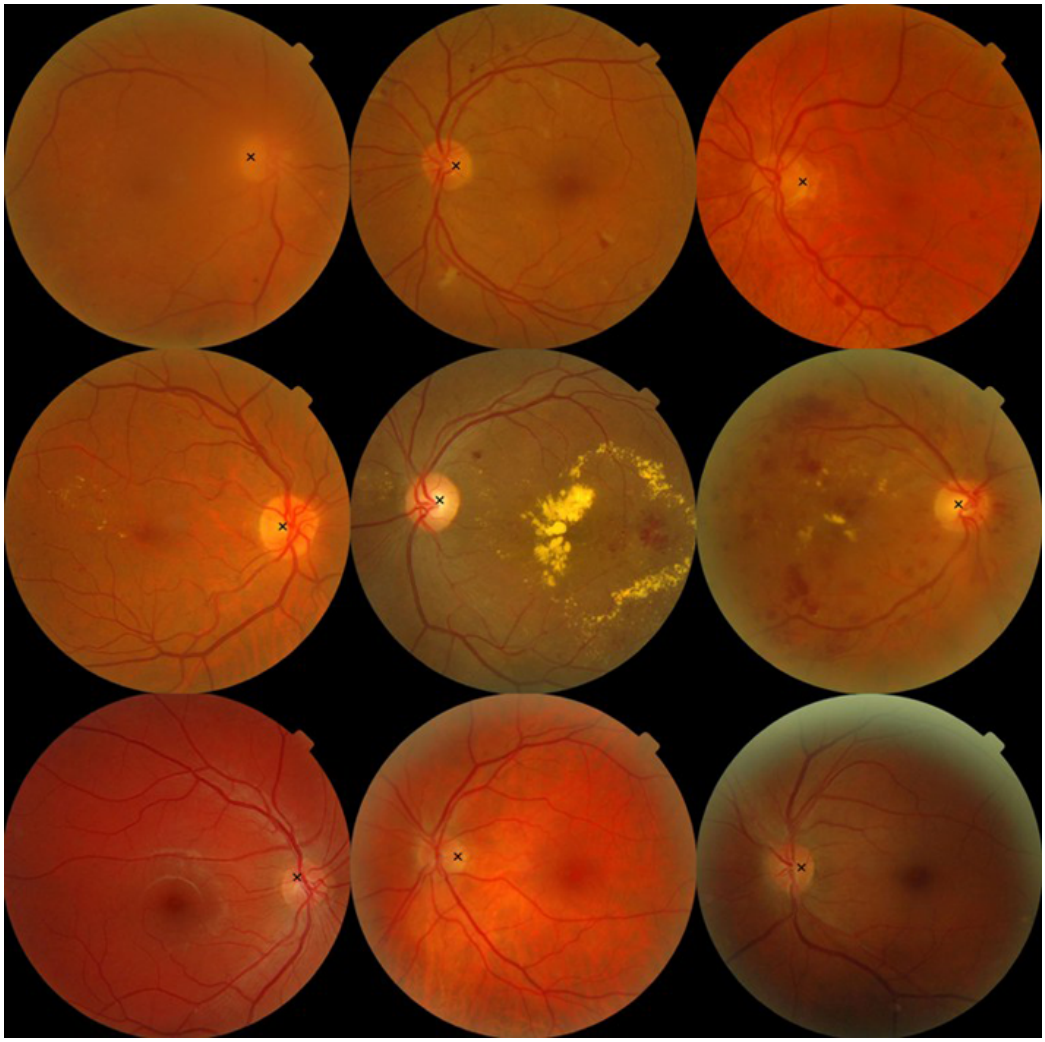


Figura 8: *Ejemplos de resultados obtenidos con la metodología de localización del DO. En la primera fila se muestran tres retinografías de pacientes afectados por RD sin riego de edema macular, en la segunda tres retinografías de pacientes con RD y riesgo de edema macular, y en la última tres retinografías de pacientes sanos.*

En su nacimiento, el árbol vascular emerge con algunos vasos principales gruesos, los cuales se van progresivamente dividiendo en otros más delgados.

En esta tesis se propone una metodología de segmentación del árbol vascular basada en un enfoque supervisado. En este enfoque se emplea un esquema de red neuronal para la clasificación de los píxeles. Para ello, la red se alimenta a su entrada con un vector 7-dimensional compuesto por características de nivel de gris y momentos invariantes. Este vector de características se computa a partir de imágenes retinales preprocesadas en la vecindad del pixel bajo consideración. Los valores de clasificación procedentes de la red (valores reales entre 0 y 1) se umbralizan para clasificar cada pixel en una de las siguientes clases: capilar o no

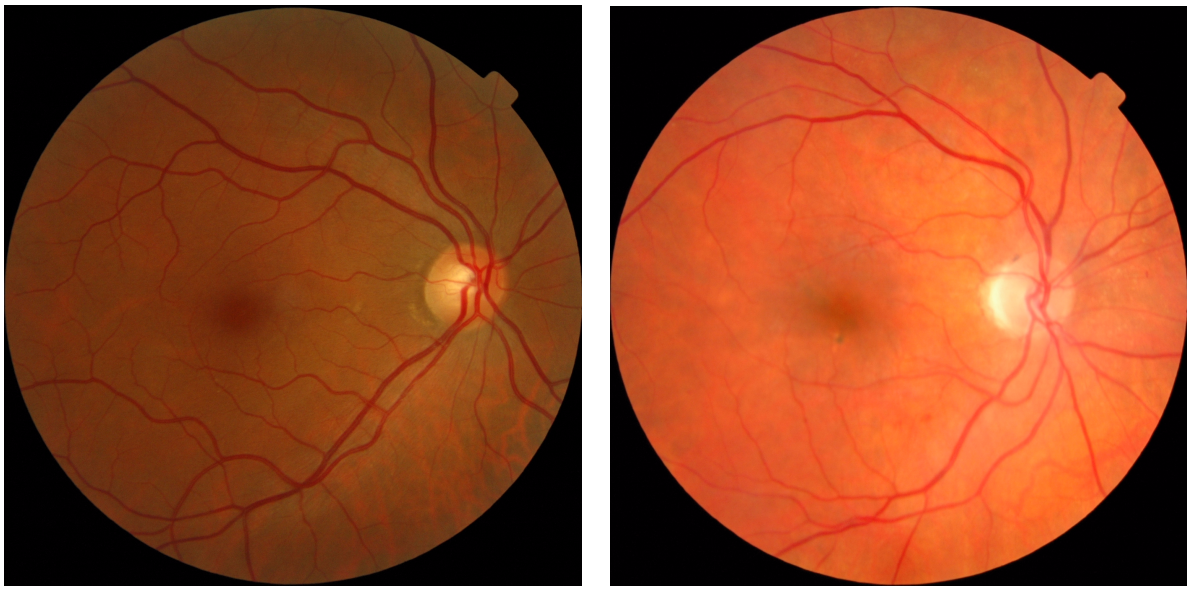


Figura 9: *Variabilidad en color del árbol vascular: (a) Árbol vascular rojo oscuro. (b) Árbol vascular anaranjado.*

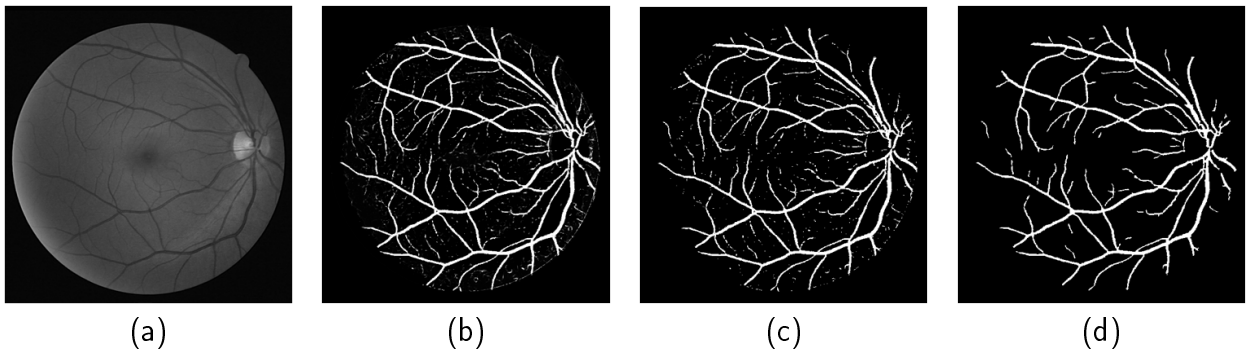


Figura 10: *(a) Canal verde de la imagen original. (b) Salida de la red representada como una imagen. (c) Imagen umbralizada. (d) Imagen posprocesada.*

capilar. Finalmente, un posprocesamiento rellena huecos en los vasos detectados y elimina falsos positivos mediante un proceso de reclasificación de aquellos píxeles aislados clasificados como capilar (ver la ilustración del proceso en la Figura 10).

La metodología fue evaluada y comparada con los métodos de segmentación más relevantes disponibles en la bibliografía. Con el objetivo de comparar la nueva propuesta con estos métodos, se usaron como medidas de rendimiento el *accuracy* y el *área bajo la curva ROC*. El hecho de que estas medidas fueron usadas por los autores de los trabajos considerados, facilita la comparación de resultados. Los valores de *accuracy* y *área bajo la curva ROC*

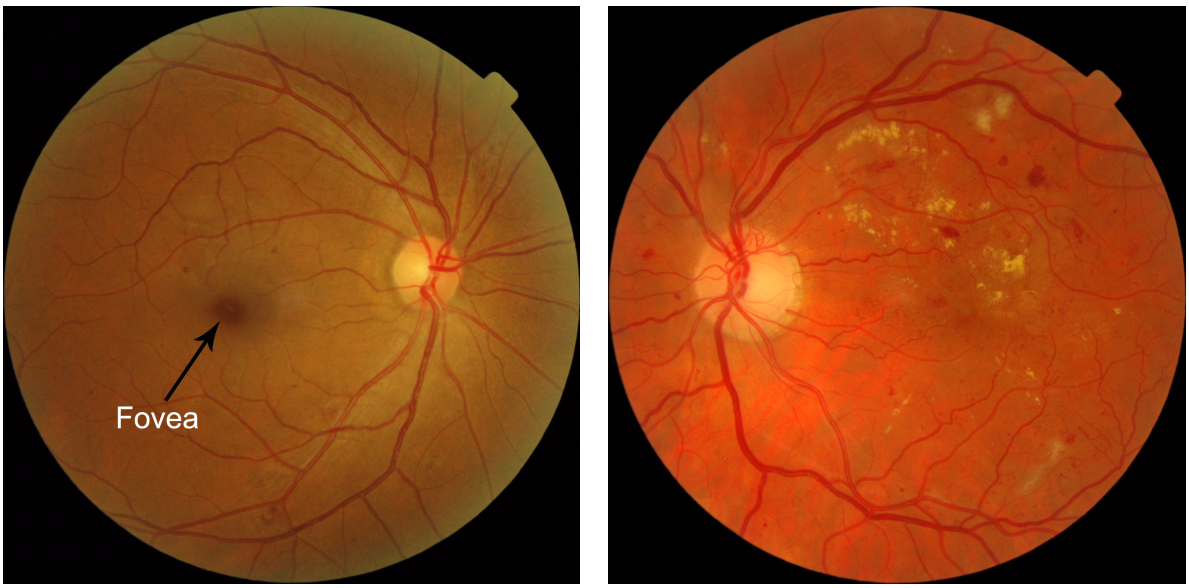


Figura 11: *Apariencia de la fóvea en imágenes de fondo de ojo: (a) En esta imagen, la fóvea es fácilmente reconocible como una región oscura redondeada. (b) Esta imagen muestra una fóvea no visible debido a la presencia de lesiones.*

proporcionados por la metodología presentada en esta tesis, calculados conjuntamente sobre la base de datos DRIVE y STARE, fueron 0.9621 y 0.9678 respectivamente. Estos resultados superan todos los trabajos de segmentación analizados y comparados. Con respecto a la complejidad computacional, el tiempo computacional medido fue menos de 90 seg. por imagen. Este dato, a pesar de que la implementación realizada es experimental, indica que la nueva propuesta es también una solución atractiva desde un punto de vista computacional.

4 Identificación de la zona macular

La mácula exhibe en imágenes retinales a color una estructura inespecífica y su aspecto varía ampliamente entre individuos debido a variaciones en los niveles de pigmentación asociados a factores como etnia, edad, dieta y estados de enfermedad. Aunque generalmente ésta no es distinguible es su totalidad, su centro, llamado fóvea, es a menudo reconocible como una región redondeada más oscura que su tejido retinal circundante (ver la Figura 11, imagen (a)). Sin embargo, esta marca puede no ser visible debido a, por ejemplo, la presencia de lesiones (ver la Figura 11, imagen (b)).

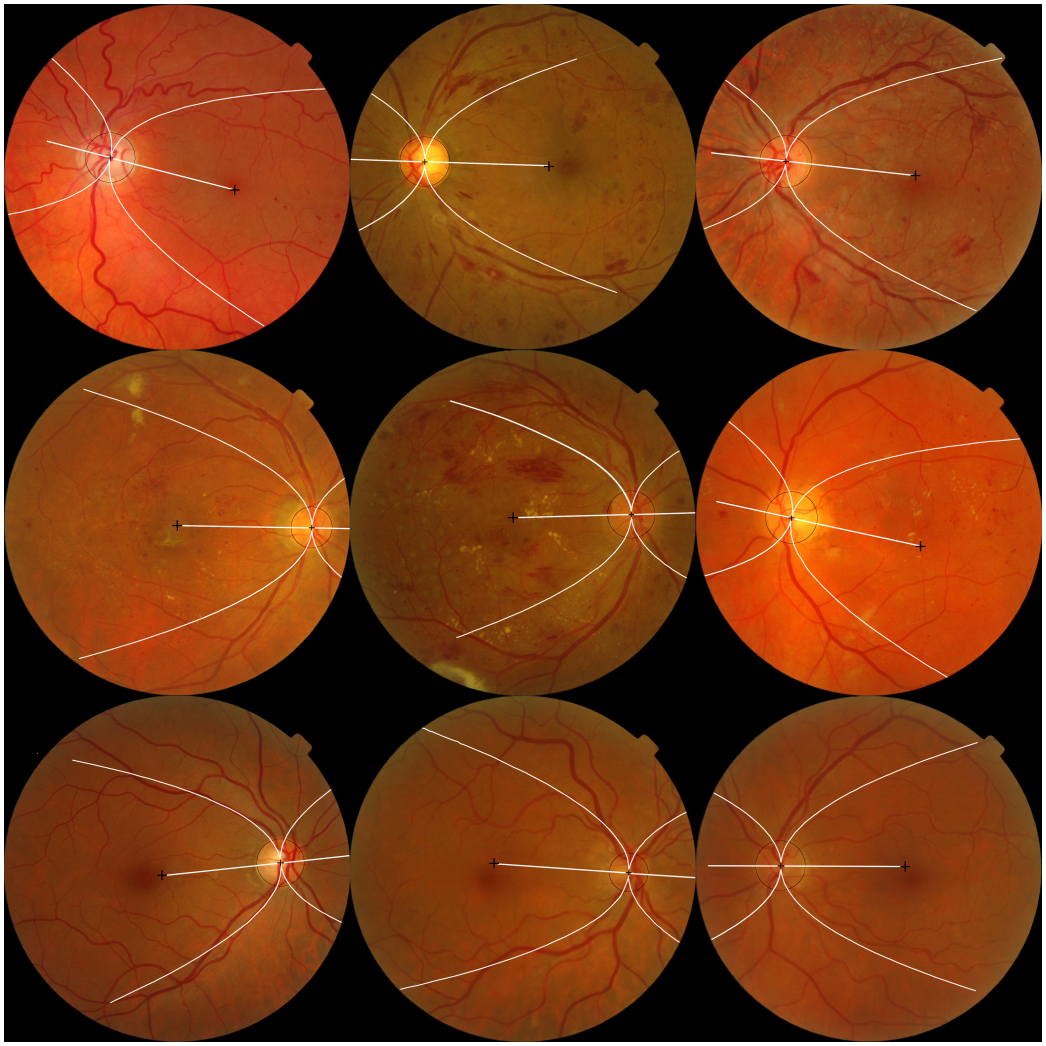


Figura 12: *Ejemplos de estimación del centro de la mácula. La primera fila muestra tres retinografías de pacientes con RD sin riesgo de edema macular, la segunda se compone de tres retinografías de pacientes afectados por RD con riesgo de edema macular y las últimas tres imágenes son de pacientes sanos.*

La metodología presentada para la identificación de la zona macular ofrece una nueva elección para su delimitación automática en retinografías digitales por medio de la obtención de su centro, esto es, el centro de la fóvea. La metodología, en primer lugar, calcula una estimación de este punto aprovechando sus características posicionales conocidas con respecto al disco óptico y al árbol vascular. Para tal efecto, las metodologías de segmentación para el disco óptico y árbol vascular propuestas anteriormente se usan para estimar el eje que delimita las regiones superior e inferior de la retina determinadas por los vasos superiores e inferiores. Una vez estimado éste, el centro de la fóvea se sitúa a 2.5 diámetros del disco óptico desde su centro siguiendo el eje (ver la Figura 12).

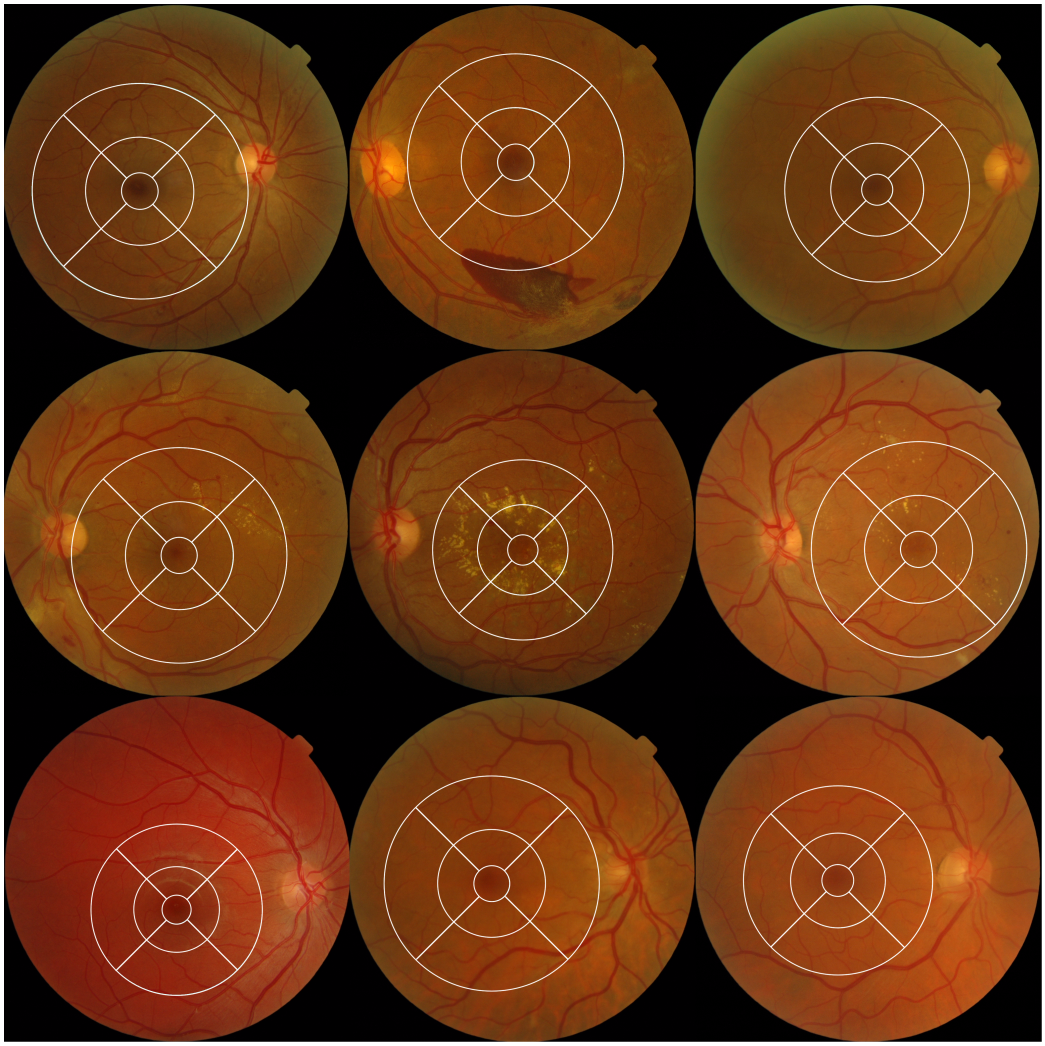


Figura 13: *Ejemplos de identificación de la zona macular. La primera fila muestra retinografías de pacientes con RD sin riesgo de edema macular, la segunda se compone de tres retinografías de pacientes afectados por RD con riesgo de edema macular y la tercera muestra imágenes de pacientes sanos.*

En segundo lugar, la estimación obtenida del centro de la fóvea se refina para obtener una posición más exacta usando un procesamiento morfológico. Este procesamiento se aplica en una subimagen que contiene a la fóvea centrada en la posición estimada de la misma. Por medio del mencionado procesamiento morfológico, la fóvea es segmentada y evaluada en términos de circularidad y tamaño. Si la fóvea segmentada satisface ambos criterios, ésta es validada y se calcula su centro preciso. En otro caso, la metodología ofrece la primera estimación obtenida. Finalmente, se identifica la zona macular usando un sistema de coordenadas polares centrado en la localización definitiva del centro de la fóvea (ver la Figura 13).

La metodología fue testeada sobre las 1200 imágenes de la base de datos MESSIDOR usando un criterio de distancia entre el centro de la fóvea obtenido y el centro real. Este criterio establece cuatro categorías de calidad de acuerdo a la magnitud de esta distancia: *Excelente*, *Bueno*, *Aceptable* y *Pobre*. La metodología obtuvo centros de fóvea en el intervalo *Excelente* - *Aceptable* (centros de fóvea ampliamente aceptados como válidos en la literatura) en el 97.42% de los casos. Es más, el 88.55% de los centros de fóvea obtenidos fueron en el intervalo *Excelente* - *Bueno*. Estos resultados superan ampliamente los publicados hasta la fecha en la bibliografía. Con respecto a la complejidad computacional, el algoritmo toma en media 10.88 seg. para ejecutar todo el proceso, excluyendo de esta medida el tiempo necesario para segmentar el DO y el árbol vascular. Este dato hace al algoritmo presentado, además de una opción altamente confiable, una solución válida desde un punto de vista computacional.

5 Una función especializada para la evaluación de la calidad de segmentaciones de vasos retinales

Además de las metodologías presentadas para la segmentación de los principales componentes anatómicos de la retina, esta tesis también propone una nueva función de evaluación de la calidad global de segmentaciones del árbol vascular.

Las métricas basadas en las tablas de contingencia para clasificación binaria han sido ampliamente usadas para evaluar la calidad de las segmentaciones de vasos retinales. Particularmente, el *accuracy* ha sido usado mayoritariamente como medida de rendimiento global de los algoritmos que obtienen tales segmentaciones y, por tanto, ha sido usado para compararlos en la literatura. Sin embargo, cuando se utilizó dicha métrica en el proceso de investigación de la metodología de segmentación del árbol vascular se le identificaron algunas debilidades. Estas debilidades pueden ser resumidas como sigue:

- Baja sensibilidad a la deficiente segmentación de vasos delgados.
- Baja sensibilidad a la pobre segmentación del árbol vascular.
- Alta dependencia con la imagen gold standard usada para medir la calidad de las segmentaciones.

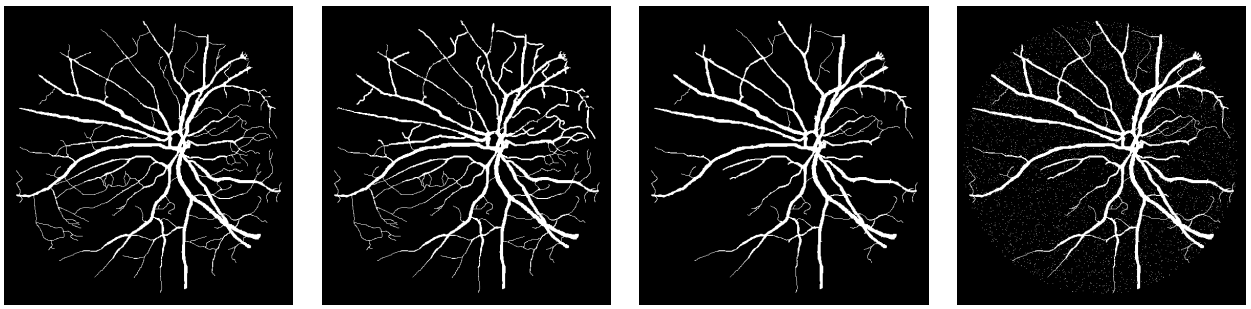


Figura 14: *Ejemplos de mala correspondencia entre la evaluación de calidad del* accuracy *y la percepción de calidad subjetiva: (a) Gold standard de referencia. (b)-(d) Imágenes con un valor de* accuracy *de* 0.9681. *Aunque el* accuracy *indica igual calidad para las tres segmentaciones, la calidad que se percibe subjetivamente es diferente para cada una de ellas.*

Así mismo, se encontró que las citadas debilidades daban lugar a una mala correspondencia entre la evaluación de calidad proporcionada por el *accuracy* y la calidad percibida subjetivamente por un observador (ver ejemplo en la Figura 14).

La función de evaluación de calidad de segmentaciones de vasos retinales que se presenta, se basa en la evaluación de tres características principales medibles que describen la vascularidad en la retina: conectividad del árbol vascular, área ocupada por el mismo y longitud de la estructura. Esta función de evaluación representa un enfoque más natural para evaluar segmentaciones vasculares y, ya que ésta se basa en la evaluación de características descriptivas del árbol vascular, también proporciona mayor interpretabilidad en los resultados, lo cual permite detectar con mayor facilidad características concretas del mismo (conectividad, área o longitud) inapropiadamente segmentadas.

6 Conclusiones y trabajo futuro

La RD es una enfermedad retinal crónica que afecta a la población diabética. Esta población se estima que se incrementará desde 171 millones de personas en el año 2000 hasta 366 millones en el 2030. Se calcula que aproximadamente el 2% de los pacientes diabéticos están ciegos y que el 10% experimenta degradación de la visión tras 15 años de afectación debido a la RD. La fotocoagulación láser es actualmente el principal tratamiento terapéutico usado para retrasar la degradación de la visión en pacientes con RD, aunque éste es eficaz sólo si se aplica cuando la enfermedad se encuentra en sus fases iniciales. Por tanto, la

detección precoz de la RD es un factor clave para la preservación de la visión en pacientes afectados. No obstante, ésto representa una dificultad importante ya que los pacientes con RD usualmente no aprecian pérdida de visión hasta que la enfermedad se encuentra en fases avanzadas. Por ello, los pacientes diabéticos son anualmente examinados en sus sistemas de salud públicos en busca de signos precoces de RD. Sin embargo, el examen de una población tan numerosa supone todo un reto para estas administraciones y tiene como consecuencia que muchos pacientes terminan no siendo apropiadamente tratados. En este contexto, un sistema que pudiera ser usado por no expertos para discriminar y filtrar con mayor confianza aquellos casos que requieren visita al oftalmólogo para mayor investigación y tratamiento de aquellos que no lo requieren, reduciría la carga de trabajo de estos especialistas, mejoraría los protocolos preventivos y, por tanto, también reduciría costes.

Esta tesis está enmarcada en un proyecto de investigación para el desarrollo de un sistema para la detección automatizada de la RD mediante el análisis automático de imágenes de fondo de ojo. La construcción de un sistema de tales características puede ser estructuralmente dividida en cuatro módulos. Estos módulos tendrían los siguientes objetivos: evaluación de la calidad de la imagen, segmentación de los principales componentes anatómicos de la retina, detección de lesiones y diagnóstico de la RD. Las investigaciones llevadas a cabo para desarrollar el segundo de los módulos mencionados ha sido la materia central de esta tesis. Este módulo comprende la segmentación del disco óptico y del árbol vascular, así como la identificación de la zona macular. De este modo, se han investigado y desarrollado metodologías automatizadas robustas y confiables para el reconocimiento de estas características retinales.

Esta tesis presenta metodologías altamente precisas para la segmentación del disco óptico, árbol vascular y zona macular. Ya que todas ellas mejoran los resultados de todos los trabajos analizados existentes en la bibliografía, el trabajo aquí presentado supone un avance en el desarrollo de sistemas para la detección automatizada de la RD.

El trabajo futuro relacionado con esta tesis se centra en la detección de lesiones producidas por la enfermedad. De hecho, este trabajo está siendo llevado a cabo en la actualidad en nuestro grupo de investigación en el marco de la segunda fase del proyecto anteriormente mencionado y financiado por la Consejería de Salud de la Junta de Andalucía (Andalucía, España). Se están realizando investigaciones para la detección de microaneurismas, hemorragias y exudados con el objetivo de disponer, a finales de 2012, de un sistema integral

para el procesamiento y análisis de imágenes retinales. En ese punto, está proyectada la implementación de un sistema experto para la clasificación de la enfermedad y la realización de una experiencia piloto en algunos centros de atención primaria del Servicio Andaluz de Salud para su testeo y validación.

7 Hitos y producción científica de la tesis

A continuación se detalla el marco de trabajo en el que se desarrolló la presente tesis y se presenta la producción científica que ha generado. Con respecto a lo primero, se describe el proyecto que financió este trabajo, así como la estancia predoctoral realizada en un centro de investigación internacional. Con respecto a lo segundo, se detallan las principales publicaciones derivadas de esta tesis en revistas internacionales y congresos.

Projectos de investigación

- **Título:** Sistema Experto para la Detección Precoz Automatizada de la Retinopatía Diabética mediante Análisis de Imágenes Digitales de la Retina.
  - **Entidad financiadora:** Consejería de Salud de la Junta de Andalucía (Andalucía, España).
  - **Entidades participantes:** Universidad de Huelva y Servicio Andaluz de Salud.
  - **Descripción del proyecto:** el objetivo final del proyecto es desarrollar un sistema para la detección precoz de la RD mediante el análisis de imágenes digitales de fondo de ojo. La Consejería de Salud de la Junta de Andalucía puso en marcha un plan integral para la prevención de la diabetes, el cual incluye protocolos para la detección precoz de la mencionada patología retinal. Estos protocolos contemplan exámenes periódicos de la retina realizados por médicos de cabecera en los centros de atención primaria del Servicio Andaluz de Salud. El sistema que se encuentra bajo investigación, haciendo uso de la red de retinógrafos interconectados existentes en los centros de atención primaria, pretende mejorar los protocolos preventivos de la enfermedad ayudando a los médicos de cabecera a filtrar con mayor confianza casos que necesitan atención por parte de los oftalmólogos de aquellos que no.
  - 1^a **Fase del proyecto**

    - **Duración:** desde el 01/01/2007 hasta el 31/12/2009.
    - **Cantidad financiada:** 150000 €.
    - **Investigador principal:** Dr. José Manuel Andújar.
    - **Objetivo de la fase:** desarrollar metodologías para la segmentación de los principales componentes anatómicos de la retina.
  - 2^a **Fase del proyecto:**
    - **Duración:** desde el 01/01/2011 hasta el 31/12/2013.
    - **Cantidad financiada:** 180000 €.
    - **Investigadores principales:** Dr. Manuel E. Gegúndez y Dr. Diego Marín Santos.
    - **Objetivo de la fase:** desarrollar algoritmos para la detección de las principales lesiones producidas por la RD e implementar un sistema experto para su diagnóstico.

- **Título:** Métodos para evaluar la segmentación y técnicas de indexado en el campo de la oftalmología retinal (MESSIDOR).
  - **Entidad financiadora:** Ministerio Francés de Investigación y Defensa.
  - **Entidades participantes:** ARMINES : Centro de Morfología Matemática, Laboratorio L3I - Universidad de la Rochelle, Departamento ITI - LaTIM INSERM U650, Laboratorio SIC - CNRS FRE 2731, Universidad Nottingham Trent: Escuela de Computación e Informática, EA 3063 / Oftalmología - Facultad de Medicina, Servicio de oftalmología Daviel - CHU Brest, Servicio de Oftalmología - Hospital Lariboisiere, ADCIS y CRIHAN.
  - **Descripción del proyecto:** en el marco de la retinopatía diabética, el principal objetivo del proyecto MESSIDOR es comparar y evaluar:
    - Los distintos algoritmos de segmentación desarrollados para la detección de lesiones presentes en imágenes retinales a color.
    - Las herramientas para indexar y gestionar las bases de datos de imágenes.

Estancia predoctoral

- **Centro:** Centro de Morfología Matemática de la Escuela Nacional Superior de Minas.

- **Dirección del centro:** 35, rue Saint Honoré, 77305 Fontainebleau CEDEX (France).
- **Director del centro:** Dr. Fernand Meyer.
- **Supervisor de la estancia:** Dr. Jean-Claude Klein.
- **Duración de la estancia:** desde el 01/10/2009 hasta el 15/03/2010.
- **Grupo de trabajo:** Dr. Jean-Claude Klein, Dr. Fernand Meyer, Dr. Etienne Decenciere y Dr. Estelle Parra-Denis.
- **Descripción de la estancia:** la estancia se dividió en dos fases. La primera se trató de un periodo de formación en *morfología matemática* consistente en la asistencia a un curso intensivo de una semana llamado "Escuela de Verano de Morfología Matemática 2009". La segunda fase, la cual duró el resto de la estancia, estuvo centrada en la investigación de la segmentación de drusas haciendo uso la mencionada técnica de procesamiento de imágenes. Las drusas son depósitos amarillentos o blanquecinos de materia extracelular cuyo origen es la degeneración macular asociada a la edad, una enfermedad producida por el deterioro de la porción central de la retina. Como resultado de este trabajo, se creó una librería llamada "OpthalMorphoTools", implementada en el entorno *Morph-M* [1]. Esta librería implementa un conjunto de operadores para el análisis de imágenes digitales de retina y para el realce de drusas.

Lista de publicaciones

- **En revistas:**

  1. **Autores:** Marín, D., Aquino, A., Gegúndez, M.E., Bravo, J.M.
     Título del artículo: A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features.
     Título de la revista: IEEE Transactions on Medical Imaging.
     Detalles de la publicación: volumen 30, issue 1, pages 146-158.
     Año: 2011.

[1] Centro de Morfología Matemática, sección de Industria, Morph-M : software de procesamiento de imágenes especializado en morfología matemática, Francia [Online]:`http://cmm.ensmp.fr/Morph-M/index_en.html`

Indicadores de calidad: la revista está incluida en cinco categorías del ranking Journal Citation Reports. Estas categorías son: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* y *Radiology, Nuclear Medicine & Medical Imaging*. Con un factor de impacto de 3.540 en 2009, la revista está incluida en el primer cuartil de todas las categorías citadas. Su mejor posición relativa es la de 9 de 245 en la categoría *Engineering, Electrical & Electronic* y su mejor posición absoluta es la de 2 en la categoría *Imaging Science & Photographic Technology*.

2. **Autores:** Aquino, A., Gegúndez, M.E., Marín, D.
   Título del artículo: Detecting the Optic Disc Boundary in Digital Fundus Images Using Morphological, Edge Detection and Feature Extraction Techniques.
   Título de la revista: IEEE Transactions on Medical Imaging.
   Detalles de la publicación: volumen 29, número 11, páginas 1860-1869.
   Año: 2010.
   Indicadores de calidad: la revista está incluida en cinco categorías del ranking Journal Citation Reports. Estas categorías son: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* y *Radiology, Nuclear Medicine & Medical Imaging*. Con un factor de impacto de 3.540 en 2009, la revista está incluida en el primer cuartil de todas las categorías citadas. Su mejor posición relativa es la de 9 de 245 en la categoría *Engineering, Electrical & Electronic* y su mejor posición absoluta es la de 2 en la categoría *Imaging Science & Photographic Technology*.

3. **Autores:** Aquino, A., Gegúndez, M.E., Marín, D.
   Título del artículo: Automated Optic Disc Detection in Retinal Images of Patients with Diabetic Retinopathy and Risk of Macular Edema.
   Título de la revista: International Journal of Medicine and Medical Sciences.
   Detalles de la publicación: volumen 1, número 2, páginas 87-92.
   Año: 2010.
   Indicadores de calidad: la revista es editada por la *World Academy of Science, Engineering and Technology* e indexada por: *SCIRUS*, *Google Scholar*, *EBSCO*,

GALE, *INTUTE*, *DOAJ*, *Ulrich's & Serials Solutions* y *Electronic Journals Library*.

4. **Autores:** Aquino, A., Gegúndez, M.E., Bravo, J.M., Marín, D.
   Título del artículo: A Specialized Function for Quality Evaluation of Retinal Vessel Segmentations.
   Título de la revista: IEEE Transactions on Medical Imaging.
   Detalles de la publicación: bajo "major revision".
   Año: -.
   Indicadores de calidad: la revista está incluida en cinco categorías del ranking Journal Citation Reports. Estas categorías son: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* y *Radiology, Nuclear Medicine & Medical Imaging*. Con un factor de impacto de 3.540 en 2009, la revista está incluida en el primer cuartil de todas las categorías citadas. Su mejor posición relativa es la de 9 de 245 en la categoría *Engineering, Electrical & Electronic* y su mejor posición absoluta es la de 2 en la categoría *Imaging Science & Photographic Technology*.

5. **Autores:** Aquino, A., Gegúndez, M.E., Marín, D., Bravo, J.M.
   Título del artículo: Automated Identification of the Macular Zone in Eye Fundus Images Using Anatomical-Based and Visual-Based Features.
   Título de la revista: IEEE Transactions on Medical Imaging.
   Detalles de la publicación: bajo revisión.
   Año: -.
   Indicadores de calidad: la revista está incluida en cinco categorías del ranking Journal Citation Reports. Estas categorías son: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* y *Radiology, Nuclear Medicine & Medical Imaging*. Con un factor de impacto de 3.540 en 2009, la revista está incluida en el primer cuartil de todas las categorías citadas. Su mejor posición relativa es la de 9 de 245 en la categoría *Engineering, Electrical & Electronic* y su mejor posición absoluta es la de 2 en la categoría *Imaging Science & Photographic Technology*.

- **En actas de congresos:**

1. **Autores:** Aquino, A., Gegúndez, M.E., Bravo, J.M., Marín, D.
   Título del artículo: A Similarity Function for Global Quality Assessment of Retinal Vessel Segmentations.
   Título de la conferencia: International Conference on Pattern Recognition and Computer Vision, Bali (Indonesia).
   Detalles de la publicación: volumen 67, páginas 7-11, ISSN 1307-6892.
   Año: 2010.
   Indicadores de calidad: conferencia organizada por la *World Academy of Science, Engineering and Technology*. Actas indexadas por: *Google Scholar*, *IE Compendex*, *INTUTE*, *GALE*, *EBSCO*, *Serials Solutions*, *DOAJ* y *Electronic Journals Library*.

2. **Autores:** Aquino, A., Gegúndez, M.E., Marín, D.
   Título del artículo: Automated Optic Disc Detection in Retinal Images of Patients with Diabetic Retinopathy and Risk of Macular Edema.
   Título de la conferencia: International Conference on Machine Vision, Image Processing and Pattern Analysis, Bangkok (Tailandia).
   Detalles de la publicación: volumen 60, páginas 85-90, ISSN 1307-6892.
   Año: 2010.
   Indicadores de calidad: conferencia organizada por la *World Academy of Science, Engineering and Technology*. Actas indexadas por: *Google Scholar*, *IE Compendex*, *INTUTE*, *GALE*, *EBSCO*, *Serials Solutions*, *DOAJ* y *Electronic Journals Library*.

3. **Autores:** Aquino, A., Gegúndez, M.E., Marín, D., Andújar, J.M.
   Título del artículo: A Robust, High Accuracy and Fast New Methodology to Locate the Optic Disc From Colour Retinal Images.
   Título de la conferencia: IADIS International Conference Computer Graphics, Visualization, Computer Vision and Image Processing, Algarve (Portugal).
   Detalles de la publicación: páginas 99-105, ISBN 978-972-8924-84-3.
   Año: 2009.
   Indicadores de calidad: conferencia organizada por el *International Association for Development of the Information Society*.

4. **Autores:** Aquino, A., Gegúndez, M.E., Marín, D., Andújar, J.M.
   Título del artículo: Sistema Experto para la Detección Precoz de la Retinopatía Diabética Mediante Análisis de Imágenes Digitales de Retina: Primeros Resultados.
   Título de la conferencia: XXIX Jornadas de Automática, Tarragona (España).
   Detalles de la publicación: volumen 29, páginas 1-10, ISBN 978-84-691-6883-7.
   Año: 2008.
   Indicadores de calidad: conferencia organizada por el *Comité Español de Automática*.

- **Patentes:**

1. **Autores:** Gegúndez, M.E., Aquino, A., Bravo, J.M., Marín, D.
   Título de la patente: Sistema y Procedimiento de Obtención de Imágenes de Alto Contraste del Árbol Vascular del Fondo de Ojo.
   Detalles de la patente: patente española, número de aplicación P200930460.
   Año: -.

Contents

List of Figures

List of Tables

Chapter 1

General Thesis Approach

This thesis was not conceived as an isolate work, but it is only a part of the works that are being carried out within the *Control y Robótica* research group (code TEP-192) in the University of Huelva. The work presented in this thesis started in 2007 with the birth of the *Expert System for Early Automated Detection of Diabetic Retinopathy by Analysis of Digital Retinal Images* project, funded by the Health Ministry of the Andalusian Regional Government (Andalusia, Spain). This project, framed within a comprehensive plan for the prevention of diabetes launched by the mentioned government, has the final aim of obtaining a system for the early automated detection of diabetic retinopathy by means of digital eye fundus image analysis. Concretely, the system under investigation is projected to be a software to be used by non-experts to discriminate and filtrate with higher confidence diagnosis cases of diabetic retinopathy. This is, the future system will filtrate those cases that require referral to an ophthalmologist for further investigation and treatment and those that do not, reducing the workload of these specialists and therefore, decreasing costs.

Within the described framework of this thesis, a stay in the Center of Mathematical Morphology of the National High School of Mines in Paris (France) was carried out by this author. The main objectives of this stay were to study theory on mathematical morphology and to work in its application in retinal imaging. For that, the stay was divided into two phases: a first phase of education on mathematical morphology and a second phase of immersion in the work of the retinal imaging group of the center.

This initial chapter is structured in three sections. The *first section* summarizes the research proposal of this thesis, describing motivation and objectives. The *second section* defines the structure and briefly summarizes contents of this book chapter per chapter.

The *third section* offers a detailed description of the framework in which this thesis has been developed and presents the main publications in journals and conference proceedings derived from this work.

1.1 Research proposal

Diabetic retinopathy is the chronic disease which nowadays constitutes the primary cause of blindness in people of working age in the developed world. This disease is characterized by a set of retinal damages caused by complications of diabetes mellitus, an illness that provokes abnormal blood glucose level elevation. This excess of glucose in blood affects the macro- and microvascular system by damaging the vessel endothelium, thus increasing vessel permeability. Usually the eye, concretely the retina, is one of the places where this becomes apparent. When the microvascular system in the eye is progressively damaged, vision loss and blindness can occur.

The retina is a light-sensitive layer at the back of the eye that covers about 65% percent of its interior surface. Photosensitive cells called rods and cones in the retina convert incident light energy into signals that are carried to the brain to be processed, thus producing the sense of vision. Three are the main anatomical components of the retina: the optic disc, vascular tree and macula. The optic disc or papilla represents the location where the nerve fibers from the retina come together and exit the eye towards the brain to carry the nerve signals. The retinal cells are fed by the retinal vessels. These retinal vessels form a vascular tree-like net that originates in the optic disc with some main gross blood vessels. These main vessels progressively split into many other smaller for covering the retina. The photosensitive cells are not equally distributed over the retina. The outlying parts of the retina are responsible for peripheral vision while the central area, the macula, is in charge of detailed central vision. The macula is a round area in the central region of the retina. There is a small depression in the centre of the macula, called the fovea, which corresponds to the region of retina with highest photosensitive cells.

The first manifestations of diabetic retinopathy are tiny capillary dilations known as microaneurysms. With the disease progression these microaneurysms can burst causing haemorrhages. In addition to leaking blood, vessels also leak lipids and proteins causing the appearance of bright spots called exudates. The leaking of fluid from vessels can produce

large exudative areas in the macula which can provoke swelling of the retina, a complication called diabetic macular edema. All this damage can produce that small parts of the retina become ischemic, deprived of blood. These ischemic areas are visible on the retina as fluffy whitish blobs called cotton wool spots. As a response to the appearance of ischemic areas in the retina, new vessels start growing to supply the retina with more oxygen. These vessels, called neovascularisations, have a greater risk of rupturing and causing larger haemorrhages than normal vessels. In the later stages of diabetic retinopathy, as consequence of the great produced damage, retinal detachment can occur.

Although diabetic retinopathy is not a curable disease, laser photocoagulation can prevent major vision loss if the illness is detected in early stages. However, diabetic retinopathy patients perceive no symptoms until visual loss develops, usually in the later disease stages, when the treatment is less effective. Therefore, to ensure the treatment is received in time, diabetic patients need annual eye fundus examination. However, this preventive action involves a huge challenge for health systems due to the huge number of patients needing ophthalmologic revision, thus preventing many patients from receiving adequate treatment and resulting in great economic costs for Administrations.

The retina is the only part of the human body where the blood vessels can be directly seen in a non-invasive manner by means of digital retinographs, which are able to take digital photos of the retinal surface (retinographies). Therefore, analysis of retinographies is used to diagnose diabetic retinopathy with a high level of reliability and comfort for the patient. This employment of digital images could be exploited for computerized early detection of diabetic retinopathy. A robust system that could be used by non-experts to filtrate cases of patients not affected by the disease, would reduce the specialists' workload, and increase the effectiveness of preventive protocols and early therapeutic treatments. Furthermore, it would also result in outstanding economic benefits for public health systems. This is why, in recent years, much effort has been spent on developing automated systems for the detection of diabetic retinopathy and some prototypes have been presented. However, none of these proposals has reached enough accuracy to be recommended for clinical practice. This conclusion is corroborated by the author of one of these proposals, which was early presented in 2008 in the journal *Diabetes Care*. Michael Abràmoff (medical doctor from the Retina Service of the Department of Ophthalmology and Visual Sciences in the University of Iowa Hospitals and Clinics, Iowa, U.S.A.), the main author of that work, concluded from their

results: "automated detection of diabetic retinopathy using published algorithms cannot yet be recommended for clinical practice. However, performance is such that evaluation on validated, publicly available datasets should be pursued. If algorithms can be improved, such a system may in the future lead to improved prevention of blindness and vision loss in patients with diabetes."

Focusing on the construction of an expert system for the automated diabetic retinopathy diagnosis by means of digital retinal image analysis, the final goal is the obtention of algorithms capable to detect lesions produced by the illness to be able to diagnose it automatically. The retinal lesions must be detected on a retinal surface substantially variable in appearance. This fact has motivated a widely used structured approach consisting on the segmentation of the three main retinal components to be used as landmarks. Thus, their a priori-known anatomical features and relationships between them are exploited to increase confidence in the classification of each object. For example, regarding the optic disc, its segmentation and identification is valuable to locate the macula by making use of the known relatively constant distance between them. Since the vascular tree borns within the optic disc, knowledge about its location can also help identifying the vessels. Another example of the use of the optic disc in the mentioned structured approach is that, since color features of the optic disc can be similar to those of retinal exudates, the optic disc segmentation can be exploited to decrease false positives in the detection of these lesions. Regarding the vascular vessels, their segmentation also plays an important role in the development of automated diabetic retinopathy diagnosis systems. For example, the vasculature origin can be used to locate the optic disc. On the other hand, since the macula center, the fovea, is an avascular zone, knowledge about location of vessels can help to locate this retinal component. Moreover, due to color similarity, information about vessels can be also used to decrease false positives in the detection of red lesions (microaneurysms and haemorrhages). Finally, the avascularity of the central macular zone can be used in the contrary way to the described before, i.e., to help locating the vessels by reducing false positives in their detection.

Identification of the retinal anatomical components is not only needed due to the exposed above, but they are also diagnosis keys in many international classifications of the illness. In these classifications, rules based on the location of abnormalities with respect to these components are used to grade diabetic retinopathy. One clear example of this is the following: since the macula is responsible for sharp central vision, the distance at which lesions are

located from this, influences their clinical relevance and therefore is used to grade severity of the illness. Therefore, for all the mentioned reasons, successful segmentation of the main anatomical components of the retina in digital eye fundus images is of vital importance for the development of an automated diabetic retinopathy diagnosis system.

Taking all these considerations into account, it can be concluded that the construction of a system for the automated diagnosis of diabetic retinopathy may be structurally divided into the four phases defined as follows:

1. Image quality assessment: here, algorithms to check whether an image has enough quality to be processed are implemented.

2. Segmentation of the main anatomical components of the retina: in this phase, algorithms for the automated segmentation of the optic disc, vascular tree and macula are developed.

3. Detection of lesions produced by diabetic retinopathy: this phase deals with the development of algorithms for the automated detection of the main lesions produced by diabetic retinopathy: microaneurysms, haemorrhages, exudates and cotton wool spots.

4. Implementation of an expert system: this phase consists on the implementation of an expert system for the diagnosis of diabetic retinopathy. This system, taking the results obtained in the previous phases, makes diagnoses based on international gradings of the illness.

This thesis is focused on the second of the defined phases. Here, novel high accurate methodologies for the segmentation of the optic disc, vascular tree and macula are proposed and described. Therefore, the work presented in this thesis constitutes the successful phase closing of the development of a comprehensive system for the early automated detection of diabetic retinopathy.

1.2 Structure of the thesis

This thesis is focused on the segmentation of the main anatomical retinal components in digital eye fundus images: the optic disc, vascular tree and macula. It is structured in seven chapters and two appendices according to the distribution described below.

Besides this **first chapter** of initial approach, the **second chapter** introduces the medical background required for an adequate understanding of this thesis. The anatomy and function of the eye, making special emphasis in the retina, are studied here. Moreover, diabetic retinopathy is studied in detail, i.e., causes and nature of the illness, incidence, diagnosis and treatment. After this, foundations of the automated detection of diabetic retinopathy are described and a review of the main published prototypes of systems for the automated diagnosis of the illness is offered. Discussion and conclusion on the results shown by these prototypes motivate the research work presented in the following chapters.

The **third chapter** presents a new template-based methodology for segmenting the optic disc in digital retinal images. The segmentation methodology starts extracting an OD-containing subimage defined by a pixel located within the optic disc and its surrounding region. For this purpose, a location methodology based on a voting-type algorithm is also proposed in this chapter. Then, in this subimage, the circular OD boundary is extracted by means of morphological and edge detection techniques and using the Circular Hough Transform.

The **fourth chapter** presents a new supervised methodology for blood vessel segmentation in digital retinal images. This method uses a neural network (NN) scheme for pixel classification and computes a 7-dimensional vector composed of gray-level and moment invariants-based features for pixel representation. This feature vector is computed from preprocessed retinal images in the neighborhood of the pixel under consideration. Real-valued classification results are then thresholded to classify each pixel into two classes: vessel and non-vessel. Finally, a postprocessing fills pixel gaps in detected blood vessels and removes falsely-detected isolated vessel pixels.

The **fifth chapter** presents a methodology to identify the macular zone in eye fundus images. Firstly, an estimate of the fovea center location is obtained making use of its a priori-known positional features with respect to the other anatomical retinal structures. To this effect, the methodologies presented in the previous chapters to segment the optic disc and vascular tree are used. Then, the obtained approximate fovea location is refined if this is detectable using morphological processing. Finally, the macular zone is identified using a polar coordinate system centered on the obtained fovea center location.

The **sixth chapter** proposes a new function for evaluating the global quality of retinal vessel segmentations. This is based on the evaluation of three main measurable features

which describe vascularity: vascular tree connectivity, area and length. This quality evaluation function represents a natural approach for evaluating vascular segmentations and, since it is based on evaluating descriptive features of the vascular tree, it provides good result interpretability to easily detect concrete features (connectivity, area or length) of the vascular tree inappropriately segmented.

The main body of this thesis ends with the **seventh chapter**. This chapter summarizes the presented work and projects future work derived from this thesis, i.e., new developments, prospects, etc.

This book ends with two appendices. The **first appendix** offers a technical description of all image databases used in the presented work. On the other hand, the **second appendix** offers the necessary theory for a complete understanding on performance evaluation based on contingency tables for binary classification, which is widely used in several chapters.

1.3 Framework and publications of the thesis

The aim of this section is to establish the framework in which this thesis has been developed as well as to remark the main results obtained in this research work. With respect to the framework of the thesis, the project that has funded this work and the predoctoral stay carried out by this doctoral student in an international research center are presented. On the other hand and regarding research results, the main publications generated by this thesis in international journals and conference proceedings are detailed. Summarizing, this section attempts to make a review of the most relevant milestones related to the presented thesis.

Research projects

- **Title:** Expert System for Early Automated Detection of diabetic retinopathy by Analysis of Digital Retinal Images.
  - **Funding body:** Health Ministry of the Andalusian Regional Government (Andalusia, Spain).
  - **Participating entities:** University of Huelva and Andalusian Health Service.
  - **Project description:** the final aim of the project is to develop a system for the early detection of diabetic retinopathy by means of digital eye fundus image

analysis. The Health Ministry of the Andalusian Regional Government launched a comprehensive plan for the prevention of diabetes, which include protocols for the early detection of diabetic retinopathy. These protocols include periodical examinations of the retina carried out by family doctors in the primary health care centers of the Andalusian Health Service. The system under investigation, taking advantage of the net of interconnected retinographs in the primary health care centers, aims to improve preventive illness protocols by helping family doctors to filtrate with higher confidence those cases needing revision and treatment by ophthalmologists and those that do not.

  - 1**st project phase:**
    - **Duration:** from 01/01/2007 to 31/12/2009.
    - **Funding amount:** 150000 €.
    - **Principal researcher:** Dr. José Manuel Andújar.
    - **Phase aim:** the concrete aim of this first phase of the project is to develop algorithms for the robust segmentation of the main anatomical components of the retina.
  - 2**nd project phase:**
    - **Duration:** from 01/01/2011 to 31/12/2013.
    - **Funding amount:** 180000 €.
    - **Principal researchers:** Dr. Manuel E. Gegúndez and Dr. Diego Marín Santos.
    - **Phase aim:** algorithms for the detection of the main lesions produced by diabetic retinopathy and an expert system for its diagnosis are investigated and implemented in this phase.

- **Title:** Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology (MESSIDOR).
  - **Funding body:** French Ministry of Research and Defense (France).
  - **Participating entities:** ARMINES : Centre de Morphologie Mathématique, Laboratoire L3I - Universite de la Rochelle, Département ITI - LaTIM INSERM U650, Laboratoire SIC - CNRS FRE 2731, Nottingham Trent University: School

of Computing and Informatics, EA 3063 / Ophtalmologie - Faculte de Medecine, Service D'Ophtalmologie Daviel - CHU Brest, Service D'Ophtalmologie - Hopital Lariboisiere, ADCIS and CRIHAN.

- **Project description:** within the scope of diabetic retinopathy, the primary purposes of the MESSIDOR project is to compare and evaluate both:
  * The various segmentation algorithms developed for the detection of lesions present in color retinal images.
  * The tools to index and manage the image databases.

Predoctoral stay

- **Center:** Center of Mathematical Morphology of the National High School of Mines.
- **Center address:** 35, rue Saint Honoré, 77305 Fontainebleau CEDEX (France).
- **Center director:** Dr. Fernand Meyer.
- **Stay supervisor:** Dr. Jean-Claude Klein.
- **Stay duration:** from 01/10/2009 to 15/03/2010.
- **Work team:** Dr. Jean-Claude Klein, Dr. Fernand Meyer, Dr. Etienne Decenciere and Dr. Estelle Parra-Denis.
- **Stay description:** the stay was divided into two phases. The first phase consisted on a period of education on *mathematical morphology* by means of a one-week intensive course called "Summer School of Mathematical Morphology 2009". The second phase, which lasted the rest of the stay, was focused on the investigation of drusen segmentation in eye fundus images using mathematical morphology. Drusen are yellowish or whitish deposits of extracellular material, whose main origin is the age-related macular degeneration, an illness caused by a deterioration of the central portion of the retina. As a result of this work, a library called "OpthalMorphoTools", implemented in the *Morph-M* environment[1], was created. This library implements a set of morphological operators for analysis of digital retinal images and drusen enhancing.

[1] Center for Mathematical Morophology, Industry section, Morph-M : image processing software specialized in mathematical morphology, France [Online]:`http://cmm.ensmp.fr/Morph-M/index_en.html`

List of Publications

- **In journals:**

1. **Authors:** Marín, D., Aquino, A., Gegúndez, M.E., Bravo, J.M.
   Paper Title: A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features.
   Journal Title: IEEE Transactions on Medical Imaging.
   Publication details: volume 30, issue 1, pages 146-158.
   Year: 2011.
   Quality indicators: the journal is included in five categories of the Journal Citation Reports raking. These categories are: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* and *Radiology, Nuclear Medicine & Medical Imaging*. With an impact factor of 3.540 in 2009, the journal is included in the first quartile of all categories. Its best relative position is 9 of 245 in the *Engineering, Electrical & Electronic* category and its best absolute position is 2 in the *Imaging Science & Photographic Technology* category.

2. **Authors:** Aquino, A., Gegúndez, M.E., Marín, D.
   Paper Title: Detecting the Optic Disc Boundary in Digital Fundus Images Using Morphological, Edge Detection and Feature Extraction Techniques.
   Journal Title: IEEE Transactions on Medical Imaging.
   Publication details: volume 29, issue 11, pages 1860-1869.
   Year: 2010.
   Quality indicators: the journal is included in five categories of the *Journal Citation Reports* raking. These categories are: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* and *Radiology, Nuclear Medicine & Medical Imaging*. With an impact factor of 3.540 in 2009, the journal is included in the first quartile of all categories. Its best relative position is 9 of 245 in the *Engineering, Electrical & Electronic* category and its best absolute position is 2 in the *Imaging Science & Photographic Technology* category.

3. **Authors:** Aquino, A., Gegúndez, M.E., Marín, D.
   Paper Title: Automated Optic Disc Detection in Retinal Images of Patients

with Diabetic Retinopathy and Risk of Macular Edema.

Journal Title: International Journal of Medicine and Medical Sciences.

Publication details: volume 1, issue 2, pages 87-92.

Year: 2010.

Quality indicators: the journal is edited by the *World Academy of Science, Engineering and Technology* and indexed by: *SCIRUS*, *Google Scholar*, *EBSCO*, *GALE*, *INTUTE*, *DOAJ*, *Ulrich's & Serials Solutions* and *Electronic Journals Library*.

4. **Authors:** Aquino, A., Gegúndez, M.E., Bravo, J.M., Marín, D.

   Paper Title: A Specialized Function for Quality Evaluation of Retinal Vessel Segmentations.

   Journal Title: IEEE Transactions on Medical Imaging.

   Publication details: under major revision.

   Year: -.

   Quality indicators: the journal is included in five categories of the *Journal Citation Reports* raking. These categories are: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* and *Radiology, Nuclear Medicine & Medical Imaging*. With an impact factor of 3.540 in 2009, the journal is included in the first quartile of all categories. Its best relative position is 9 of 245 in the *Engineering, Electrical & Electronic* category and its best absolute position is 2 in the *Imaging Science & Photographic Technology* category.

5. **Authors:** Aquino, A., Gegúndez, M.E., Marín, D., Bravo, J.M.

   Paper Title: Automated Identification of the Macular Zone in Eye Fundus Images Using Anatomical-Based and Visual-Based Features.

   Journal Title: IEEE Transactions on Medical Imaging.

   Publication details: under revision.

   Year: -.

   Quality indicators: the journal is included in five categories of the *Journal Citation Reports* raking. These categories are: *Computer Science, Interdisciplinary Applications*; *Engineering, Biomedical*; *Engineering, Electrical & Electronic*; *Imaging Science & Photographic Technology* and *Radiology, Nuclear*

Medicine & Medical Imaging. With an impact factor of 3.540 in 2009, the journal is included in the first quartile of all categories. Its best relative position is 9 of 245 in the *Engineering, Electrical & Electronic* category and its best absolute position is 2 in the *Imaging Science & Photographic Technology* category.

- **In Conference Proceedings:**

1. **Authors:** Aquino, A., Gegúndez, M.E., Bravo, J.M., Marín, D.
   Paper Title: A Similarity Function for Global Quality Assessment of Retinal Vessel Segmentations.
   Conference Title: International Conference on Pattern Recognition and Computer Vision, Bali (Indonesia).
   Publication details: volume 67, pages 7-11, ISSN 1307-6892.
   Year: 2010.
   Quality indicators: conference organized by the *World Academy of Science, Engineering and Technology*. Proceedings indexed by: *Google Scholar*, *IE Compendex*, *INTUTE*, *GALE*, *EBSCO*, *Serials Solutions*, *DOAJ* and *Electronic Journals Library*.

2. **Authors:** Aquino, A., Gegúndez, M.E., Marín, D.
   Paper Title: Automated Optic Disc Detection in Retinal Images of Patients with Diabetic Retinopathy and Risk of Macular Edema.
   Conference Title: International Conference on Machine Vision, Image Processing and Pattern Analysis, Bangkok (Thailand).
   Publication details: volume 60, pages 85-90, ISSN 1307-6892.
   Year: 2010.
   Quality indicators: conference organized by the *World Academy of Science, Engineering and Technology*. Proceedings indexed by: *Google Scholar*, *IE Compendex*, *INTUTE*, *GALE*, *EBSCO*, *Serials Solutions*, *DOAJ* and *Electronic Journals Library*.

3. **Authors:** Aquino, A., Gegúndez, M.E., Marín, D., Andújar, J.M.
   Paper Title: A Robust, High Accuracy and Fast New Methodology to Locate the Optic Disc From Colour Retinal Images.
   Conference Title: IADIS International Conference Computer Graphics, Visua-

lization, Computer Vision and Image Processing, Algarve (Portugal).
Publication details: pages 99-105, ISBN 978-972-8924-84-3.
Year: 2009.
Quality indicators: conference organized by the *International Association for Development of the Information Society.*

4. **Authors:** Aquino, A., Gegúndez, M.E., Marín, D., Andújar, J.M.
Paper Title: Sistema Experto para la Detección Precoz de la Retinopatía Diabética Mediante Análisis de Imágenes Digitales de Retina: Primeros Resultados.
Conference Title: XXIX Jornadas de Automática, Tarragona (Spain).
Publication details: volume 29, pages 1-10, ISBN 978-84-691-6883-7.
Year: 2008.
Quality indicators: conference organized by the *Spanish Committee of Automatic.*

- **Patents:**

1. **Authors:** Gegúndez, M.E., Aquino, A., Bravo, J.M., Marín, D.
Patent Title: Sistema y Procedimiento de Obtención de Imágenes de Alto Contraste del Árbol Vascular del Fondo de Ojo.
Patent details: spanish patent, application number P200930460.
Year: -.

Chapter 2

Medical Background and Motivation

DIABETIC retinopathy is a chronic disease, caused by complications of diabetes mellitus , which nowadays constitutes the primary cause of blindness in people of working age in the developed world. Although diabetic retinopathy is not a curable disease, laser photocoagulation can slow down the progression to blindness if the illness is detected in its early stages. In spite of it, this is not an easy task because people affected by diabetic retinopathy do not perceive symptoms until visual loss develops and this usually happens in the later stages of the disease, when treatment is less effective. Therefore, to ensure that treatment is received in time, diabetic patients need to undergo a yearly eye fundus examination . However, this preventive action involves a huge challenge for health systems due to the huge number of patients needing ophthalmologic revision, thus preventing many patients from receiving adequate treatment. Since diabetic retinopathy is diagnosed by means of the analysis of digital retinal images, this problem could be alleviated by taking advantage of a computerized approach. This is, a system to filtrate automatically cases of patients not affected by the disease or, even, to grade automatically the stage of the illness, would reduce the specialists' workload and increase the effectiveness of preventive protocols and early therapeutic treatments.

This chapter, structured in three sections, establishes the medical background and motivation of this thesis. The **first section** provides a general anatomical and functional description of the eye. Moreover, once the retina is briefly described as an eye component, it is then studied more in detail.

The **second section** closes the medical background of the chapter studying diabetic retinopathy. The origin, incidence, nature, diagnosis, treatment and classification of the illness are treated in this section.

The **third section** deals with the automated diagnosis of diabetic retinopathy by means of automated systems. These systems are described and their benefits are presented and discussed. Finally, the main prototypes present in the literature are reviewed and their current weaknesses are identified to motivate and frame the work presented in this thesis.

2.1 The eye

This section is divided into three subsections studying the structure and composition of the eye as well as its function in the process of vision. The first subsection provides a general description of the eye, defining its main features and components. Since the retina is a central subject of this thesis, the second subsection is focused on its study, extending the brief description introduced in the first subsection. Finally, the third subsection summarizes the process of vision, making special emphasis in the eye's role.

2.1.1 General description of the eye

The eye or eyeball (Figure 2.1 shows a diagram with all the eye's components described along this section) is one of the main parts of the vision system. It has an irregular spherical form and its front part, the cornea, constitutes a prominence with form of spherical segment whose radius is lower than the radius of the rest of the eyeball. The human eye weighs 7 gr, has a volume of 6.5 ml and its axial length (distance between the anterior and posterior pole) is 24 mm.

The eyeball is composed of a content and a wall:

- The content: it is formed by the transparent material in the eye.

  The most important component of the eye's content is the vitreous gel, which is a substance of gelatinous consistency. It fills two third parts of the eye's posterior pole, it is into contact with the retina in all its extension and it is limited at the front by the lens.

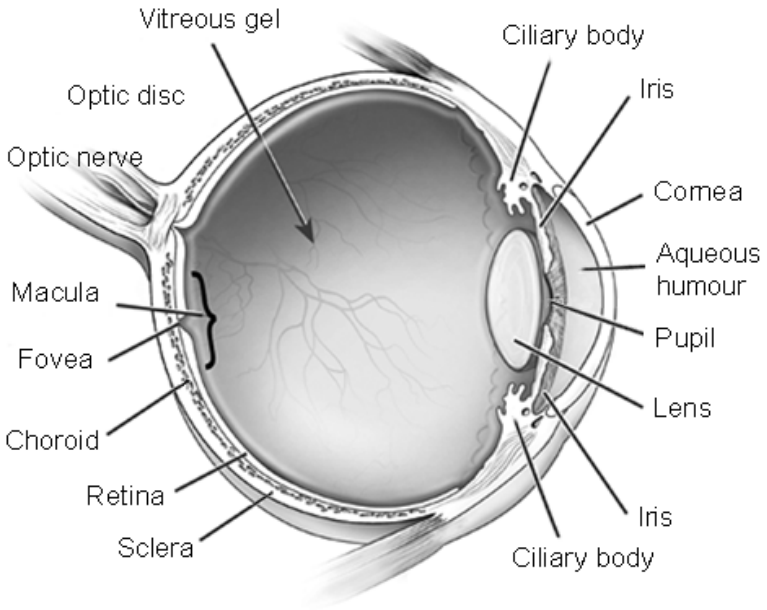


Figure 2.1: *Basic anatomy of the eye.*

The eye lens is a biconvex crystalline lens located between the iris and the vitreous gel. Its convergence capacity is variable and it allows to project images on the retina.

The aqueous humour, a thick watery substance, fills the space between the lens and the cornea.

- The wall: it is composed of three concentric membranes.

  The external membrane is a fibrous and resistant cover called sclera. It is unextending, poorly vascularized and it has pearly white color. The sclera forms four fifth parts of a hollow sphere. The resting one fifth part of this sphere is represented by the cornea, which is a transparent and curved surface fixed to the anterior portion of the sclera. The hole through which the optic nerve goes out the eye towards the brain is in the posterior pole of the sclera.

  The middle membrane, vascularized and with connective tissue, is called uvea and is formed by the choroid, ciliary body and iris. The posterior part of the uvea is the choroid, a richly vascularized membrane which is mainly responsible of feeding the retina. The ciliary body and the iris are at the anterior pole. The ciliary body secretes the aqueous humour and contains the ciliary muscle, which changes the shape of the lens when the eye focuses on an object. It lies just behind the iris, a circular colored

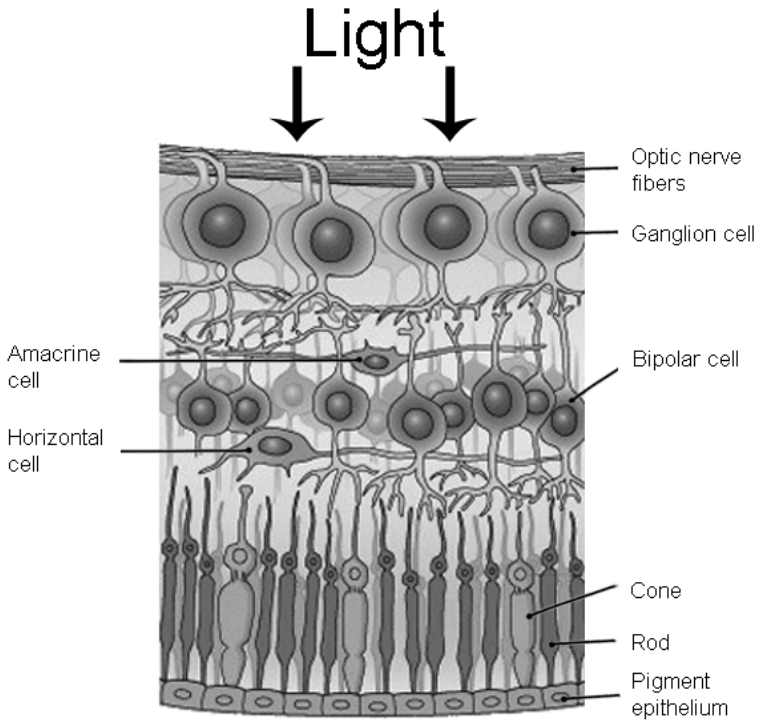


Figure 2.2: *Cross-sectional retina diagram.*

structure responsible for controlling the diameter and size of the pupil. The pupil is the hole in the iris through which the light enters the eye.

The internal sensorial membrane is the retina, the responsible of transforming light into nerve signals to be interpreted by the brain. The main parts visible on the retinal surface are the optic disc, macula and vascular tree. For the aims of this thesis, it is interesting to study in depth this membrane of the eye. Therefore, the next subsection offers a detailed description of the retina.

2.1.2 The retina

The retina is the main eye's membrane involved in the process of vision. It is approximately 0.5 mm thick, lines the internal back of the eye and is composed of a set of layers, delimiting at the outside with the choroid and at the inside with the vitreous gel. The retina is smooth and semi-transparent; it modifies its opacity when is exposed to direct sunlight.

Figure 2.2 shows a cross-sectional diagram of the retina showing its main components. These components are described as follows:

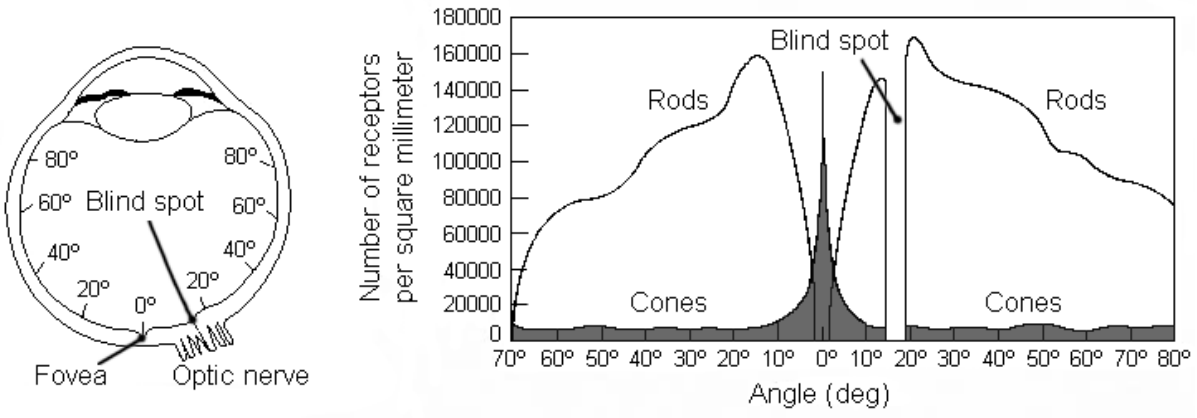


Figure 2.3: *Rods and cones distribution on the retina.*

- Pigment epithelium: it is attached to the underlying choroid, so this layer constitutes the outer blood-retinal barrier. Its main function is to nourish the photoreceptors.

- Photoreceptors: the retinal photoreceptors are a specialized type of neuron capable of converting light into chemical messages, and are attached to the pigment epithelium. Two types are distinguished, rods and cones:

  - Rods are very light sensitive cells, responsible for night vision but, in spite of this sensitivity, they are not sensitive to color. These cells are absent from the fovea (see its detailed description below) and at a few degrees away from its center their density rises to a high value and spreads over a large area of the retina (see Figure 2.3). This is why rods are mainly used for peripheral vision.

  - Cones are less light sensitive than rods but they are sensitive to color; there are three types of cones, each type absorbing either red, green, or blue light preferentially. These photoreceptors are responsible for day vision and are also in charge of detailed vision, since they are able to perceive finer detail and more rapid changes in images due to their response times to stimuli are faster than those of rods. Since, contrary to rods, cones are densely found and mostly packed in the fovea (see Figure 2.3), cones are used in central vision. Cones are much less numerous than rods (1:20 ratio).

- Amacrine, horizontal and bipolar cells: these types of neurons are located between the ganglion cells and the photoreceptors. They transmit information and also combine signals from several photoreceptors. This combination is carried out in a detailed way

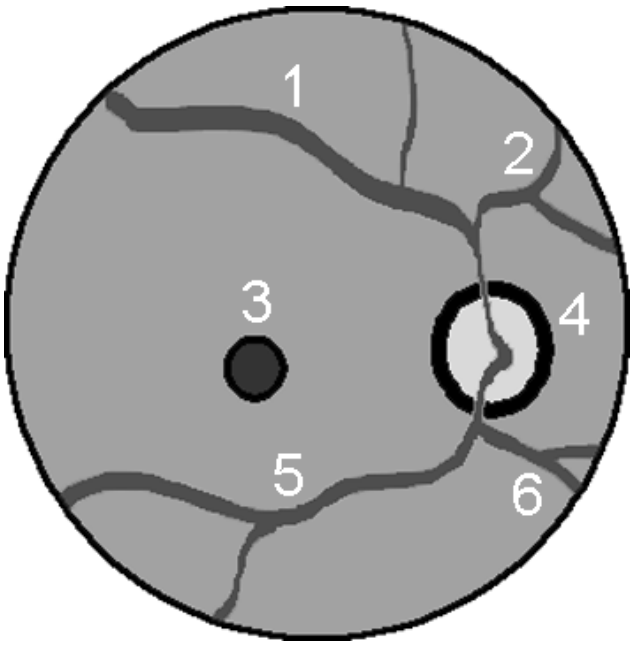


Figure 2.4: *Diagram of the retinal surface: (4) Optic disc. (3) Macula center (fovea). (1), (2), (5) and (6) Retinal vascular tree: (1) superior temporal blood vessels, (2) superior nasal blood vessels, (5) inferior temporal blood vessels and (6) inferior nasal blood vessels.*

on the precise spatial and temporal patterns of light which the ganglion cells critically rely on. Bipolar cells receive input from the cones, and many feed into the retinal ganglion cells. Horizontal cells link cones and bipolar cells, and amacrine cells, in turn, link bipolar cells and retinal ganglion cells.

- Ganglion cells: these are the final output neurons of the retina. Ganglion cells collect information from bipolar cells and amacrine cells. This information is in the form of chemical messages, which are transformed here into intracellular electrical signals.

- Optic nerve fibers: these fibers collects all the axons of the ganglion cells to transport electrical signals to the brain visual centers.

Studying now the retinal surface, its frontal view shows three main anatomical features: the optic disc (number (4) in Figure 2.4), macula (the macula center, called the fovea, is number (3) in Figure 2.4) and vascular tree (numbers (1), (2), (5) and (6) in Figure 2.4). The description of these retinal components is the following:

- Optic disc: the optic disc (see diagram in Figure 2.5 to follow the description), also called the optic nerve head or papilla, is located on the nasal side of the macula and represents the location where the nerve fibers come together, form the optic nerve, and

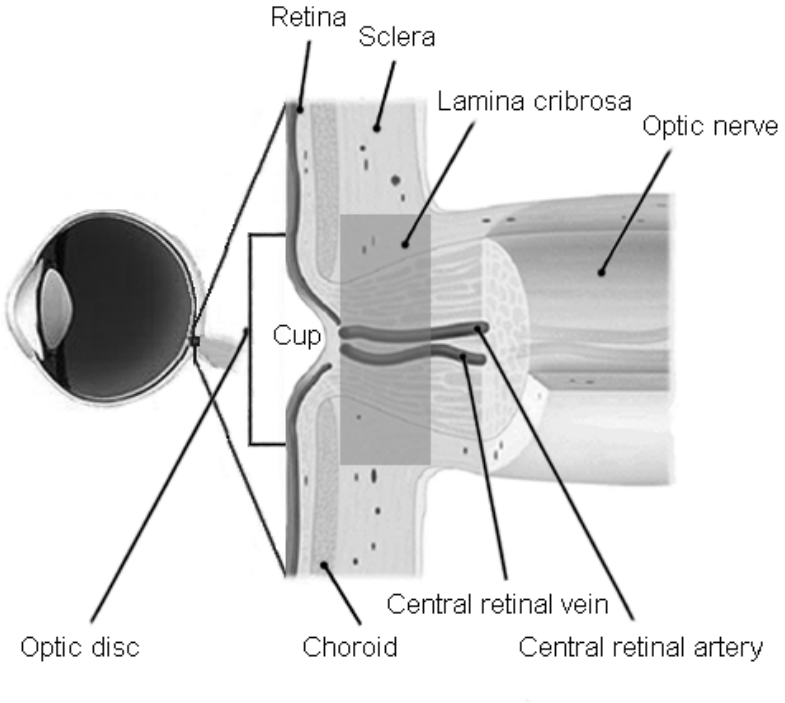


Figure 2.5: *Cross-sectional optic nerve diagram.*

exit the eye. It is slightly elliptical in shape, with a diameter varying between 1.5 and 2 mm. The optic disc is also the entrance and exit site of the blood vessels that feed the retina. There are no photoreceptors in the OD (Figure 2.3), so it represents the blind spot in the retina. The optic nerve fibers do not occupy the whole area of the optic nerve head as they are located at the optic nerve head border, which leads to an excavation in the center. This excavation resembles a small "cup". When is visible, the cup appears in white color as result from exposure of the hyper reflective tissue, the lamina cribrosa. The lamina cribrosa is a continuation of the sclera and provides support to nerve fibers and blood vessels.

- Macula: the macula is a round area in the central region of the retina, which measures about 3 to 4 mm in diameter. The macula contains mainly cones and only some rod, so this area is the one providing the most distinct vision and is responsible of central vision. There is a small depression in the centre of the macula measuring about 1 mm in diameter and visible as a round dark area called the fovea (see diagram in Figure 2.6). As mentioned above, the fovea contains the largest concentration of cone cells and do

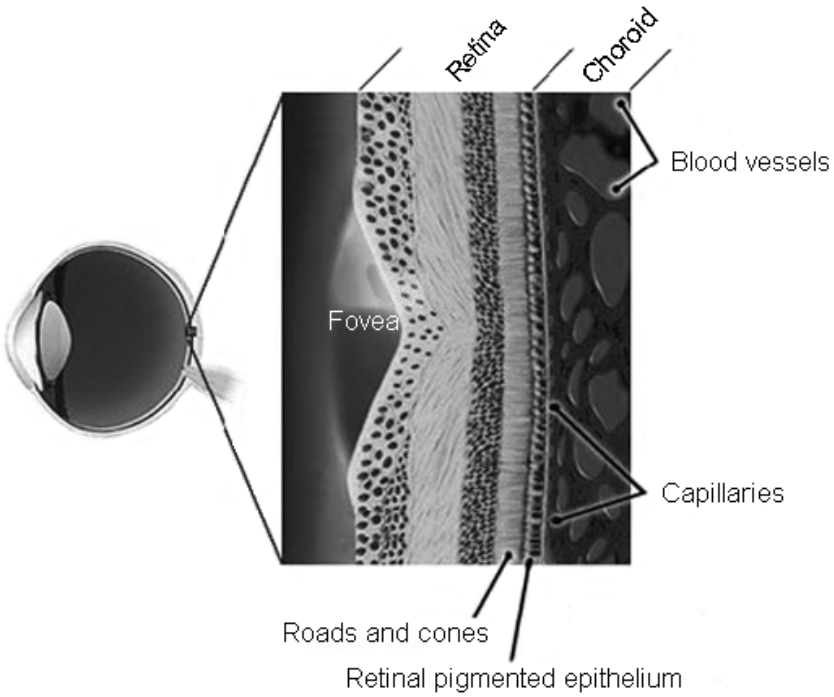


Figure 2.6: *Cross-sectional macula diagram*

not contain rods, so it is responsible for our sharp central vision and it is also the main portion of the retina used for color discrimination.

- Vascular tree: the retinal vascular tree is a vascular tree-like connected structure which is in charge of nourishing the inner parts of the retina. This vascular structure originates in the optic disc with the central retinal artery and vein (see Figure 2.5), which progressively split into many other smaller vessels for covering the retina. Structurally, the retinal vascular tree can be divided into four type of vessels: superior temporal blood vessels (number (1) in Figure 2.4), superior nasal blood vessels (number (2) in Figure 2.4), inferior temporal blood vessels (number (5) in Figure 2.4) and inferior nasal blood vessels (number (6) in Figure 2.4).

A digital image of the retinal surface showing the optic disc, macula and vascular tree is presented in Figure 2.7.

2.1.3 The process of vision

A simplified description of the process of vision can be summarized as follows (see Figure 2.8 to follow the description):

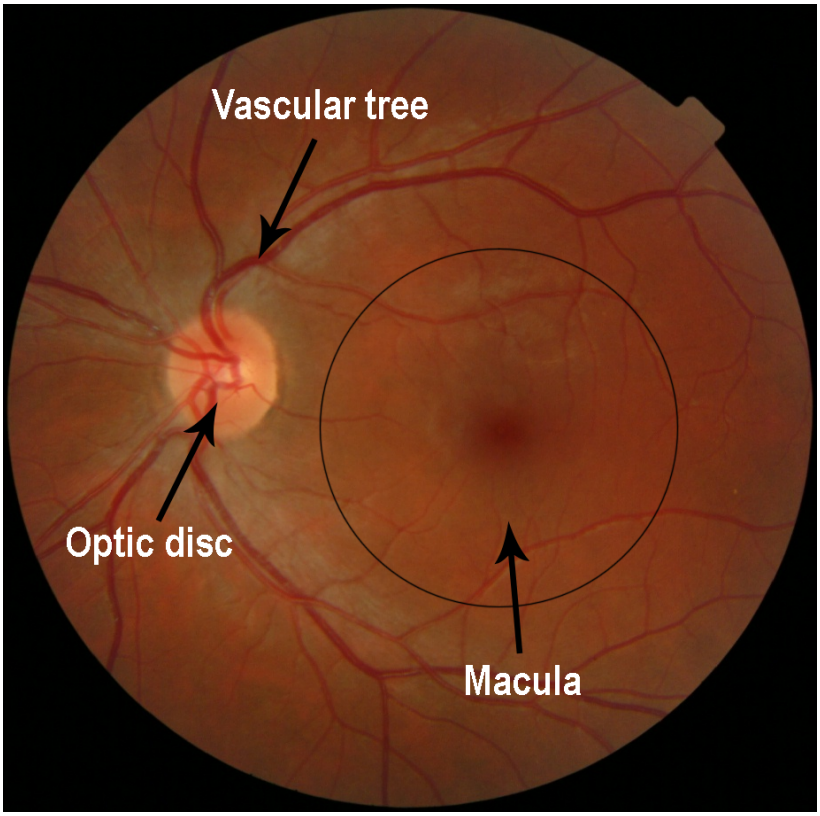


Figure 2.7: *Digital retinal image showing the optic disc, macula and vascular tree.*

1. Light waves from an object enter the eye through the cornea and then progresses through the pupil. The amount of light entering the eye is controlled by the iris, which controls the size of the pupil.

2. Then, the light waves are bent or converged first by the cornea, and then further by the crystalline lens to a nodal point (N) located immediately behind the back surface of the lens. At that point, the image becomes reversed (turned backwards) and inverted (turned upside-down).

3. Light continues through the vitreous gel to the retinal surface, which turns more or less transparent according to the incident amount of light.

4. Light passes through the different layers of the retina to photoreceptors, which are excited and turn light into chemical messages. These messages travel in the opposite direction to ganglion cells, which ultimately change chemical messages into electrical signals.

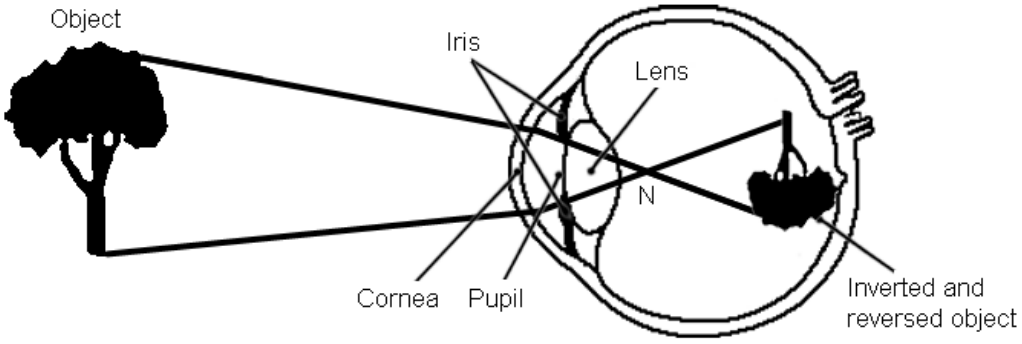


Figure 2.8: *Simplified diagram of the process of vision.*

5. Finally, electrical signals from the ganglion cells are collected by the optic nerve fibers and sent through the optic nerve, along the visual pathway, to the occipital cortex at the posterior of the brain. Here, the electrical signals are interpreted or "seen" by the brain as a visual image.

2.2 Diabetic retinopathy

Metabolism is the process of converting the food we eat into energy. It begins when food is broken down during digestion to create glucose, which is the main source of fuel for the body. Glucose metabolism in the body is controlled by insulin, a hormone produced by the pancreas. When glucose passes into the bloodstream, insulin allows it to get into the cells.

Diabetes mellitus, commonly known as diabetes, is a metabolic disorder that interferes with the body's ability to use and store glucose. The determining factor in diabetes is insulin, and people with diabetes present blood glucose level elevation as a consequence of low insulin production, abnormal cell response to insulin or both reasons.

There are two major types of diabetes:

- Type I or insulin-dependent diabetes: it is sometimes referred to as juvenile diabetes because it most often begins in childhood (although it may also occur in adults). In this disease, the body makes little or no insulin, so daily injections of this hormone are needed.

- Type II or adult-onset-diabetes: it is more common than type I. It usually occurs in adulthood, but young people are increasingly being diagnosed with this disease. In

this diabetes type, the pancreas does not make enough insulin to keep blood glucose levels normal, often because the body does not respond well to the hormone.

The excess of glucose in blood produced by diabetes damages the macro- and microvascular system and usually the eye is one of the first places where this becomes apparent. Diabetic Retinopathy (DR) is the retinal disease derived from complications caused by the abnormal high level of glucose in blood produced by diabetes and is the most frequent microvascular complication of this disease. When the microvascular system in the eye is progressively damaged, vision loss and, ultimately, blindness can occur. Actually, DR is nowadays the leading ophthalmic pathological cause of blindness among people of working age in developed countries (Taylor and Keeffe, 2001).

All people affected by diabetes are at risk for the development of DR and most of them will eventually develop the illness. In fact, although the presence of diabetes does not necessarily involve vision impairment, about 2% of the patients affected by this disorder are blind and 10% undergo vision degradation after 15 years of diabetes (Klein et al., 1995; Massin et al., 2000) as a consequence of DR complications. The estimated prevalence of diabetes for all age groups worldwide was 2.8% in 2000 and 4.4% in 2030, meaning that the total number of diabetic patients is forecasted to rise from 171 million in 2000 to 366 million in 2030 (Wild et al., 2004). This figures of incidence make DR, besides a great public health problem, a great economic problem for Administrations since for instance, only in U.S., cost of ophthalmic chronic complications caused by diabetes exceeded 1 billion dollars in 2007 (American Diabetes Association, 2008).

As said before, DR damages the eye's vascular system due to the abnormally high level of glucose in blood. This vascular damage consists on the deterioration of the vessel endothelium, resulting this in vessel permeability and fragility increasing. As a consequence, with the progression of the illness and this damage, a set of characteristic and univocal clinical signs of the illness can appear on the retinal surface. Accurate assessment of DR requires the ability to detect and identify the clinical features described as follows (see Figure 2.9 for recognizing each feature described below):

- Microaneurysms: these are the first manifestation of DR. Microaneurysms are capillary dilations produced by local weakening of capillary walls that make them to emerge in a balloon-like fashion. They are visible on the retinal surface as tiny round dark red dots and their size varies from 10 to 100 microns in diameter.

- Haemorrhages: the disease and damage progression can produce the burst of microaneurysms, which leaks blood and therefore causes haemorrhages. Retinal haemorrhages can be discernable either as small red dots or blots indistinguishable from microaneurysms or as larger red flame-shaped patches.

- Exudates: besides the leaking of blood, vessels also leak lipids and proteins causing the appearance of bright spots called exudates. These are typically manifested as yellowish/white patches of varying sizes, shapes and locations. The leaking of fluid from vessels can progress producing large exudative areas in the macula, which can provoke swelling of this. This complication is called diabetic macular edema and can seriously affect central vision.

- Cotton wool spots: with the progression of DR and its associated lesions, parts of the retina can be deprived of blood and consequently can suffer ischemia. These ischemic areas are called cotton wool spots and are visible on the retina as fluffy whitish blobs.

- Neovascularization: the appearance of ischemic areas in the retina provokes as response the growing of new vessels to supply the retina with more oxygen. These new vessels are called neovascularizations and have a greater risk of rupturing and causing larger haemorrhages than normal vessels.

In the later stages of DR, as a consequence of the great produced damage, the retina can detach from its supporting tissues, phenomenon known as retinal detachment. This is a serious situation what can lead to permanent vision loss if it is not repaired shortly.

DR is not a curable disease, but it can be treated and controlled if it is detected in early stages (Lee et al., 2001; Taylor and Keeffe, 2001). Currently, treatment of DR is still predominantly based on laser photocoagulation. It consists on the application of a strong light beam to certain areas on the retina. For instance, the laser can be directly applied to leaking microaneurysms to prevent further hemorrhaging. Another example is that it can also be applied in a grid pattern over a larger part of the retina with the purpose of reducing the overall need for oxygen and diminishing the load on the damaged microvasculature.

Laser photocoagulation can prevent major vision loss, however, it can not restore visual acuity already lost. Therefore, the detection of the illness in early stages is of vital importance to apply the treatment in the most effective way possible. The main problem of this early illness detection is that DR patients perceive no symptoms until visual loss develops, and this

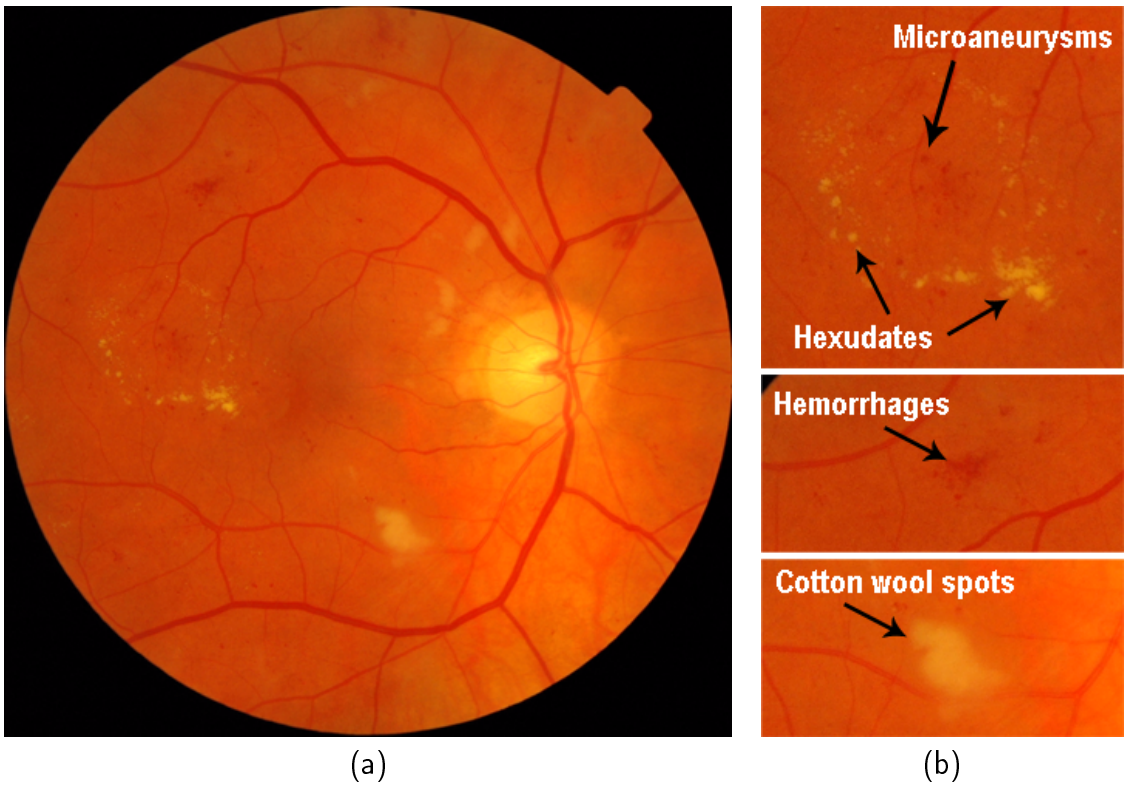


Figure 2.9: *Examples of the most common lesions produced in the retina by DR:(a) Photography of the retinal surface. (b) Close ups of zones of image (a) showing examples of microaneurysms, haemorrhages, exudates and cotton wool spots.*

usually occurs in the later disease stages. This has motivated a widely accepted preventive protocol based on the periodical retinal examination of diabetic patients (typically yearly examination) (Bresnick et al., 2000; Early Treatment Diabetic Retinopathy Study Research Group, 1991a; Fong et al., 2003; Kinyoun et al., 1989), which is the population potentially at risk. Concretely, this consists on screening sessions where the retinas of both eyes are examined by physicians looking for early signs of DR. If any sign of the illness is found, the patient is referred to an ophthalmologist who decides the necessary treatment. Otherwise, the patient continues attending to the periodical examinations.

The retina is the only part of the human body where the blood vessels can be directly seen in a non-invasive manner (Rassam et al., 1994). In the context of DR detection it is carried out by means of mainly two types of screening methods: ophthalmoscopy and retinal photography. Ophthalmoscopy consists on the direct examination of the retina through the pupil by means of a portable instrument, called ophthalmoscope, composed of a light source and a set of lenses (2.10, image (a)). On the other hand, retinal photography is the

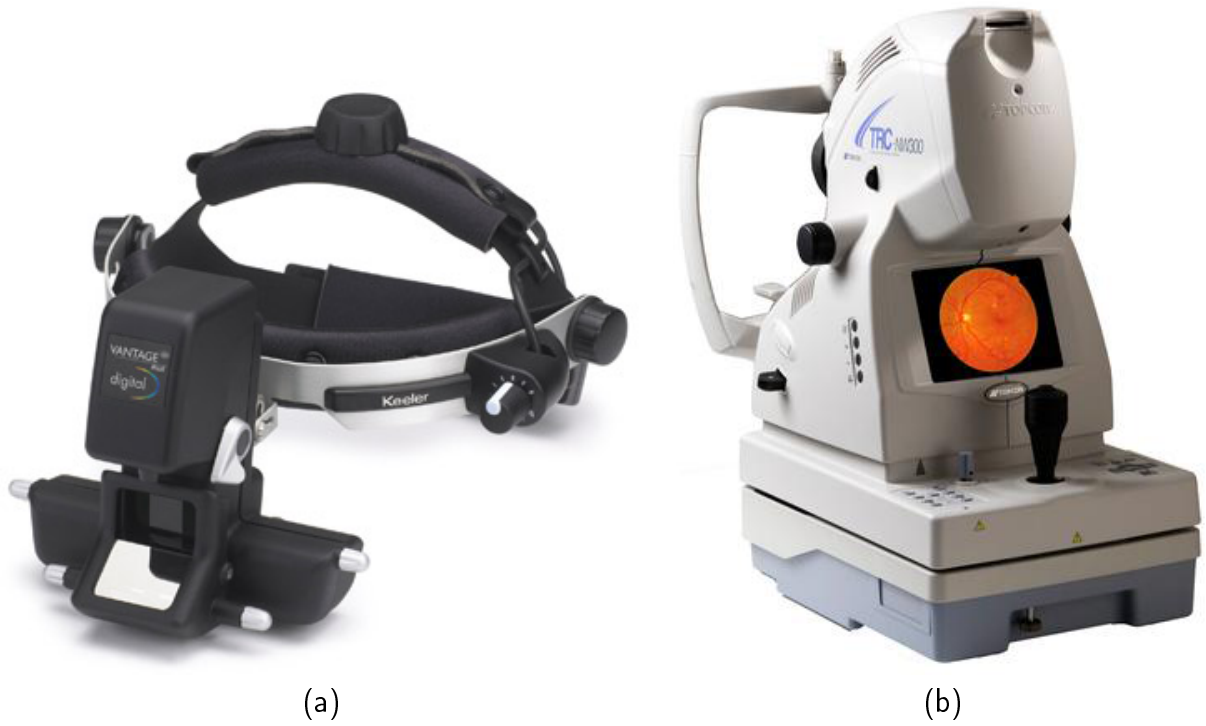


Figure 2.10: *Main instruments used for retinal inspection: (a) Indirect ophthalmoscope manufactured by Keeler Instruments Inc. (b) Non-Mydriatic retinograph manufactured by Topcon Corporation.*

procedure of taking digital photos of the retinal surface (retinographies) by using special CCD based digital cameras, called retinographs , with special optics that allow to make color photographs of the eye fundus through the pupil (2.10, image (b)). Since the pupil is narrow and does not allow much light to enter the eye, a good strategy to facilitate the retinal visualization is to lower the ambient light intensity. Besides this, ophthalmoscopy and retinal photography can also be performed with mydriasis (dilation of the pupil with eye drops). However, it is important to point out that mydriasis is not needed with modern retinographs in the majority of cases.

Retinal photography offers outstanding advantages for large scale retinal screening. This modality allows instantaneous examination of the retina and quick storage and access of the images. Besides this, since digital retinographs offer interfaces for connecting them to PCs, acquisition and interpretation stages of the screening can be decoupled. The images can be acquired anywhere and sent instantaneously to a specialized institution for their interpretation. An example of such a screening program is the Andalusian Comprehensive Plan for the Prevention of Diabetes (Aguilar Diosdado et al., 2009), which has motivated

the investigations presented in this thesis. This plan includes periodical retinal examination of the diabetic population. Thus, this large population is screened in primary health care centers of the Andalusian Health Service by physicians. Images are acquired on site and are sent over the internet to a central server. A trained reader specialized in DR will log into this server and grade the images. Patients needing to see an ophthalmologist or needing to attend the screening more frequently are notified.

Regarding classification of DR, two different approaches have emerged during last years: full disease classifications and population screening classifications. The main differences between both types of classifications are their aim and complexity. A brief description of both types of classifications is the following:

- Full disease classifications: classifications in this category are aimed to be used by ophthalmologists. They are complex classifications considering a high number of levels of DR to cover the full range of the illness exhaustively. The main classification in this group is the original Airlie House classification developed for the Early Treatment Diabetic Retinopathy Study (ETDRS) (Early Treatment Diabetic Retinopathy Study Research Group, 1991b). Other full disease classifications are mainly modifications of this. The cited classification distinguishes a great number of levels of DR severity grouped into two global stages of severity. The first stage, called Non-proliferative Diabetic Retinopathy (NPDR), reffers to the earliest retinal changes caused by DR. In this stage, a person affected by DR can present microaneurysms, retinal haemorrhages, exudates and cotton wool spots. Vision loss can develops in this stage as a consequence of macular edema or macular isquemia. On the other hand, the appearance of neovascularization marks the beginning of the second and severe stage of DR called Proliferative Diabetic Retinopathy (PDR). In this stage, haemorrhages in the vitreous and, ultimately, retinal detachment can occur.

- Population screening classifications: these classifications are designed to be used in large screening programs. They propose simplified gradings, with only a few levels of illness severity, easier than the previous classifications to be managed in the context of such programs. The main goal of these classifications is not to grade DR accurately, but to assess when a patient presents the illness in a stage sufficiently severe to merit referral for expert ophthalmological opinion and possible treatment. Typically, these

classifications are designed by countries depending on their needs to make viable their screening programs.

An example of this type of classification is the one used in chapters 3 and 5 of this thesis. This classification, designed within the framework of the Messidor - TechnoVision Project [1], grades DR into four stages and, additionally, proposes a classification to evaluate the risk of macular edema. The two simplified gradings were elaborated by experts from the recommendations provided by: Alfediam (French association for the study of diabetes and metabolic diseases) (Massin et al., 1996, 1997), ETDRS (Early Treatment Diabetic Retinopathy Study) (Early Treatment Diabetic Retinopathy Study Research Group, 1991b) and Diabetic Retinopathy Screening Services in Scotland (Diabetic Retinopathy Screening Services in Scotland, 2003). The descriptions of the proposed gradings are:

- DR grade:
  - 0 (Normal): $(\mu A = 0)$ and $(H = 0)$.
  - 1: $(0 < \mu A \leq 5)$ and $(H = 0)$.
  - 2: $((5 < \mu A < 15)$ or $(0 < H < 5))$ and $(NV = 0)$.
  - 3: $(\mu A \geq 15)$ or $(H \geq 5)$ OR $(NV = 1)$.

  where μA is the number of microaneurysms, H is the number of haemorrhages, and $NV = \{0, 1\}$ denotes the absence or presence of neovascularization respectively.

- Risk of Macular Edema:
  - 0 (No risk): no visible hard exudates.
  - 1: shortest distance between macula and hard exudates $>$ one optic disc diameter.
  - 2: shortest distance between macula and hard exudates $\leq$ one optic disc diameter.

[1] Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology - TECHNOVISION Project. Available: `http://messidor.crihan.fr/index-en.php`.

2.3 Automated detection of diabetic retinopathy

The benefits and cost-effectiveness of screening programs for early detection of DR are well recognized (Javitt et al., 1994). However, their effective development is a great challenge for health systems as a consequence of the great amount of human and material resources needed to put them into place (Mason, 2003). These difficulties and the generalized use of retinal photography in screening programs has motivated during the last 15 or 20 years an effort in the investigation of algorithms for the diagnosis of DR with sufficient accuracy to be able to implement them into these programs. If an automated system could be able to exclude a large number of those patients who have no DR, it would reduce the workload of the specialists and thus enhance preventive protocols. Furthermore, it would also result in economic benefits for public health systems, since cost-effective treatments associated to early illness detection lead to remarkable cost savings (American Academy of Ophthalmology, 2008; Lairson et al., 1992).

The construction of a generic comprehensive system for the diagnosis of DR can be structurally and logically divided into four general phases. These phases are described as follows:

1. Image quality assessment: since retinal images are taken using retinographs, the image acquisition is completely decoupled from the system. Many factors can affect negatively the acquisition process and can give as a result images without enough quality to be processed. For example, eye movement at the moment of retinography capture may lead to blurred images, or insufficient pupil dilation as a consequence of inappropriate illumination or inefficient mydriasis (if it is needed) may produce too dark and low contrasted images. With all these considerations, this phase defines the preconditions of the system and its entry point. This is, the way the system receives an image and the decision rules to assess if this has enough quality to be successfully processed, are designed and implemented in this phase.

2. Segmentation of the main anatomical components of the retina: the key of success for an automated system for the diagnosis of DR is the accurately detection and identification of lesions produced by DR (this is faced in the following phase). These lesions must be detected against a highly variable background retinal image that includes: the optic disc, the vascular tree, the macula, and distractors such as choroidal vessels, laser scars or light artifacts among other. Furthermore, other features inherent to the retina

such as its differences of pigmentation in different ethnic races increase this variability (Usher et al., 2003). This fact has motivated a widely used structured approach in which the main anatomical components in the retina are segmented and used as landmarks. Thus, their a priori-known anatomical features and relationships between them are exploited to increase confidence in the classification of each object. Besides this, the segmentation of the anatomical components are also needed for DR diagnosis in many classifications of the illness, such as the population screening classification detailed in the previous section and used in this thesis. Therefore, taking into account the motivations previously exposed, this phase faces the investigation of algorithms for the optic disc, macula and vascular tree segmentation.

3. Detection of lesions produced by diabetic retinopathy: taking advantage of the information previously obtained, the aim of this phase is the identification of signs of DR by means of the analysis of the image. There are two main challenges to overcome in this task. The first one is the obtention of algorithms capable to identify all the different lesions produced by DR. In this sense, a system for the diagnosis of DR should be able to detect microaneurysms, haemorrhages, exudates and cotton wool spots, and also should be able to distinguish them. The second challenge derives from the fact that the image to be analysed can come from a healthy patient or from a patient affected by the illness. Therefore, the system can produce false positives, if it detects non-existing lesions, and false negatives, if it does not detect existing lesions. Therefore, the second challenge is the minimization of both types of errors.

4. Implementation of an expert system: this is the last development phase of the system. At this point, the image was accepted to be processed, the retinal anatomical components were segmented and the results of the image analysis for the identification of lesions are available. Therefore, in this phase, an expert system for the identification of DR is designed and implemented. Taking all the information resulting from the previous phases, this expert system has to provide a diagnosis of the illness. The complexity of this system is the main design decision. It can be implemented simply as a "yes/no" decision system, as a system based on any population screening classification or, even, as a system based on any complex full disease classification.

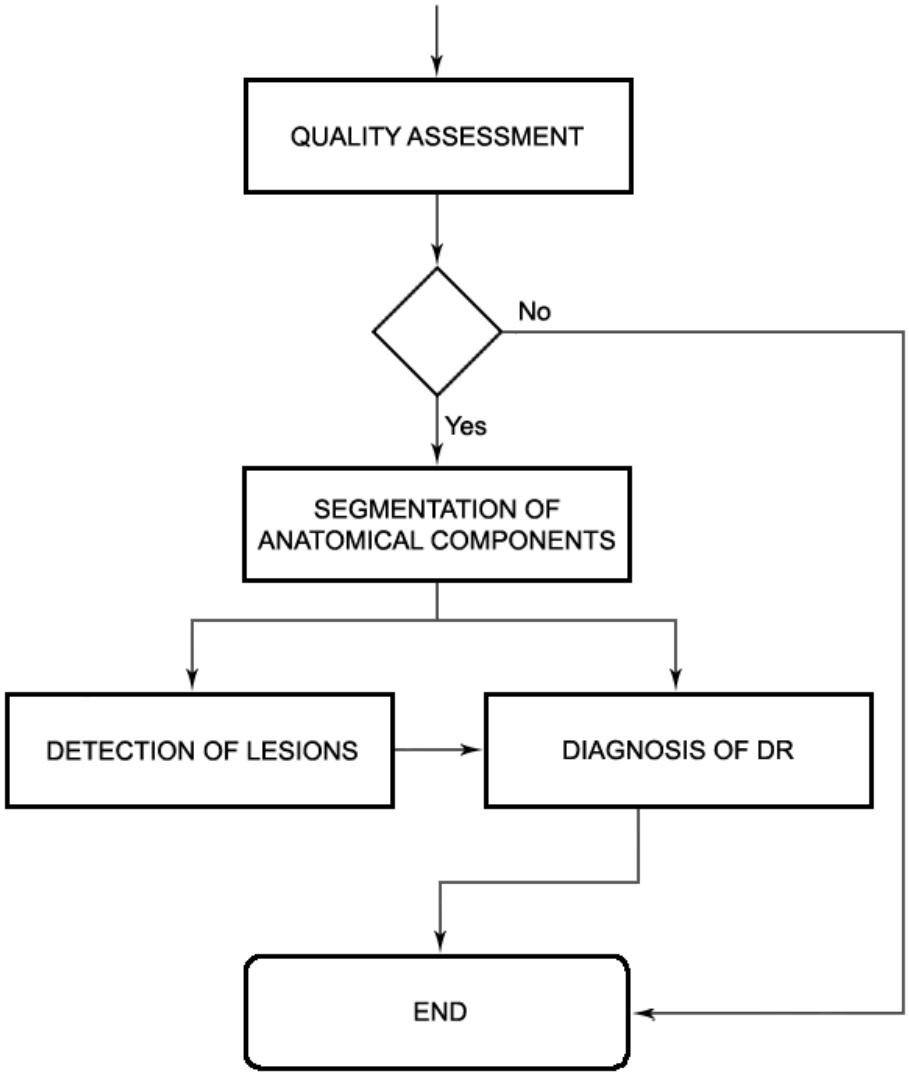


Figure 2.11: *Flow diagram of a generic system for the automated diagnosis of DR.*

Each of the described phases produce as a result a logical module of the system which gives support to its underlying modules. This results in a strong unidirectional coupling between modules, what argues for a bottom-up development of the system. Figure 2.11 shows a flow diagram of a generic comprehensive system for the automated diagnosis of DR based on the exposed descriptions. In this diagram, the modules resulting from the described phases can be identified and it can be checked how these modules interact unidirectionally from the bottom to the top.

Some prototypes of automated systems for the detection of DR can be found in the recent literature. Usher et al. (2003) presented a "yes/no" decision system based on the detection of microaneurysms, haemorrhages and exudates. Thus, each image was classified using a neural

network as normal or abnormal (affected by DR) according to the presence or absence of lesions. The system was trained using a set of images from 500 patients and evaluated by comparing its performance with a human grader on a test set of images from 773 patients. Results obtained by this system were 94.8% *sensitivity* and 52.8% *specificity*. These results mean that 94.8% of patients affected by DR were correctly detected. However, only 52.8% of patients not affected by the illness were correctly classified, what means that 47.2% of healthy patients were erroneously labeled as affected by DR. Another "yes/no" decision system was presented later by Philip et al. (2007). The system, called "disease/no disease" grading system, used a neural network for patient classification according to the detection of red lesions (microaneurysms and haemorrhages) exclusively. The neural network was trained using 1067 retinographies and tested on a set of 14406 images from 6722 patients. The system obtained a *sensitivity* of 90.5%, lower than the previously commented system, and a *specificity* of 67.4%, what supposes an increasing of 14.6% in healthy patients correctly detected. In spite of this increasing, 32.6% of these patients remain erroneously classified as affected by DR. Recently, Abràmoff et al. (2008) presented a system for the automated detection of DR, built from published algorithms, also based in a "yes/no" illness presence scheme according to the detection of microaneurysms, haemorrhages, exudates and cotton wool spots. This system extracted several features from the image under evaluation and, using them, created two probabilities: the probability that an exam has sufficient quality for evaluation and the probability that an exam shows referable DR. These probabilities were ultimately thresholded at an optimum threshold value, calculated to obtain the maximum sensitivity and specificity in the system, to decide a "yes/no" illness presence response. The system was tested on a large group of patients from a true screening population, i.e., with a high proportion, of up to 90%, of normal retinas and a low proportion, of up to 30%, of detectable DR. On this test set, the system obtained 84.0% *sensitivity* and 64.0% *specificity*. These figures indicate that 16.0% of patients affected by DR were erroneously classified as normal and that 36.0% of healthy patients were erroneously recommended for ophthalmic revision as affected by DR.

Results of the reviewed systems for the automated detection of DR suggest that accuracy currently provided by these systems is not still high enough to include them in screening programs for the automated management of patients. The way to progress in the development of such systems is definitely associated with the improvement of algorithms for the

segmentation of the optic disc, macula and vascular vessels, and for the detection of lesions. These conclusions agree with those exposed by Abràmoff et al. (2008) regarding the results of their system. Literally, they concluded from their results: "automated detection of diabetic retinopathy using published algorithms cannot yet be recommended for clinical practice. However, performance is such that evaluation on validated, publicly available datasets should be pursued. If algorithms can be improved, such a system may in the future lead to improved prevention of blindness and vision loss in patients with diabetes."

This thesis aims to contribute to the development and improvement of systems for the automated detection of DR by proposing novel high accurate algorithms for the segmentation of the optic disc, macula and vascular tree. This work is framed within the development of such a system for the Health Ministry of the Andalusian Regional Government (Andalusia, Spain) whith the aim of enhancing effectiveness of its screening program for DR in the coming years.

Chapter 3

Optic Disc Segmentation

OPTIC disc segmentation is an important step in developing systems for automated diagnosis of diabetic retinopathy. In this sense, optic disc segmentation has been used in the literature as a key preprocessing component in many algorithms designed to identify other fundus features. To this respect, knowledge on the optic disc has been useful to estimate the location of the macula and has served as initial information for segmenting the vascular tree. Additionally, finding the optic disc has also been reported valious to decrease false positives in the detection of white lesions related to diabetic retinopathy such as, for instance, retinal exudates.

This chapter presents a new methodology for segmenting the optic disc from digital retinal images. The methodology requires a pixel located within this as initial information. For this purpose, a location methodology is also presented.

The chapter is structured in **six sections**. The **first section** characterizes the optic disc in digital eye fundus images and motivates its segmentation in the context of automated diagnosis of diabetic retinopathy.

The **second section** reviews the optic disc location methods present in the literature. Moreover, this is also done for the optic disc segmentation methods, categorizing and summarizing the most relevant published works. There are mainly two types of methods for optic disc segmentation: *template-based* methods and methods based on *deformable models* or *snakes*. Template-based methods are those that approximate the optic disc boundary by means of the application of a circular or an elliptical model. On the other hand, methods based on deformable models or snakes use models with a high number of degrees of freedom to attempt to find its exact contour.

The **third section** presents the proposed location and segmentation methodologies. The location methodology is a voting-type algorithm composed of three independent methods for locating the optic disc. The segmentation methodology starts extracting an optic disc-containing subimage. This subimage is defined by an optic disc pixel, decided by the location methodology, and its surrounding region (a surrounding region wide enough to include it completely). In this subimage, the segmentation methodology obtains a circular optic disc boundary approximation.

The algorithms are evaluated in the **fourth section**. The optic disc location methodology was subjectively evaluated by medical experts of the Andalusian Health Service. Results denoted as "success" were those in which the pixel found was located within the optic; the contrary cases were denoted as "failure". Regarding the optic disc segmentation methodology, it was subjectively and objectively evaluated. The subjective evaluation was made by medical experts. For each segmentation produced by the algorithm, they labeled a "success" when the circular optic disc boundary approximation appropriately fitted its boundary, and a "failure" when the approximation was not representative of its contour. On the other hand, the objective evaluation was made by measuring the common area overlapping between automated segmentations and true optic disc regions in a created gold standard set.

The **fifth section** discusses on the advantages and disadvantages of the models more generally used for optic disc segmentation. The choice of a model to segment the optic disc is not trivial. Since it is not exactly circular in shape, the use of a circular model may entail descriptive potential loss but, however, its simplicity, with only three degrees of freedom, may provide more robustness. As the number of degrees of freedom increase, the descriptive potential of the model also increases. However, it is also expectable loss of robustness, since the model may be more sensitive to distractors in the image or irregular contrast in the optic disc frontier.

The **sixth section** ends this chapter with the author's discussion and conclusion.

3.1 Introduction

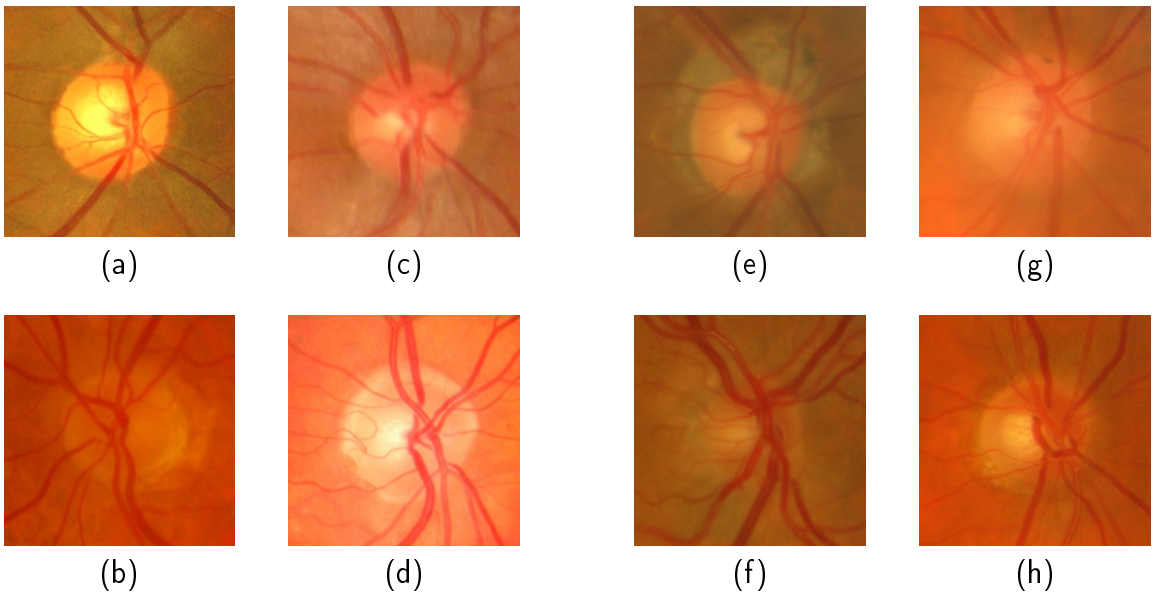
Optic disc (OD) segmentation plays an important role in developing automated diagnosis expert systems for diabetic retinopathy (DR) as it is a key preprocessing component in many algorithms designed to identify other fundus features. Since the OD and macula center are

at a quasi-constant distance, this feature can be used to help estimate the location of the latter (Sinthanayothin et al., 1999; Tobin et al., 2007). On the other hand, to segment the vascular tree, vessel tracking methods need an initial seed vessel point. For this, pixels of vessels within the OD or in its vicinity have been used (Gagnon et al., 2001; Li and Chutatape, 2001). In addition, OD segmentation can be useful in diagnosing automatically some diseases caused by DR. Finding the OD can be used to decrease false positives in the detection of regions of retinal exudates (Osareh et al., 2003), and these injuries are a diagnostic key to grading the risk of macular edema (ME).

Besides DR, OD segmentation is also relevant for automated diagnosis of other ophthalmic pathologies. One of them and maybe the most noteworthy is Glaucoma. This illness is the second most common cause of blindness worldwide (Quigley and Broman, 2006) and, as occurs with DR, its progression is usually asymptomatic in the early stages and, if it is not detected early, OD damage is irreversible and can lead to blindness. Glaucoma is identified by recognizing the changes in shape, color or depth that it produces in the OD (Li and Chutatape, 2003). Thus, its segmentation and analysis can be used to detect evidence of Glaucoma automatically.

The OD can be distinguished in eye fundus images as a slightly elliptical shape. Its size may vary significantly and different estimations have been made. Whereas Sinthanayothin et al. (1999) stated that it occupies about one seventh of the entire retina, alternatively other authors have pointed out that OD size varies from one person to another, occupying about one tenth to one fifth of the retina (Li and Chutatape, 2001). In color fundus images, the OD usually appears as a bright yellowish region, although this feature may also experience significant variations (Figure 3.1).

OD segmentation is not an easy matter. Besides the variations in OD shape, size and color pointed out previously, there are some additional complications to take into account. Contrast all around the OD boundary is usually not constant or high enough piecewise. This is due to outgoing vessels partially obscuring portions of the rim producing “shadows”. This phenomenon occurs in a more pronounced form in macula-centered retinographies. Furthermore, this kind of image is especially sensitive to non-appropriate pupil dilation. As the OD is not centered in the camera lens, if the pupil is not appropriately dilated, it does not receive enough light, producing a fundus image with an OD not visible at all. Another distractor is produced when peripapillary atrophy is present. This condition is caused by



OD color variability: (a) Yellowish OD. (b) Brownish OD. (c) Reddish OD. (d) Whitish OD.

OD segmentation distractors: (e) OD with paripapillary atrophy. (f) Poorly illuminated OD. (g) Slightly blurred OD. (h) OD rim partially obscured by vessels.

Figure 3.1: *Examples illustrating variability in OD appearance.*

the atrophic changes in the retinal pigment epithelium that afford a clear view of the hyper reflective layer, the sclera, thus producing bright areas just outside the OD rim which distort its shape. On the other hand, eye movement at the moment of retinography capture may also lead to slightly blurred images, making their automated analysis even more difficult. This problem can be avoided by simply discarding these images and re-taking new ones. However, this action is not usually applied as their quality is usually good enough for human visual inspection.

3.2 State of art

The available works related to OD processing in eye fundus color images can be grouped into two distinct categories: *location* and *segmentation* methods. The former works focus on finding an OD pixel, generally representative of its center. On the other hand, the latter works estimate the OD boundary. Within this category, a general distinction can be made between *template-based* methods (methods for obtaining OD boundary approximations) and

methods based on *deformable models* or *snakes* (methods for extracting the OD boundary as exactly as possible).

With regard to location methods, Chaudhuri et al. (1989a) and Lee et al. (1999) assumed for locating the OD that this is usually the brightest component on the eye fundus. Thus, a cluster of high intensity pixels with a high gray-scale value was used to identify the OD location. Sinthanayothin et al. (1999) presented a method where the images were preprocessed by applying an adaptative local contrast enhancement to the intensity channel of the HSI color space. The OD center location was then identified using the variance of intensity produced by the blood vessels within the OD. Other works that used the confluence of blood vessels at the OD are those by Akita and Kuga (1982) and Hoover and Goldbaum (1998). Another work of Hoover and Goldbaum (2003) proposed later to locate the OD center by using the vasculature origin. They proposed to determine where all the vessels converged by means of a voting-type algorithm called fuzzy convergence. Another method that uses the convergence of the vessels to detect the OD center was proposed by Foracchia et al. (2004). The four main vessels originating from the OD were geometrically modeled using two parabolas, and the OD position was located as their common vertex. Principal component analysis was used by Li and Chutatape (2004) to detect the OD position. Firstly, the brightest pixels in the fundus were clustered as candidate OD regions. Then, using these candidates, a training set was produced and principle component analysis was then applied to project a new image to the "disc space". The OD center location was finally found by calculating the minimum distance between the original retinal image and its projection. Inspired by previous works, Youssif et al. (2008) presented an OD location method based on a vessels' direction matched filter. As a first step, a binary mask was generated followed by image brightness and contrast equalization. Finally, the retinal vasculature was segmented, and the directions of the vessels were matched to the proposed filter representing the expected vessels' directions in the vicinity of the OD.

With regard to segmentation methods and concretely to works based on deformable models, Osareh et al. (2002) located the OD center by means of template matching and extracted its boundary using a snake initialized on a morphologically enhanced region of the OD. Lowell et al. (2004a) also localised the OD by means of template matching and, moreover, selected a deformable contour model for its segmentation. Specifically, they used a global elliptical model and a local deformable model with variable edge-strength depen-

dent stiffness. Another deformable model-based approach was presented by Xu et al. (2007). The snake behaviour against vessel occlusion was improved and the model was extended in two aspects: knowledge-based clustering and smoothing update. Thus, the snake deformed to the location with minimum energy and then self-clustered into two groups, which were finally updated by the combination of both local and global information. Regarding template-based methods, Wong et al. (2008) proposed: 1) OD location by means of histogram analysis and initial contour definition according to the previously obtained location, and 2) A modified version of the conventional level-set method was subsequently used for OD boundary extraction from the red channel. This contour was finally fitted by an ellipse. Another template-matching approach for OD segmentation is the Hausdorff-based template matching presented by Lalonde et al. (2001). Initially, they determined a set of OD candidate regions by means of multi-resolution processing through pyramidal decomposition. For each OD region candidate, they calculated a simple confidence value representing the ratio between the mean intensity inside the candidate region and inside its neighborhood. The Canny edge detector and a Rayleigh-based threshold were then applied to the green-band image regions corresponding to the candidate regions, constructing a binary edge map. As final step, using the Hausdorff distance between the edge map regions and circular templates with different radii, they decided the OD among all the candidates.

On the other hand, although they do not belong to the two reviewed categories, the works of Walter and Klein (2001), Reza et al. (2008) and Abràmoff et al. (2007) proposed other relevant OD segmentation methods. Walter and Klein (2001) found the OD contour through the watershed transformation. For OD detection, its center was previously approximated as the centroid of the largest and brightest connected object in a binary image obtained by thresholding the intensity channel. Reza et al. (2008) also used the watershed transformation for OD segmentation. Firstly, the green channel was preprocessed for image condition enhancement. Then, morphological opening, extended maxima operator and minima imposition were finally used to apply the watershed transformation for bright structure segmentation. Finally, although applied to stereo images, it is worth mentioning the novel OD segmentation approach presented by Abràmoff et al. (2007). Pixel feature classification by means of a K-nearest neighbor classifier was used in this case for OD segmentation in stereo color photographs.

3.3 Methodology

The aim of this chapter is to introduce a new methodology for OD segmentation that obtains a circular boundary approximation. It needs as initial information the coordinates of a pixel located within the OD. To this effect, an OD location methodology is also proposed to obtain the required OD pixel. It must be stressed that any other location method could be used for this purpose.

All values of parameters, constants and window sizes given in this section refer to retinas of 1046 pixels in diameter. For different image resolutions, all of these values would have to be scaled.

3.3.1 Optic disc location

The location methodology obtains a pixel (called Optic Disc Pixel, ODP, hereafter) that belongs to the OD. It comprises three independent detection methods. Each method obtains its own OD candidate pixel. The final ODP is selected by taking into account the three previous candidate pixels and their location with respect to their average point (centroid). For this, a voting procedure comprising the following cases is applied:

- If the three OD candidate pixels are close to the centroid (closer than one fifth of the retina, maximum OD diameter estimation (Li and Chutatape, 2001)): the selected ODP is the centroid.

- If only two candidates are close to the centroid: the selected ODP is the average point of these two referred pixels.

- Otherwise: the selected ODP is the candidate pixel obtained with the most reliable method (performed tests show this is the *maximum variance method* described below).

The three developed methods work on the green channel of the RGB color space as this is the one that provides the best contrast (Walter and Klein, 2005). Let this gray scale image be denoted as I. Thus, a description of these methods, illustrated in Figure 3.2 by three examples of their application to different eye fundus images, is presented as follows:

- Maximum Difference Method: the OD usually appears as a bright region in eye fundus images. Moreover, the vascular tree formed by the "dark" blood vessels emerges in the

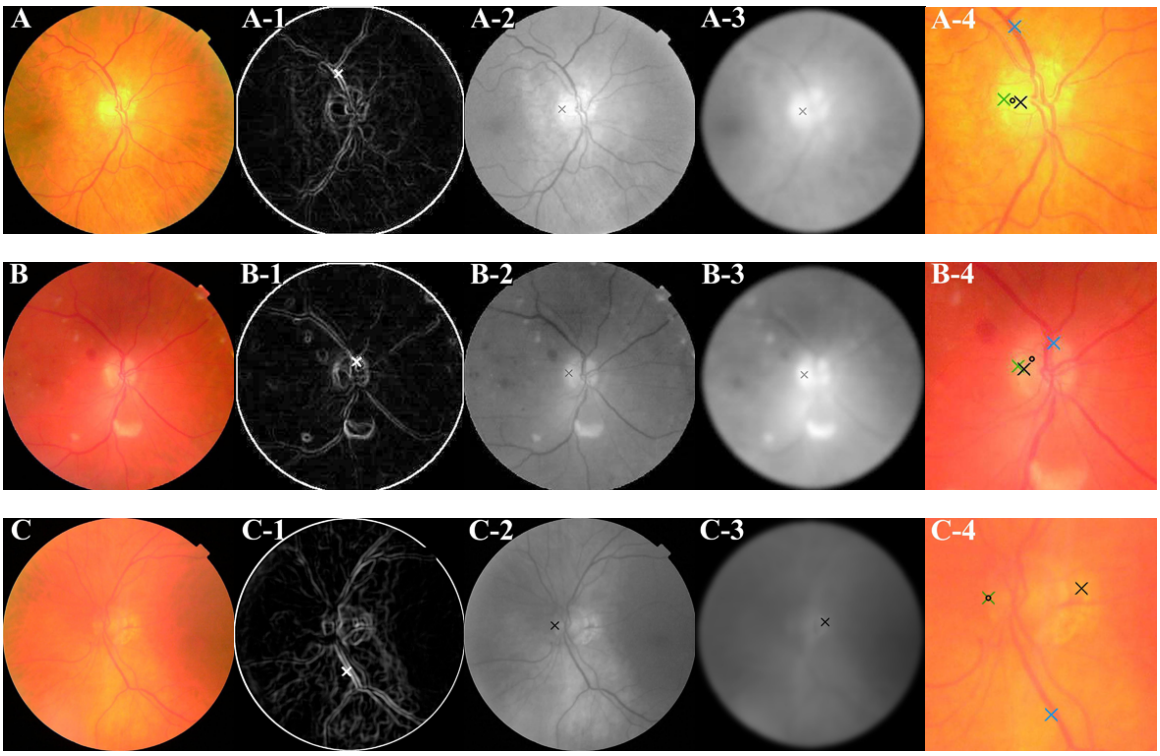


Figure 3.2: *ODP determination: (A), (B) and (C) Original images. (A-1), (B-1) and (C-1) OD pixels rendered by the maximum difference method. (A-2), (B-2) and (C-2) OD pixels rendered by the maximum variance method. (A-3), (B-3) and (C-3) OD pixels rendered by the low-pass filter method. (A-4), (B-4) and (C-4) Final ODP determination (black circumference): in (A-4), although the OD pixel of the maximum difference method is not properly selected, the ODP is successfully located; in (B-4) the ODP selected is the centroid of the three OD pixels; in (C-4), since the three OD pixels are far from their centroid, the ODP selected is the OD pixel from the maximum variance method.*

disc. This is why the maximum variation of the gray levels usually occurs within the OD. This maximum is used by this method to select its OD pixel.

A median filter of 21×21 is applied beforehand to I in order to remove non-significant peaks in the image. If I_M denotes this filtered image, the OD pixel from this method is decided according to the following equation:

$$P_{MDM} = \underset{(x,y)}{\arg} \left(\max \left\{ \max_{(s,t) \in S^{21}_{x,y}} \{I_M(s,t)\} - \min_{(u,v) \in S^{21}_{x,y}} \{I_M(u,v)\} \right\} \right) \tag{3.1}$$

where $\underset{(x,y)}{\arg}$ is the argument operation, which returns the coordinates fulfilling the function input, and $S^w_{x,y}$ denotes the set of coordinates in a square window of size $w \times w$

centered on point (x, y) (see examples of this method in Figure 3.2, images A-1, B-1 and C-1).

- Maximum Variance Method: this method is based on the same properties as the previous one. It calculates the statistical variance for every pixel by using a 71×71 centered window. On the other hand, a set of "bright" pixels is obtained by automatic blue-channel thresholding according to the Otsu method (Otsu, 1978). The OD pixel returned by this method is the maximum variance pixel located in a bright region. This restriction forces this pixel to be close to a bright region (ideally the OD). It was checked that it was enough to consider that a pixel is located in a bright region if it has at least 10 "bright" pixels in its neighborhood (established by means of a 101×101 pixel window) (see Figure 3.2, images A-2, B-2 and C-2). The window sizes selected to compute the variance and to establish the neighborhood criteria, were set with the aim of obtaining the best location performance on the set of images used for testing the methodology, which is the set of 1200 fundus images in the MESSIDOR database (see its description in appendix A).

- Low-Pass Filter Method: the OD pixel of this method is the maximum gray-level pixel in a low-pass filtered image. Although the OD is usually the brightest area in a retinography, the pixel with the highest gray level could not be located within it. In many cases, this pixel may be inside other small bright regions. In order to smooth out these distractors, the image I is transformed to the frequency domain and filtered by the Gaussian low-pass filter defined as follows:

$$H(u, v) = \exp\left(\frac{-D^2(u, v)}{2D_0^2}\right) \tag{3.2}$$

  where $D(u, v)$ is the Euclidean distance between the point (u, v) and the origin of the frequency plane, and D_0 is the cut-off frequency with a value of 25. The highest gray-level pixel in the filtered image returned to the spatial domain is the result of this method (Figure 3.2, images A-3, B-3 and C-3).

The result of the final ODP selection process is illustrated by the three examples of application of the methodology shown in Figure 3.2. In the first example (Figure 3.2, images A to A-4), it can be confirmed that although the pixel returned by the maximum difference

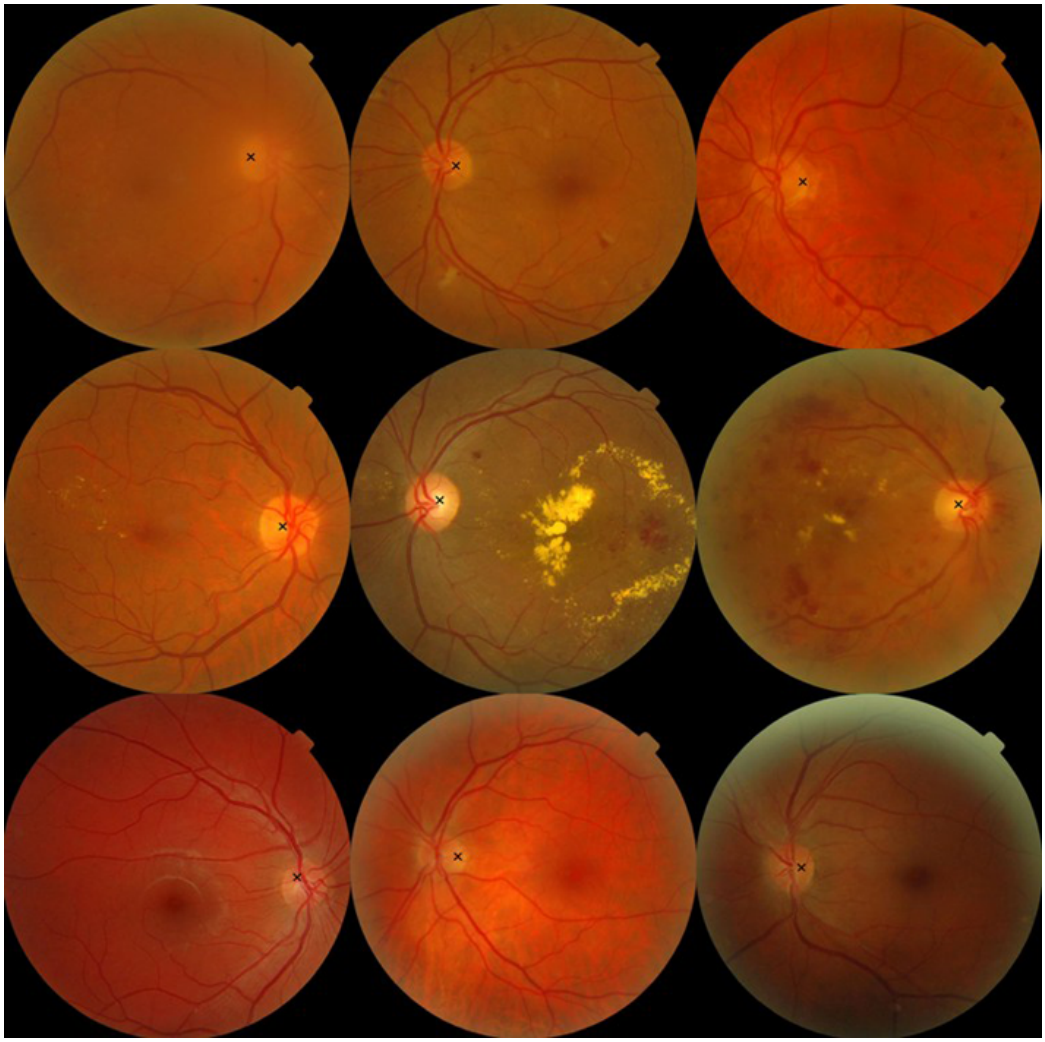


Figure 3.3: *Examples of final results obtained with the OD location methodology. The first row shows three retinographies from patients with DR and with no risk of ME, the second is composed of three retinographies from patients affected by DR and with risk of ME and the last three images are from healthy retinas.*

method is outside the disc, it is discarded and the ODP is successfully selected. In the second example (Figure 3.2, images B to B-4), pixels returned by the three methods are closer than one fifth of the diameter of the retina, so the location of the ODP is the location of their centroid. In the last example (Figure 3.2, images C to C-4), the three partial results are far from their centroid, so the final ODP is the OD pixel found by the maximum variance method. Figure 3.3 shows some final results of the OD location methodology on healthy retinas as well as on retinas affected by DR with and with no risk of ME.

3.3.2 Optic disc boundary segmentation

The method proposed is performed on an RGB subimage of the original retinography. By this way, robustness and efficiency in OD segmentation are increased as it reduces the search space and decreases the number of artifacts and distractors present in the whole image. So, as a first step, a 400×400 RGB subimage is extracted centered on an OD pixel provided by the OD location methodology previously presented.

Although the green component of an RGB retinography is the one with highest contrast (Walter and Klein, 2005), the OD is often present in the red field as a well-defined white shape, brighter than the surrounding area. When contrast between the OD shape and its environment in this color field is high enough, the OD can usually be segmented better than in the green field. At other times, the OD is not discernable in the red component and has to be segmented in the green one. To exploit this feature, the OD segmentation is performed in parallel on the two components and the "better" of the two segmentations is ultimately selected. The proposed procedure firstly employs an special morphological processing to eliminate blood vessels. Then, a binary mask of the OD boundary candidates is obtained by applying edge detection and morphological techniques. Finally, the Circular Hough Transform is used to calculate the circular approximation of the OD.

Elimination of blood vessels

Consider the gray-level image I from the red or green field of the subimage containing the OD (Figure 3.4, images R and G). As it was discussed, blood vessels within the OD act as strong distractors, so they should be erased from the image beforehand. The vasculature is piecewise linear and can be considered as a structure composed of many such connected linear shapes with a minimum length L and a maximum width W, where usually $W \ll L$ (see Heneghan et al., 2002). These linear shapes are formed, as a general rule, by a set of pixels with an almost constant gray-level value, with this value being somewhat lower than the gray-level values of non-vessel pixels in their vicinity. Using a rotating linear structuring element B of width 1 and length $L_1 \gg W$, a linear shape can be detected by calculating the statistical variance of gray-level values of pixels along it. The rotation associated with the minimum value will be that in which the vessel contains B and, conversely, the rotation with the highest value will refer to the situation in which B crosses the linear shape (see Figure 3.5). Thus, vessels can be removed from image I by finding, for every pixel, the rotation which

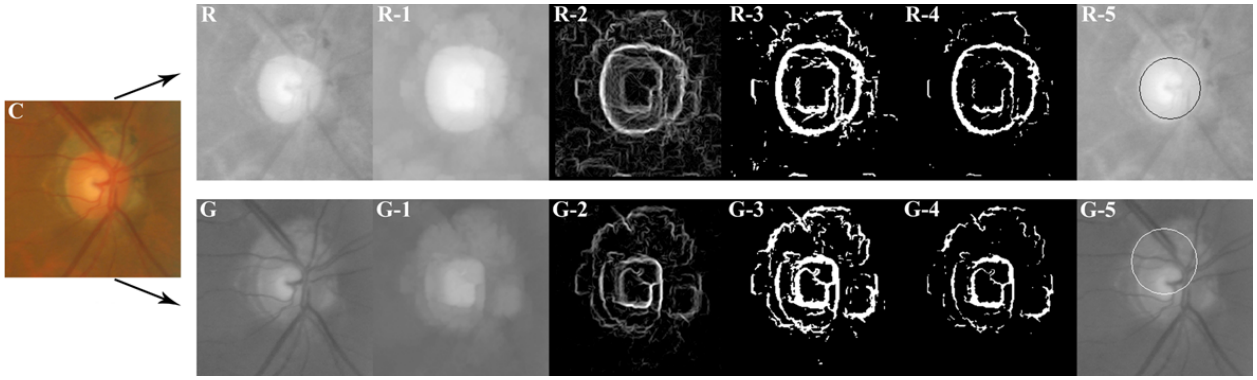


Figure 3.4: *Illustration of the process for the calculation of the circular OD boundary approximation: (C) Initial RBG subimage containing an OD affected by peripapillary atrophy. On the right, the top row shows the process performed on the red channel, whereas the bottom corresponds to the process applied to the green component. (R) and (G) Sub-images extracted from the red and green channels of (C), respectively. (R-1) and (G-1) Vessel elimination. (R-2) and (G-2) Gradient magnitude image. (R-3) and (G-3) Binary image. (R-4) and (G-4) Cleaner version of the binary image. (R-5) and (G-5) Circular OD boundary approximation. The scores obtained in the Circular Hough Transform algorithm are,* 264 *for segmentation in (R-5) and* 130 *for segmentation in (G-5), so the segmentation selected would be the one performed on the red channel.*

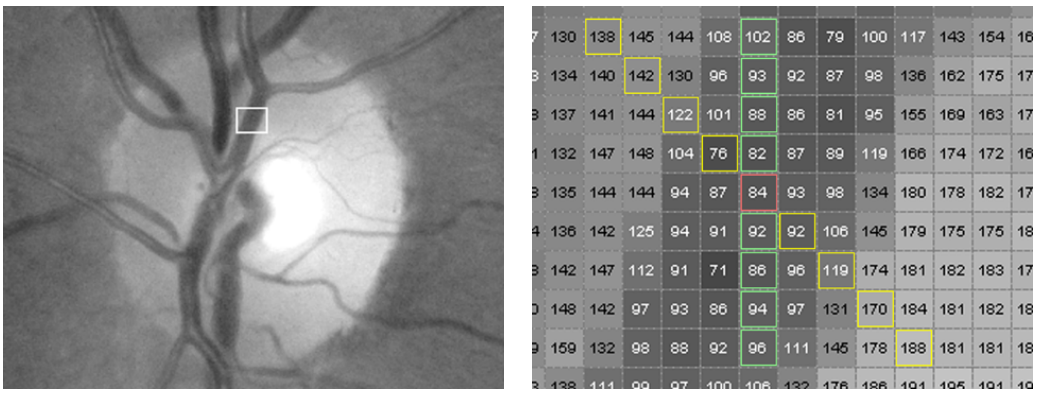
produces the maximum gray level variance value and taking the maximum gray-level in the neighborhood defined by the structuring element at that rotation (see effects at Figure 3.4, images R-1 and G-1). Mathematically, the obtention of an image with removed vessels, $I_{\neg V}$, can be expressed as

$$I_{\neg V}(x,y) = \max_{(s,t)\in B_{x,y}^{L_1;r}} \{I(s,t)\} \tag{3.3}$$

where $B_{x,y}^{L_1;r}$ represents structuring element B of length L_1, centered on point (x,y) and at rotation r. r is determined by finding the rotation k associated to the maximum variance considering twelve rotations of B each 15 ° apart:

$$r = \arg_k \left(\max \left\{ \operatorname{var}_{(s,t)\in B_{x,y}^{L_1;k}} \{I(s,t)\},\ k = 1,\ldots,12 \right\} \right) \tag{3.4}$$

where $\arg_k$ is the argument operation. The length L_1 of the structuring element should be chosen so as to cross vessels in order to erase them by applying (3.3). Taking into account the study proposed by Heneghan et al. (2002), the value for this parameter was set to 27. This operation produces an OD enlargement of L_1 pixels in all directions. This will be considered at the end of the processing.



(a) (b)

Figure 3.5: *Illustration of vessel detection using a rotating linear structuring element and statistical variance of gray levels. (a) Green channel of an image section containing the OD. (b) Enlarged image of the zone delimited by the white square in image (a). Each cell corresponds to a pixel of the image and contains a number representing its gray level. For a linear structuring element of length* 9*, centered on the pixel marked in red and with twelve rotations each* 15 *degrees apart, green represents the rotation in which the minimum statistical variance of gray levels is obtained, and yellow represents the rotation associated with the maximum value.*

Obtaining optic disc boundary candidates

The OD boundary represents the frontier between the OD and the background. It is characterized by a sudden variation in gray levels, with these values higher within the OD than in its surroundings. So, the OD boundary can be detected by measuring the gradient magnitude of gray-level changes in small neighborhoods of the image. Firstly, a 3×3 mean filter is applied to obtain I_S, in which pixel values unrepresentative of their environment are smoothed:

$$I_S(x,y) = \underset{(s,t)\in S^3_{x,y}}{\text{mean}} \{I_{\neg V}(s,t)\} \tag{3.5}$$

The Prewitt edge detector (Gonzalez and Woods, 2002) estimates the magnitude and orientation of an edge. It uses two 3×3 convolution kernels to calculate, respectively, approximations of the derivatives for horizontal (3.6) and vertical changes (3.7),

$$I_{G_x}(x,y) = I_S(x,y) * P_x \tag{3.6}$$

$$I_{G_y}(x,y) = I_S(x,y) * P_y \tag{3.7}$$

where $*$ denotes the convolution operation and the two Prewitt kernels P_x and P_y are defined as

$$P_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}; P_y = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \tag{3.8}$$

A gradient magnitude image I_{GM} is obtained by computing for every pixel (x, y) the module of its partial derivative values,

$$I_{GM}(x, y) = \sqrt{I_{G_x}^2(x, y) + I_{G_y}^2(x, y)} \tag{3.9}$$

Thus, I_{GM} is an image which contains information on edges, specifically on the location and intensity of local gray-level variations (Figure 3.4, images R-2 and G-2). As the blood vessels were previously erased, in general the most significant edges in the gradient image correspond to the OD boundary. Thus, a binary mask of OD boundary candidates can be produced by thresholding the image I_{GM}.

As stated before, there is great variability in OD appearance, and the contrast level between the OD and the background may vary quite substantially. That is why it is not suitable to establish a unique threshold for any image. The Otsu thresholding method automatically decides a threshold for a gray-level image by assuming that it is composed of two sets, the background and the foreground. Then, the method establishes the optimum threshold T_{OTSU} by maximizing the between-class variance. Using this threshold, a first binary mask of OD boundary candidates is given by a simple binarization operation (see Figure 3.4, images R-3 and G-3),

$$I_B(x, y) = \begin{cases} 0 & \text{if } I_{GM}(x, y) \leq T_{OTSU}, \\ 255 & \text{if } I_{GM}(x, y) > T_{OTSU}. \end{cases} \tag{3.10}$$

This image can contain some noise caused by small rims present in the original image and detected in I_{GM}. So, the definitive binary mask of OD boundary candidates is obtained by cleaning I_B by means of morphological erosion (Figure 3.4, images R-4 and G-4),

$$I_{BM}(x, y) = \min_{(s,t) \in C_{x,y}^D} \{I_B(s, t)\} \tag{3.11}$$

where $C^D_{x,y}$ is a circular structuring element with a diameter of D pixels, centered on point (x, y) and whose value was fixed to 5. This operation reduces the OD radius in 2 pixels. This reduction along with the enlargement produced in eq (3.3) will be corrected at the end of the processing.

Final optic disc boundary segmentation

The Hough Transform is widely used in Computer Vision and Pattern Recognition for detecting geometrical shapes that can be defined by parametric equations. Based on the primitive Hough Transform (Hough, 1962), the Circular Hough Transform was outlined by Duda and Hart (1972) and later improved and extended by Kimme et al. (1975). It aims to find circular patterns within an image. It is used to transform a set of feature points in the image space into a set of accumulated votes in a parameter space. Then, for each feature point, votes are accumulated in an accumulator array for all parameter combinations. The array elements that contain the highest number of votes indicate the presence of the shape. A circumference pattern is described by the parametric equation of the circumference, defined as

$$(x-a)^2 + (y-b)^2 = r^2 \tag{3.12}$$

where (a, b) are the coordinates of the circle center and r is the radius. So, the circular shapes present in I_{BM} can be obtained by performing the Circular Hough Transform on this image. It can be defined as

$$(P_c, r) = CHT\left(I_{BM}, r_{min}, r_{max}\right) \tag{3.13}$$

where $P_c = (i_c, j_c)$ and r are respectively the center position and the radius that define the circular shape with the highest punctuation in the Circular Hough Transform implemented by CHT. The radius r is restricted to be between r_{min} and r_{max}, values which are one tenth and one fifth of the image (Li and Chutatape, 2001) divided by two (as these measurements refer to OD diameter estimation). The minimum radius restriction reduces the probability of considering the OD cup, while the maximum radius restriction eliminates candidates with too wide areas. The obtained r value must be corrected due to the effects of (3.3) and (3.11). The vessel elimination performed in (3.3) enlarged the OD 27 pixels and the erosion operation in (3.11) produced a 2-pixel reduction, so the r value has to be reduced by 25.

As previously commented, this processing is applied in parallel to the green and red channels. Thus, two OD approximations are obtained. The one with the higher score in the Circular Hough Transform algorithm is then selected as the definitive circular OD boundary approximation (Figure 3.4, images R-5 and G-5). This score quantifies the point by point matching degree between the estimated circumference and the fitted shape in I_{BM}. Therefore, higher scores generally involve better OD border extraction and, hence, better segmentation quality. Moreover, the selection of the correct candidate is also favoured by the fact that the score of this algorithm is an absolute and not a relative measure. This implies that the selected maximum-score criterion tends to select longer candidate circumferences. This is especially useful when the OD cup is wide enough to be considered a candidate, as it leads to an increased probability of selecting the correct candidate between the cup and the true OD boundary.

Figure 3.6 shows OD segmentation examples produced by the presented methodology. The images used for these examples are the same used in Figure 3.3.

3.4 Experimental results

The proposed OD location and segmentation methodologies were tested using the MESSIDOR database described in appendix A. To prevent the inclusion of any kind of skew, although some images in the set are not suitable for OD processing (i.e., images too blurred or with severe enough cataract), no exclusion criteria was applied. Since the MESSIDOR database do not provide an OD gold standard set, to make evaluation of the algorithm performance on this database possible, experts of the Andalusian Health Service created this set by manually delimiting the OD rim in all images. On the other hand, images in the MESSIDOR database have resolutions of 1440×960, 2240×1488 or 2304×1536. To this respect, tests on image resolution, scaling down the images, were performed. These tests revealed that the results provided by the location and segmentation methodologies are stable in spite of decreasing image resolution down to 300×300 for OD location and 640×640 for OD segmentation. Therefore, the methodologies presented are actually applied to images of these sizes. Any image of a given resolution is reduced to 300×300 and 640×640 for OD location and OD segmentation respectively and the processes are performed scaling the

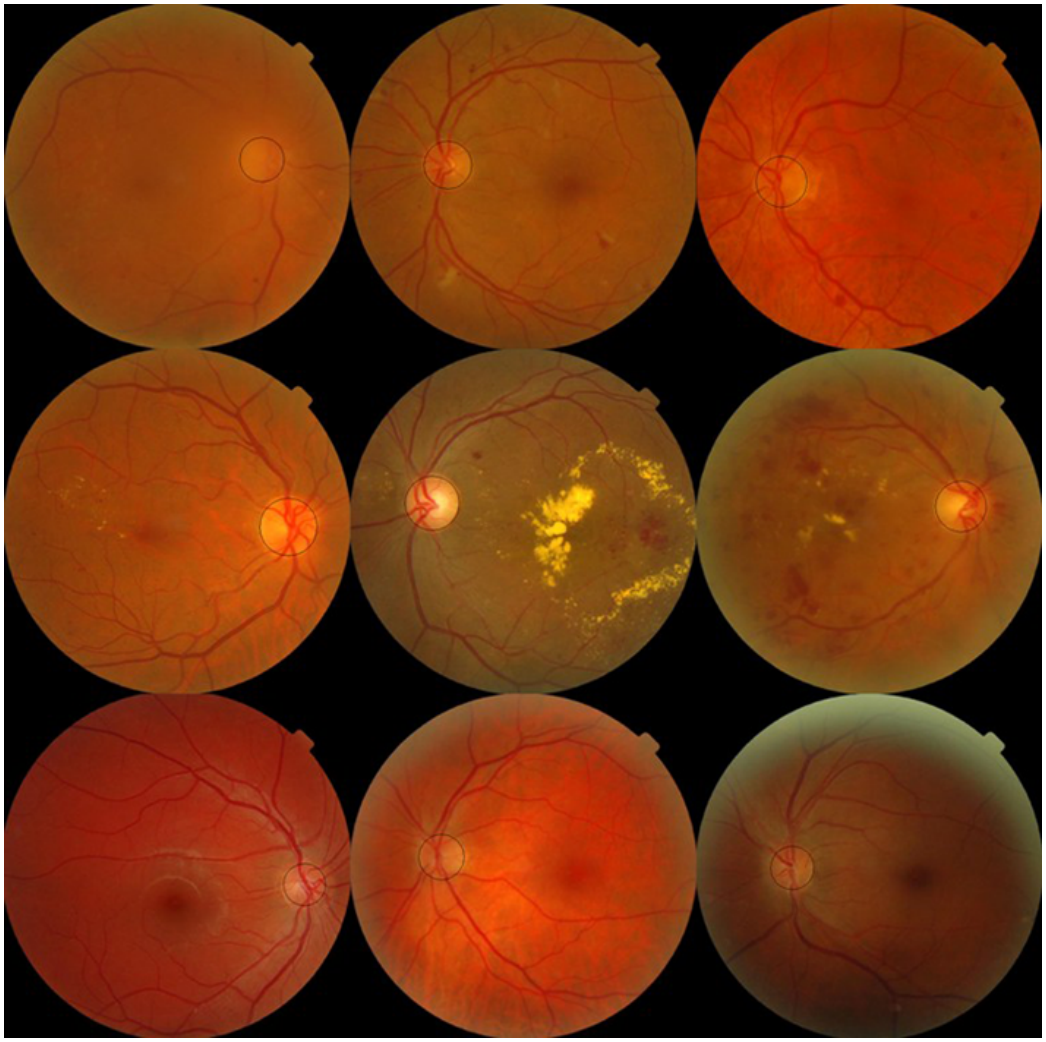


Figure 3.6: *Examples of results obtained with the OD segmentation methodology on the same images described in Figure 3.3. The first row shows three retinographies from patients affected by DR and with no risk of ME, the second row shows three retinographies from patients affected by DR and with risk of ME and the last row is composed of three images from healthy retinas.*

window sizes and parameters to these resolutions. So, the results described in this section were obtained applying these reductions.

This section is divided into three subsections. The first one shows the testing results provided by the location methodology. The second and third subsections describe the results obtained with the segmentation methodology. The second subsection is focused on the subjective evaluation of the algorithm performance. Here, the criterion used for this purpose is defined and the obtained results are presented and discussed. The third subsection deals with the objective evaluation of the algorithm performance. A metric for the mathematical

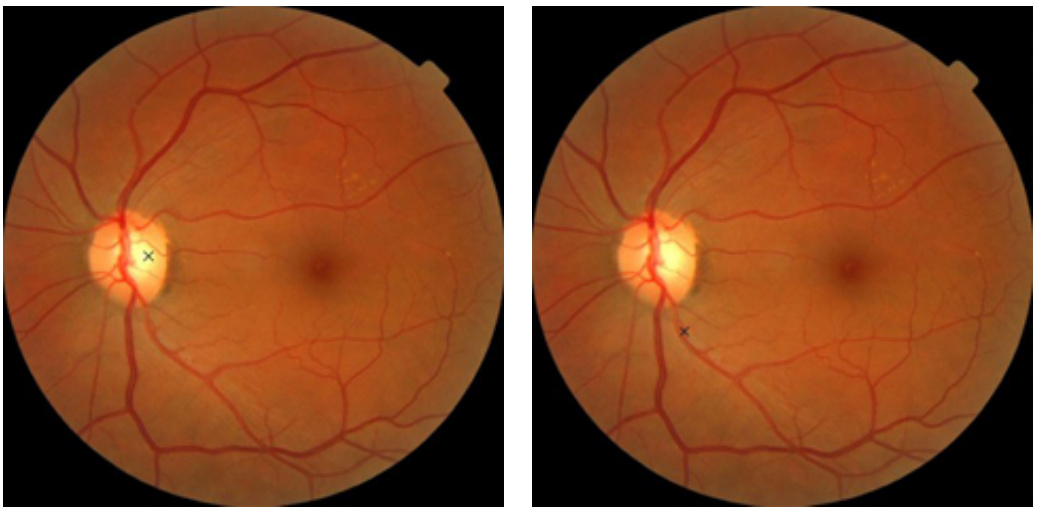


Figure 3.7: *Illustration of the criterion used for evaluating results the OD location methodology. (a) A result that would be labeled a "success". (b) A result that would be labeled a "failure".*

evaluation of the performance is defined and results are presented, discussed and compared to another circular template-based algorithm for OD segmentation present in the literature.

3.4.1 Results of the OD location methodology

Results of the OD location methodology [1] were checked by ophthalmologists from the Andalusian Health Service, providing a failure-success result to every image. A result was denoted a "success" when the pixel found was located within the OD and otherwise, a "failure" (see examples of this criterion in Figure 3.7).

The results obtained with the location methodology are summarized in Table 3.1. This table shows the average percent of successful OD locations for each diagnosis case of DR Grade - Risk of ME, cases established according to the gradings described in section 2.2 (DR condition from 0, illness absence, to 4, maximum condition severity; risk of ME edema from 0, no risk, to 2, highest risk). This methodology successfully found the OD in 1186 cases out of the whole set of 1200 analysed images, which means a 98.83% success rate. Classifying the images by retinas affected by DR or not, the achieved success rates were 98.33% and 99.44% respectively. For images from patients affected by DR, the success rate was 96.51% for images diagnosed with risk of ME and 99.3% for those images without this risk. These

[1] OD location results for all images are available at: `http://www.uhu.es/retinopathy`, in the Optic Disc Results subsection of the Results section

DR Grade - Risk of ME	Cases	LocationMethod (this work)
0-0	540	99.44%
1-0	142	99.30%
1-1	5	100.0%
1-2	6	100.0%
2-0	182	100.0%
2-1	28	100.0%
2-2	37	100.0%
3-0	107	98.13%
3-1	42	100.0%
3-2	111	92.79%
Total	1200	98.83%

Table 3.1: *Results of the OD location methodology. Results are detailed per diagnosis case and are given as the average percent of obtained successful OD locations.*

	Hôpital Lariboisière Paris		Faculté de Médecine St. Etienne		LaTIM - CHU de Brest	
DR Grade - Risk of ME	Cases	LocationMethod (this work)	Cases	LocationMethod (this work)	Cases	LocationMethod (this work)
0-0	151	100.0%	186	98.39%	203	100.0%
1-0	26	100.0%	68	98.53%	48	100.0%
1-1	2	100.0%	1	100.0%	2	100%
1-2	2	100%	2	100.0%	2	100%
2-0	49	100.0%	62	100.0%	71	100.0%
2-1	13	100%	9	100.0%	6	100.0%
2-2	8	100%	20	100.0%	9	100.0%
3-0	48	95.83%	21	100.0%	38	100.0%
3-1	26	100.0%	7	100.0%	9	100.0%
3-2	75	89.0%	24	100.0%	12	100.0%
Total	400	97.50%	400	99.0%	400	100.0%

Table 3.2: *Extended results of the OD location methodology. Results are detailed per diagnosis case and ophthalmic center and are given as the average percent of obtained successful OD locations.*

figures show a stable behaviour of the proposed location methodology upon illness presence or severity. Table 3.2 offers an extended report showing the obtained results detailed per ophthalmic center.

Tests concerning computational efficiency were also done. The tests were performed on a current mid-range PC equipped with an Intel Core2Duo CPU at 2.13 GHz and 2 GB of RAM capacity. The algorithm was implemented in Java programming language and tested running

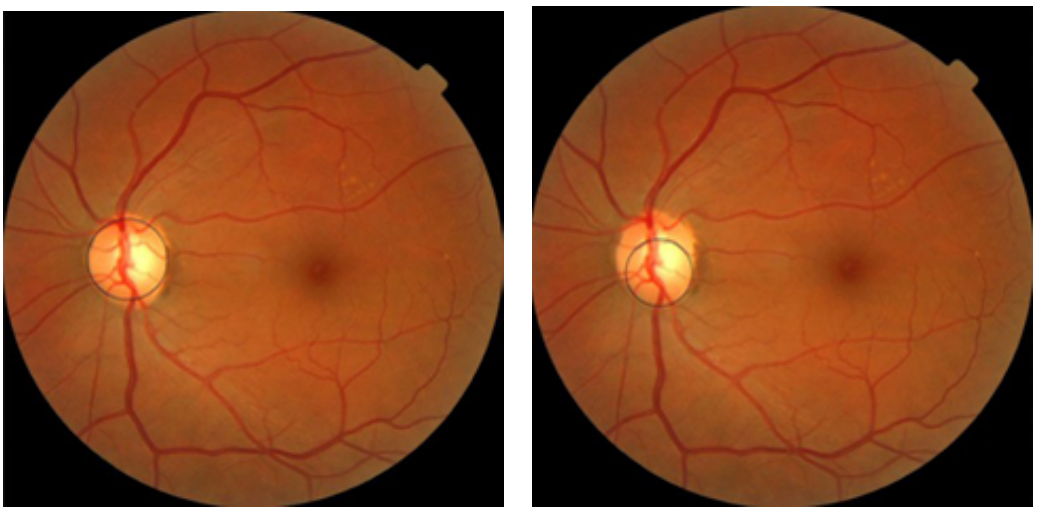


Figure 3.8: *Illustration of the criterion used for subjectively evaluating results the OD segmentation methodology. (a) A result that would be labeled a "success". (b) A result that would be labeled a "failure".*

on the Java Virtual Machine version 6. For a total of 1200 executions of the algorithm, the average computational time obtained was 1.67 sec. with a standard deviation of 0.14 sec.

Finally, it should be pointed out that, for testing the OD segmentation algorithm (results described below), the location methodology presented here was used to obtain the required initial OD location. As said before, this algorithm correctly found the OD in 1186 out of the 1200 images. OD location was manually determined in the 14 images in which the OD location algorithm failed for checking the segmentation algorithm.

3.4.2 Results of the OD segmentation methodology: subjective evaluation

To subjectively evaluate the OD segmentation algorithm, as it was done to evaluate the location methodology, results produced by this methodology were checked by ophthalmologists from the Andalusian Health Service. They provided a failure-success result to every image according to the following criterion: a resulting image given by the boundary segmentation methodology was labeled a "success" when the circular OD boundary approximation appropriately fitted its boundary and a "failure", when the approximation was not representative of its contour (see examples in Figure 3.8).

DR Grade - Risk of ME	Cases	CircularHough (this work)
0-0	540	95.93%
1-0	142	93.66%
1-1	5	100.0%
1-2	6	100.0%
2-0	182	95.6%
2-1	28	89.29%
2-2	37	94.59%
3-0	107	91.59%
3-1	42	85.71%
3-2	111	94.59%
Total	1200	94.58%

Table 3.3: *Results of the OD segmentation methodology obtained by subjective evaluation. Results are detailed per diagnosis case and are given as the average percent of obtained good OD segmentations.*

Table 3.3 shows the results obtained by the segmentation methodology [2] (referred to as *CircularHough* hereafter). For each diagnosis case of DR Grade - Risk of ME, the average percent of obtained good OD segmentations is given. As it can be checked, the segmentation methodology obtained a good OD boundary approximation in 94.58% of cases, which means segmentation success in 1135 out of 1200 images. For retinas affected by DR, the success rate was 94.34% whereas, for healthy retinas cases, the segmentation was appropriate in 95.93%. Finally, for patients affected by DR and with risk of ME, a 92.58% success rate was achieved, and for images diagnosed as affected with DR with no risk of ME this figure was 93.97%. These results show evidence that the segmentation algorithm performance is not sensitive to the illness presence or absence. Furthermore, since success rate does not show a decreasing pattern related to illness severity increasing, algorithm performance does not either show sensitivity to condition severity when the illness is present. Table 3.4 completes the evaluation report showing the obtained results detailed per ophthalmic center.

3.4.3 Results of the OD segmentation methodology: objective evaluation

To objectively evaluate algorithm performance, the overlapping degree between the true OD regions in "gold standard" images and the approximated regions obtained with the

[2] OD segmentation results for all images are available at: `http://www.uhu.es/retinopathy`, in the Optic Disc Results subsection of the Results section

	Hôpital Lariboisière Paris		Faculté de Médecine St. Etienne		LaTIM - CHU de Brest	
DR Grade - Risk of ME	Cases	CircularHough (this work)	Cases	CircularHough (this work)	Cases	CircularHough (this work)
0-0	151	96.03%	186	93.55%	203	98.03%
1-0	26	96.15%	68	92.65%	48	93.75%
1-1	2	100.0%	1	100.0%	2	100.0%
1-2	2	100.0%	2	100.0%	2	100.0%
2-0	49	89.80%	62	96.77%	71	98.59%
2-1	13	100.0%	9	88.89%	6	66.67%
2-2	8	100.0%	20	95%	9	88.89%
3-0	48	95.83%	21	90.48%	38	86.84%
3-1	26	88.46%	7	85.71%	9	77.78%
3-2	75	100.0%	24	79.17%	12	91.67%
Total	400	95.75%	400	92.75%	400	95.25%

Table 3.4: *Extended results of the OD segmentation methodology obtained by subjective evaluation. Results are detailed per diagnosis case and ophthalmic center and are given as the average percent of obtained good OD segmentations.*

described approach was measured. Lalonde et al. (2001) proposed to use the Jaccard coefficient (Jaccard, 1901) with this purpose: an overlapping score S is defined to measure the common area between a true OD region T and a detected region D as

$$S = \frac{Area\,(T \cap D)}{Area\,(T \cup D)} \tag{3.14}$$

Thus, the *CircularHough* algorithm presented in this chapter was applied to calculate the overlapping S for the 1200 images in the MESSIDOR database. Table 3.5 shows the measured average S, $\overline{S}$, between the automatically-generated segmentations and the true OD regions in "gold standard" images detailed per diagnosis case of DR Grade - Risk of ME. As it can be observed, the segmentation methodology obtained a total average overlapping of 0.86. Classifying the images by retinas affected by DR or not, measured average overlapping for both cases was 0.85 and 0.87 respectively. For patients affected by DR and with risk of ME, the average overlapping was 0.86, and for patients affected by DR and with no risk of ME this figure was 0.85. These figures lead to similar conclusions to those reached with the subjective evaluation: the measured values of average overlapping for healthy retinas and for retinas affected by DR with risk of ME show evidence that the segmentation algorithm performance

DR Grade - Risk of ME	Cases	CircularHough (this work)
0-0	540	$\overline{S}_{0-0} = 0.87$
1-0	142	$\overline{S}_{1-0} = 0.86$
1-1	5	$\overline{S}_{1-1} = 0.80$
1-2	6	$\overline{S}_{1-2} = 0.90$
2-0	182	$\overline{S}_{2-0} = 0.87$
2-1	28	$\overline{S}_{2-1} = 0.88$
2-2	37	$\overline{S}_{2-2} = 0.89$
3-0	107	$\overline{S}_{3-0} = 0.83$
3-1	42	$\overline{S}_{3-1} = 0.83$
3-2	111	$\overline{S}_{3-2} = 0.84$
Total	1200	$\overline{S} = 0.86$

Table 3.5: *Results of the OD segmentation methodology obtained by objective evaluation. Results are detailed per diagnosis case and are given as the average common overlapping between the automatically-generated segmentations and the true OD contours.*

	Hôpital Lariboisière Paris		Faculté de Médecine St. Etienne		LaTIM - CHU de Brest	
DR Grade - Risk of ME	Cases	CircularHough (this work)	Cases	CircularHough (this work)	Cases	CircularHough (this work)
0-0	151	$\overline{S}_{0-0} = 0.86$	186	$\overline{S}_{0-0} = 0.86$	203	$\overline{S}_{0-0} = 0.88$
1-0	26	$\overline{S}_{1-0} = 0.87$	68	$\overline{S}_{1-0} = 0.84$	48	$\overline{S}_{1-0} = 0.86$
1-1	2	$\overline{S}_{1-1} = 0.88$	1	$\overline{S}_{1-1} = 0.69$	2	$\overline{S}_{1-1} = 0.83$
1-2	2	$\overline{S}_{1-2} = 0.86$	2	$\overline{S}_{1-2} = 0.93$	2	$\overline{S}_{1-2} = 0.94$
2-0	49	$\overline{S}_{2-0} = 0.86$	62	$\overline{S}_{2-0} = 0.87$	71	$\overline{S}_{2-0} = 0.88$
2-1	13	$\overline{S}_{2-1} = 0.90$	9	$\overline{S}_{2-1} = 0.88$	6	$\overline{S}_{2-1} = 0.84$
2-2	8	$\overline{S}_{2-2} = 0.90$	20	$\overline{S}_{2-2} = 0.90$	9	$\overline{S}_{2-2} = 0.83$
3-0	48	$\overline{S}_{3-0} = 0.86$	21	$\overline{S}_{3-0} = 0.83$	38	$\overline{S}_{3-0} = 0.79$
3-1	26	$\overline{S}_{3-1} = 0.84$	7	$\overline{S}_{3-1} = 0.87$	9	$\overline{S}_{3-1} = 0.79$
3-2	75	$\overline{S}_{3-2} = 0.89$	24	$\overline{S}_{3-2} = 0.79$	12	$\overline{S}_{3-2} = 0.85$
Total	400	$\overline{S} = 0.87$	400	$\overline{S} = 0.85$	400	$\overline{S} = 0.85$

Table 3.6: *Extended results of the OD segmentation methodology obtained by objective evaluation. Results are detailed per diagnosis case and ophthalmic center and are given as the average common overlapping between the automatically-generated segmentations and the true OD contours.*

is stable against illness presence or absence. Table 3.6 completes the information in Table 3.5 showing the obtained results detailed per ophthalmic center.

It is important to point out that, since the proposed algorithm segments the OD by approximating its shape by a circumference, for a better evaluation of its behaviour, it is interesting to get to know how far its results are from the maximum results reachable with

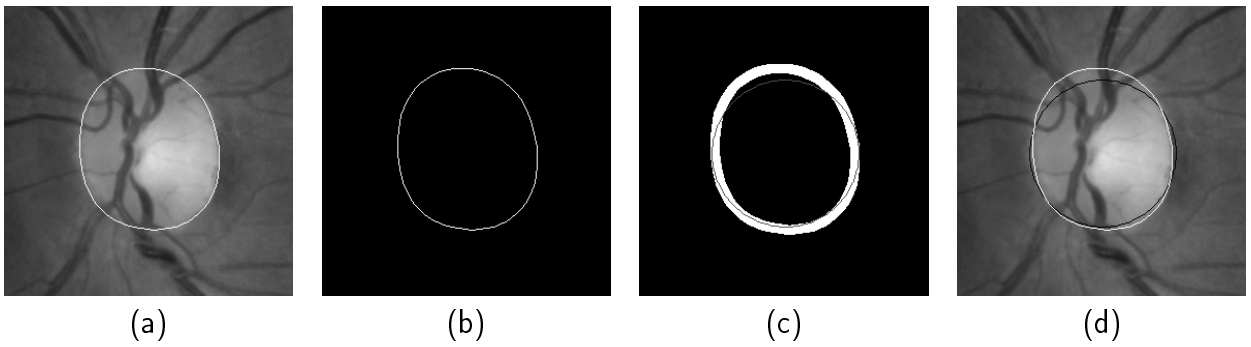


Figure 3.9: *Illustration of the obtention of a circular gold standard: (a) Optic disc with its contour traced by medical experts in white color. (b) Gold standard image of the optic disc contour. (c) In white color is represented the optic disc contour in (b) dilated to* 10 *pixels in width; in dark gray color is represented the obtained best circular approximation to the dilated contour, i.e., the circular gold standard. (d) Optic disc with the true contour traced in white color and the obtained circular gold standard traced in black color.*

this circular template-based approach. With this purpose, a "circular gold standard" set was created by calculating the best circular approximations for all true OD contours in the "gold standard" set by using (3.13). For this, true OD contours were dilated to have 10 pixels in width so as to get tolerance to obtain the best circular adjustment to shapes (see illustration in Figure 3.9). Then, the common area between the regions in the "gold standard" and these best circular approximations in the "circular gold standard" were calculated also according to (3.14) for all the images in the MESSIDOR database. Therefore, the average of these values may be considered the upper limit of average common overlapping for an automatic OD segmentation using a circular approximation. So, generalizing, this experiment, apart from being interesting for better algorithm evaluation, provides an interesting objective measurement of the maximum OD area which can be covered by a circle. Therefore, it is an appropriate measure to assess the general suitability of OD segmentation using a circular approximation.

Table 3.7 summarizes the results for the presented segmentation algorithm (*CircularHough*) and for the "circular gold standard" set (*CircularGoldStandard*). This table shows the percentage of images for different intervals of S values, as well as average overlapping $\overline{S}$ for the whole set of images. These overlapping measures corresponding to the *CircularHough* algorithm normalized by the results of the *CircularGoldStandard* are also shown in the last row. As shown in this table, overlapping between the hand-labeled OD region and *CircularHough* algorithm-segmented one is higher than or equal to 0.75 for 90% of the images in the

	$S \geq 0.95$	$S \geq 0.90$	$S \geq 0.85$	$S \geq 0.80$	$S \geq 0.75$	$S \geq 0.70$	$\overline{S}$
CircularHough (this work)	**7%**	**46%**	**73%**	**84%**	**90%**	**93%**	**0.86**
Hausdorff-BasedTemplateMatching (Lalonde et al., 2001)*	2%	25%	45%	70%	77%	85%	0.81
"CircularGoldStandard"	27%	87%	96%	99%	99%	100%	0.92
CircularHough / "CircularGoldStandard"	26%	53%	76%	85%	91%	93%	0.93

* This results were obtained on a local database composed by 40 images.

Table 3.7: *Results of CircularHough in terms of percentage of images per overlapping interval and average overlapping of the whole set.*

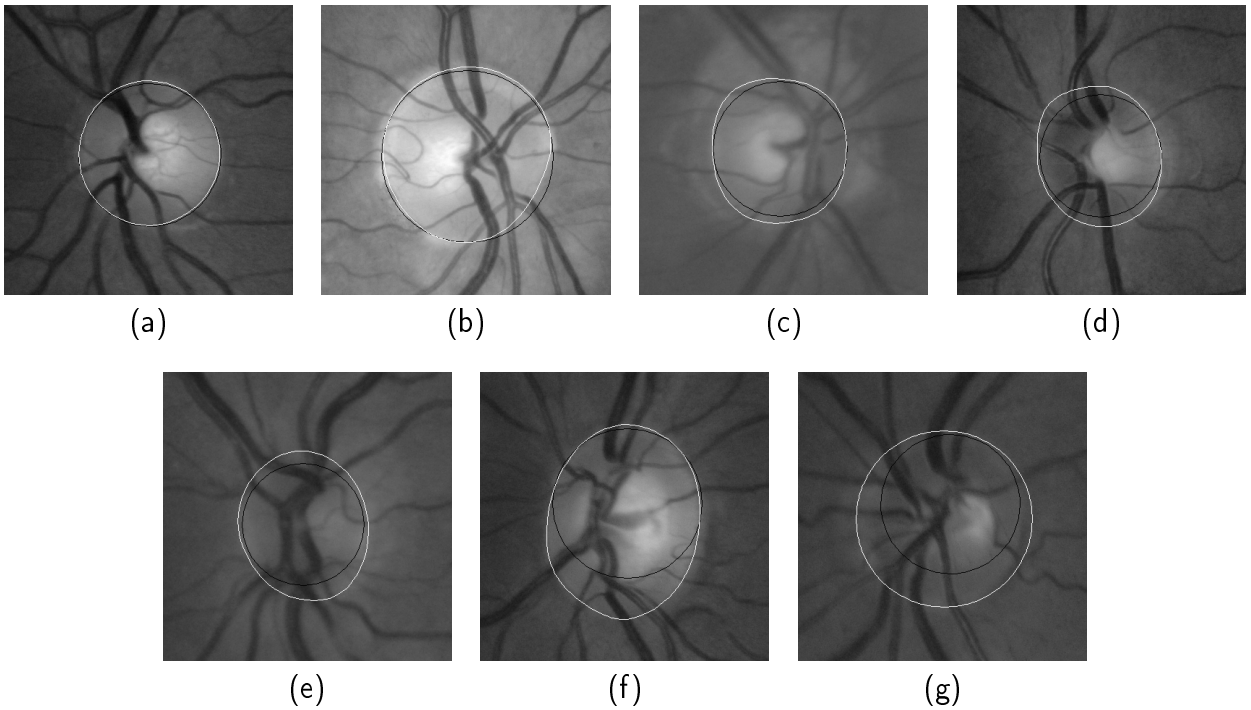


Figure 3.10: *Sample segmentations obtained with the presented methodology. (a) $S = 0.96$. (b) $S = 0.93$. (c) $S = 0.89$. (d) $S = 0.83$. (e) $S = 0.76$. (f) $S = 0.74$. (g) $S = 0.65$. Black line: segmentations produced by the algorithm; white line: true OD areas labeled by specialists.*

database. Average overlapping obtained for the whole set of images was 0.86. In this respect, note that the maximum overlapping reachable with a circular template-based approach is 0.92. Figure 3.10 shows segmentation examples obtained by the proposed methodology for all the overlapping intervals defined in Table 3.7.

The second row in Table 3.7 shows the performance results published by Lalonde et al. (Lalonde et al., 2001) for their circular template-based methodology. These results were obtained on a different database comprising 40 images taken from 27 patients attending a

diabetic retinal-screening programme. As in the case of the MESSIDOR database, this set of images includes good, fair and bad visual quality images (Lalonde *et al.* considered bad quality images as those blurred and/or containing abnormally dark or bright regions). Our algorithm could not be tested on this database as we had no access to it. The algorithm proposed by Lalonde *et al.* (Hausdorff-BasedTemplateMatching) is based on minimizing the Hausdorff distance between previously obtained candidate edge map regions and a set of circular templates. As it can be observed, the *CircularHough* method exceeds the results of this methodology for all considered overlapping intervals and also enhances average overlapping.

Finally, regarding algorithm computational efficiency, tests using the same PC architecture and features used for evaluating performance of the location methodology (it is defined at the end of section 3.4.1) were performed. For a total of 1200 executions of the algorithm, the average computational time obtained was 5.69 sec. with a standard deviation of 0.54 sec.

3.5 Discussion on optic disc segmentation models

In this section, we discuss on the advantages and disadvantages of the most widely-used OD segmentation models: circular and elliptical template-based methods as well as deformable model-based methods. The first discussion focuses on the use of elliptical or circular models. For that, behaviour of both approaches is studied and compared using the circular approximation presented here and four different elliptical implementations. In the second subsection, the proposed circular approach and three deformable model-based algorithms are used to compare advantages and disadvantages of both models.

3.5.1 Circular template-based approach vs elliptical template-based approach

According to OD shape features, the use of an elliptical model for OD segmentation, at least theoretically, should render higher segmentation potential. For empirical evaluation of this issue, the experiment used for circular model potential assessment was now reproduced and applied to elliptical models. That is, an "elliptical gold standard" set was generated by applying an elliptical model for fitting the true OD contours in the "gold standard" set of the MESSIDOR database described in appendix A (see illustration of the obtention of elliptical gold standards in Figure 3.11). So, the average overlapping of both sets may be considered the maximum segmentation potential of an elliptical model. Table 3.8 summarizes

	$\overline{S}$
Deformable Models	100%
Elliptical Template-Based Models	97%
Circular Template-Based Models	92%

Table 3.8: *Comparison of segmentation potential of different models in terms of average overlapping.*

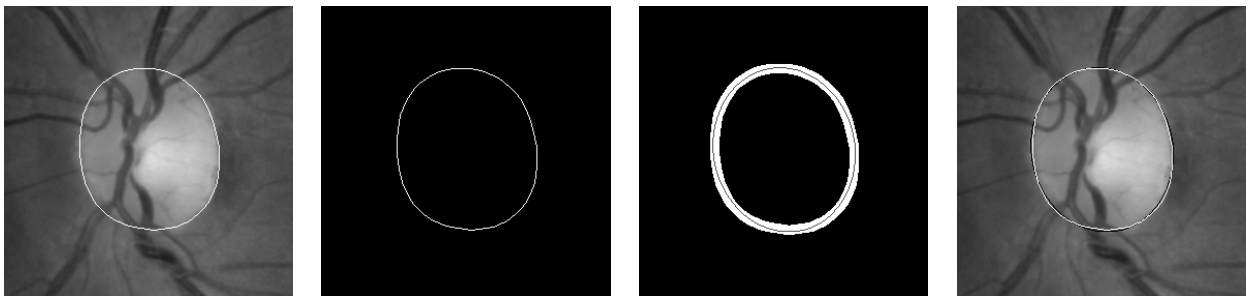


Figure 3.11: *Illustration of the obtention of an elliptical gold standard: (a) In white color is represented the optic disc contour traced by medical experts. (b) Gold standard image of the optic disc contour. (c) In white color is represented the optic disc contour in (b) dilated to* 10 *pixels in width; in dark gray color is represented the obtained best elliptical approximation to the dilated contour, i.e., the elliptical gold standard. (d) Optic disc with the true contour traced in white color and the obtained elliptical gold standard traced in black color.*

the potential of all models discussed in this section. As it can be observed, the use of an elliptical model renders a theoretical overlapping increase of 5% relative to the potential rendered by a circular model. Nevertheless, since the elliptical approach has two degrees of freedom more than the circular model (rotation angle and additional radius), its results may show a less robust and stable behaviour.

To assess elliptical model behaviour, four elliptical template-based variants were implemented and applied to the I_{BM} obtained from operation (3.11). Three of these approaches were based on minimizing the algebraic distance (Fitzgibbon et al., 1999; Gander et al., 1994) and one was based on minimizing the geometric distance (Gander et al., 1994). Each approach was tested on the MESSIDOR database and results were measured using the metric defined in (3.14). Thus, the elliptical and circular approaches can be compared under the same conditions. Table 3.9 summarizes the results obtained in this test. Data are presented in columns as the percentage of images per interval of overlapping values, the penultimate column is the total average overlapping and the last one indicates the percentage of model-

	$S \geq 0.95$	$S \geq 0.90$	$S \geq 0.85$	$S \geq 0.80$	$S \geq 0.75$	$S \geq 0.70$	$\overline{S}$	SF^*
EllipticalGeometricalMin (Gander et al., 1994)	2%	11%	20%	30%	39%	51%	0.67	89%
EllipticalAlgebraicMin1 (Fitzgibbon et al., 1999)	2%	9%	17%	26%	36%	48%	0.66	99%
EllipticalAlgebraicMin2 (Gander et al., 1994)	2%	10%	19%	27%	36%	47%	0.65	87%
EllipticalAlgebraicMin3 (Gander et al., 1994)	2%	9%	18%	26%	35%	48%	0.66	98%
CircularHough (this work)	**7%**	**46%**	**73%**	**84%**	**90%**	**93%**	**0.86**	**100%**

* SF (Solutions Found): percentage of cases for which the model found a solution.

Table 3.9: *Comparison between the presented circular template-based method and four elliptical template-based approaches. Data in terms of percentage of images per overlapping interval and average overlapping of the whole set.*

solved cases, as the tested methods do not always converge to a solution. The first row shows the results for the elliptical approach based on minimizing the geometric distance, the following three rows show the results obtained with the three variants based on minimizing the algebraic distance, and the last row show the results of the circular approach presented in this chapter. Importantly, the average values for each elliptical method were calculated considering only the cases for which they had found a solution. Another important issue with regard to results is that no automated criterion was used for selecting for each image between the segmentations from the red and green channels. For each image, both candidates were measured and the one with the highest overlapping score was selected for method total average calculation. Hence, the results shown in Table 3.9 are the highest possible scores for each elliptical approach. Figure 3.12 shows segmentation examples obtained with the elliptical approaches per considered overlapping interval.

The results from this study indicate that the segmentation performance of all elliptical approaches is similar. The greater difference between them was observed at the percentage of solved cases. Interestingly, their performance is significantly poorer than that obtained with the circular model. This performance degradation is the result of elliptical models' higher sensitivity to poor OD border contrast and, therefore, to poor border extraction. These models require the extraction of a great amount of OD representative borders to provide correct estimations of the two radiuses and rotation angle. On the other hand, the circular model only needs some portions of the OD contour to obtain a fitting circumference (see

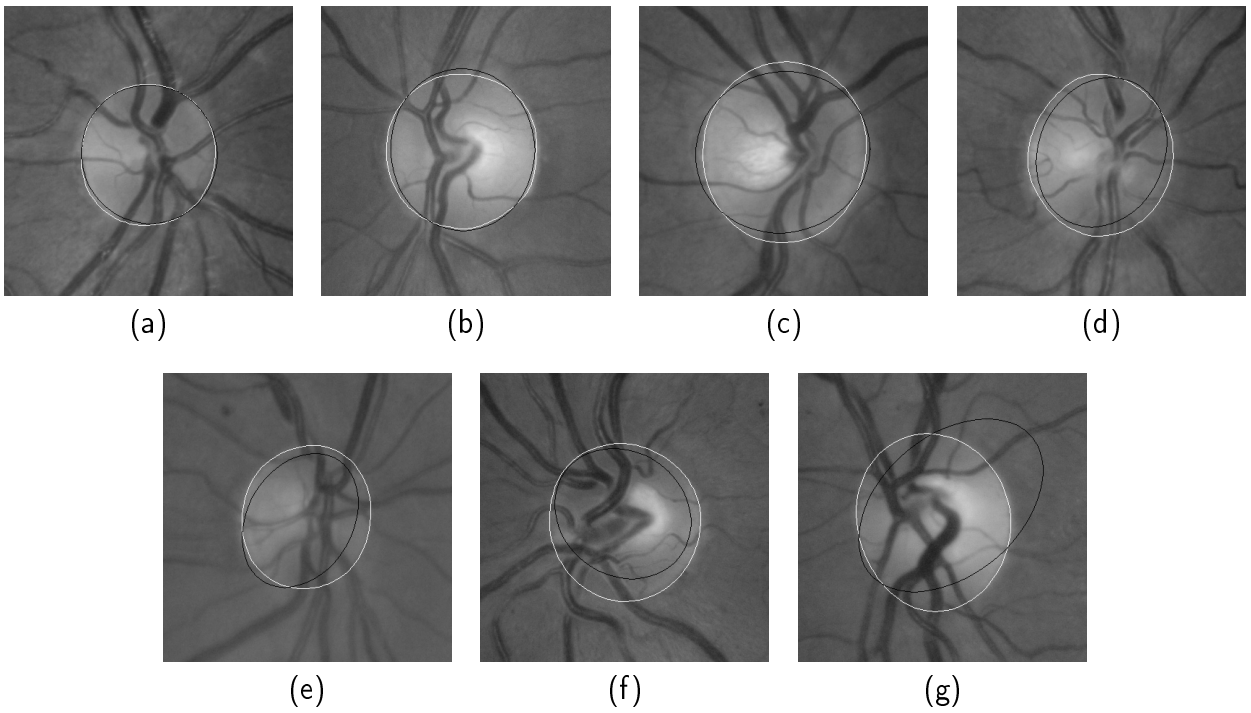


Figure 3.12: *Sample segmentations obtained with the elliptical approaches. (a)* $S = 0.96$. *(b)* $S = 0.93$. *(c)* $S = 0.89$. *(d)* $S = 0.83$. *(e)* $S = 0.76$. *(f)* $S = 0.74$. *(g)* $S = 0.65$. *Black line: segmentations produced by the algorithms; white line: true OD areas labeled by specialists.*

illustration in Figure 3.13). Although the preprocessing designed in this work to be used in the context of a circular model may not be the most appropriate for elliptical models, and maybe it has not maximized their potential, the behaviour comparison between models shown here suggests that the direct application of elliptical models for OD segmentation offer less stable and homogeneous behaviour.

3.5.2 Circular template-based approach vs deformable model-based approach

The main advantage of using a deformable model instead of a template-based model for OD segmentation is that, theoretically, 100% of overlapping areas between the automated segmentation and the ground truth may be achieved. As shown in Table 3.8, it involves an 8% increase relative to a circular model. This is why deformable models have much more degrees of freedom than template-based models to fit the desired shape. However, these additional degrees of freedom make these models more sensitive to irregular or low OD boundary contrast.

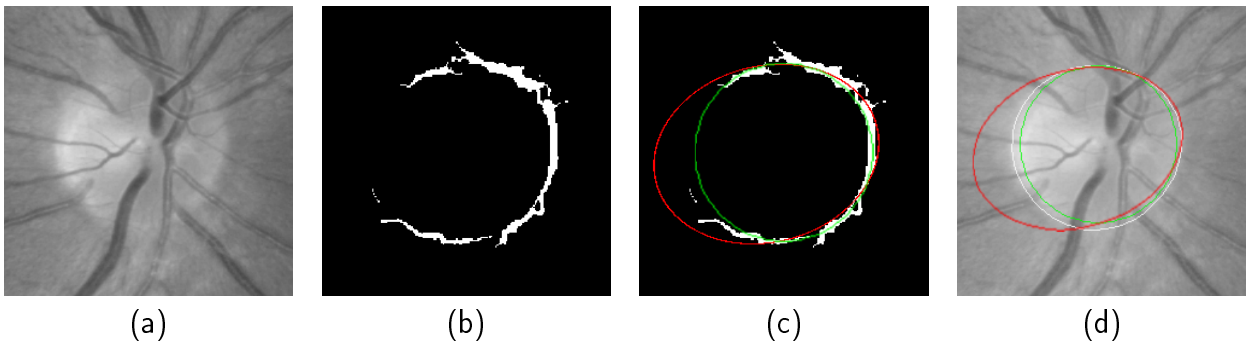


Figure 3.13: *Example of the sensibility of elliptical models to poor OD border extraction. (a) Original image containing an OD. (b) Extracted OD contour. (c) Results of the elliptical fitting, in red color, and the circular fitting, in green color. (d) Original image with the results shown in (c) and the true OD contour traced in white color.*

As a basis for this discussion, the circular template-based method proposed here was compared to the three OD segmentation approaches based on deformable models proposed in Lowell et al. (2004a). For the sake of comparison rigorousness, we used the same database, "gold standard" set and metric as those used in this work. Thus, the algorithms can be compared under identical conditions. The database used is the LOWELL database described in appendix A. The "gold standard" set used by these authors is composed of four manual segmentations of the OD contour made by clinicians. With respect to the metric of segmentation quality, they calculated the mean and radial standard deviations of these contours and then, they defined the discrepancy metric δ^j as

$$\delta^j = \sum_i \frac{m_i^j - \mu_i^j}{\sigma_i^j - \epsilon} \tag{3.15}$$

where μ_i^j and σ_i^j summarize the clinicians choice of rim location on spoke i of image j and m_i^j is segmentation location on spoke i for image j. Spokes are points belonging to the OD rim; 24 spokes were considered taken each 15°. Division by σ compensates for uncertainty in rim position and ϵ is a small factor to prevent division by zero where the clinicians are in exact agreement. In addition to this measure, they defined four categories corresponding to their subjective perception of quality. These four categories are *Excellent*, *Good*, *Fair* and *Poor*, referencing images with disparity up to one, two, five, or more, respectively. They assessed their algorithms performance considering the percentage of segmentations classified

	Excellent	Good	Fair	Poor	Excellent-Fair
CircularHough (this work)	40%	39%	18%	3%	**97%**
TemporalLock (Lowell et al., 2004a)	42%	31%	10%	17%	83%
DV-Hough (Lowell et al., 2004a)	39%	22%	20%	19%	81%
Simple (Lowell et al., 2004a)	9%	8%	30%	53%	47%

Table 3.10: *Comparison between the presented circular template-based method and three deformable model-based approaches in terms of percentage of images per subjective category.*

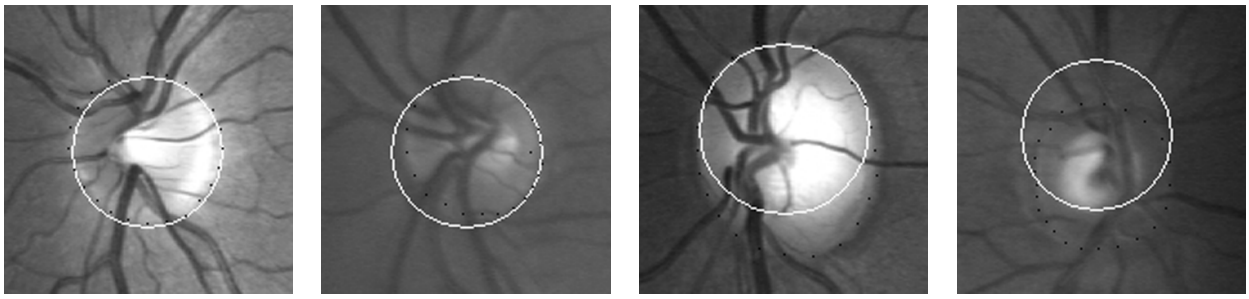


Figure 3.14: *Sample segmentations. (a) Excellent:* $\delta = 0.67$. *(b) Good:* $\delta = 1.62$. *(c) Fair:* $\delta = 3.17$. *(d) Poor:* $\delta = 5.69$. *The solid white lines are algorithm segmentations and the black dotted lines are clinician boundaries.*

in the range *Excellent* - *Fair*. Figure 3.14 shows segmentation examples produced by our algorithm in each defined subjective category.

Table 3.10 summarizes the performance on this subjective scale for their three alternative algorithms and the one proposed in this chapter. Results are expressed in terms of the percentage of images per subjective category. As it can be observed, with the best deformable model approach, 2% more of excellent segmentations were obtained than with the presented approach. Nevertheless, the template-based approach provided a significant enhancement in the percentage of obtained "valid" segmentations, thus reaching the 79% vs 73% of segmentations within the *Excellent* - *Good* range, and 97% vs 83% of segmentations within the *Excellent* - *Fair* range.

The results of this experiment, in spite of not being applicable to all template-based and deformable models, are a good example of comparable segmentation performance of both approaches on a particular common case. Therefore, the main conclusion of this experiment would be that, for OD segmentation, under appropriate OD background-contrast conditions, deformable models render more accurate OD segmentations. On the contrary case, when

contrast conditions are not so favourable, the circular approach may turn out to be a more robust and reliable solution.

3.6 Discussion and conclusion

This chapter presents a new template-based approach for OD segmentation by means of a circular OD-boundary approximation. In addition, an OD location methodology for obtaining the OD position needed by the segmentation algorithm as initial information is also proposed.

The location methodology obtained successful OD locations in 98.83% of the 1200 images in the MESSIDOR database (see it description in appendix A). As shown in tables 3.1 and 3.2, the methodology rendered a robust behaviour against all clinic case situations. Regarding the methodology efficiency, the measured average computational time was 1.67 sec. with a standard deviation of 0.14 sec. These results make the presented tool an attractive solution for locating the OD in digital retinal images.

The segmentation methodology was subjectively and objectively evaluated on the MESSIDOR database. Both evaluations, according to data in tables 3.3, 3.4, 3.5 and 3.6, lead to the same conclusion: the proposed methodology offers a reliable and robust solution for OD segmentation in retinographies from healthy patients as well as from patients affected by DR with risk of ME. Focusing on results of objective evaluation and according to the results in Table 3.7, the overlapping between the "true" OD region and the one segmented by our algorithm is over 0.75 in 90% of the 1200 images, being the average overlapping 0.86 for the whole set of images. This result is more valuable taking into account that maximum overlapping with a circular template-based approach is 0.92. Referentially, it can be pointed out that performance comparisons with the circular template-based approach by Lalonde et al. (2001) show that the overlapping obtained by our proposal was clearly higher, in spite of the fact that our results were obtained on a set of 1200 images while Lalonde *et al.* used only 40 (results are shown in Table 3.7). Besides these results, the measured average computational time of the methodology implementation, 5.69 sec. with a standard deviation of 0.54 sec., indicates that the proposed tool also offers an efficient solution for OD segmentation.

Moreover, discussion in section 3.5 on different OD segmentation models, strengthens the hypothesis of the suitability of circular models for this purpose. The tests summarized in

Table 3.8 indicate that the overlapping area between ground-truth OD segmentations and those obtained by elliptical models is higher than those obtained by circular models (97% and 92%, respectively). However, when testing with numerous and different real cases, the simplicity of circular models generally favours obtaining a more robust behaviour. Thus, performance comparisons between the proposed circular template-based method and four elliptical template-based approaches for the MESSIDOR images (Table 3.9) indicate that the circular approach renders a clearly higher average overlapping. The same conclusion was drawn for deformable models. Although these models can theoretically obtain 100% overlapping, the performance results presented in terms of discrepancy grades in Table 3.10, indicate that our proposed circular model outperforms the deformable model proposed by Lowell et al. (2004a).

On the other hand, it must be mentioned that, to facilitate performance comparison between OD segmentation methods, the generated hand segmentations of the OD rim for the 1200 images in the MESSIDOR database are currently available for researchers at `http://www.uhu.es/retinopathy`, in the Sample Databases subsection within the Results section by selecting the MESSIDOR database. The original database is available at: `http://messidor.crihan.fr/download-en.php`. To the best of our knowledge, such a set of "gold standard" images from a clinically labeled database is not available for the research community at the present time.

The main conclusions of this chapter can be summarized as follows:

1. The performance results obtained by the proposed methodology on a huge digital retinal database indicate that simple methods, based on basic image processing techniques, seem to suffice for OD location and segmentation.

2. A circular modelling for the OD boundary, compared to elliptical and deformable models, was shown to offer good compromise between success rate, quality and efficiency, as shown by comparing its segmented area to experts' free-drawn areas.

Despite all of this, the existence of some specific cases in which, due to their exceptional ellipsity degree, the circular approach does not reach the performance results of the elliptical approach has been reported (see, for instance, the case in Figure 3.10 - (f)). With the aim of enhancing the overlapping rates obtained in this work for these isolated cases, the authors' current research is focused on the development of a methodology for performing a controlled

elliptical deformation of the obtained circumference. Within the framework of this study, whether any preprocessing modification is necessary or even whether post-processing would be appropriate for assuring deformation process stability is currently under study.

Chapter 4

Vascular Tree Segmentation

RETINAL blood vessel assessment is an important diagnosis key to detect and evaluate many pathologies leading to vascular anomalies such as diabetic retinopathy. As a previous step, vessel assessment requires segmentation of the vascular tree from the background for further processing. In the field of automated detection of diabetic retinopathy, blood vessel segmentation has been reported useful to reduce the number of false positives in microaneurysm and haemorrhage detection as well as to detect other fundus features such as the optic disc and the macula.

This chapter presents a new methodology for blood vessel detection in digital retinal images. **Six sections** compose this chapter. The **first section** characterizes the retinal vascular tree in digital retinographies and introduces and motivates its automated segmentation.

The **second section** makes a review of the retinal vessel segmentation topic. In this section, the most relevant methods present in the literature are categorized and briefly summarized.

The new methodology for blood vessel segmentation, a supervised approach, is presented in the **third section**. This methodology uses a neural network scheme for pixel classification and computes a 7-dimensional vector composed of gray-level and moment invariants-based features for pixel representation.

The proposed segmentation method is evaluated and compared to other published works in the **fourth section**. The metrics used to objectively evaluate the algorithm performance were *sensitivity*, *specificity*, *positive predictive value* and *negative predictive value* for partial performance evaluation, and *accuracy* and *area under ROC curve* for global performance evaluation and comparison.

Finally, this chapter is closed in the **fifth section** with the author's conclusion and discussion.

4.1 Introduction

The retinal vascular tree can be recognized in digital eye fundus images as a net of connected either deep red or orange-red filaments originating within the optic disc (see Figure 4.1). Here, the vascular tree borns with some main wide vessels that progressively split into other thinner vessels.

As it was previously described in this thesis, the main cause of diabetic retinopathy (DR) is the abnormal elevation of glucose in the blood. It produces damages in the vessels endothelium, resulting in an increase in their permeability. The first manifestations of the disease are microaneurysms, which are tiny dilations of the capillaries. DR progression also causes in the retinal vascular net blood vessel changes in diameter, branching angles or tortuosity, as well as venous beading, neovascularization or haemorrhages. Therefore, since vascular anomalies are one of DR manifestations, automatic assessment of retinal blood vessels is necessary for automated detection of DR. As a previous step, vessel assessment demands vascular tree segmentation from the background for further processing.

Besides being a DR diagnosis key, segmentation of blood vessels has been also reported useful for other objectives related to automated detection of DR. Knowledge on blood vessel location can be used to reduce the number of false positives in microaneurysm and haemorrhage detection (Frame et al., 1998; Larsen et al., 2003; Niemeijer et al., 2005; Spencer et al., 1996). Furthermore, the vascular tree can also be useful as valuable information to locate other fundus features such as the optic disc (Foracchia et al., 2004; Hoover and Goldbaum, 2003; Youssif et al., 2008) and the macula (Li and Chutatape, 2004; Tobin et al., 2007). Moreover, besides these applications motivated by automated early detection of DR, vascular tree segmentation proves useful for other clinical purposes: extraction of reference points of the vasculature for image registration (Matsopoulos et al., 2004; Zana and Klein, 1999), evaluation of the retinopathy of prematurity (Heneghan et al., 2002), arteriolar narrowing (Grisan and Ruggeri, 2003; Hatanaka et al., 2004), vessel tortuosity to characterize hypertensive retinopathy (Foracchia et al., 2001), vessel diameter measurement for the diagnosis of hypertension and cardiovascular diseases (Goa et al., 2001; Lowell et al., 2004b;

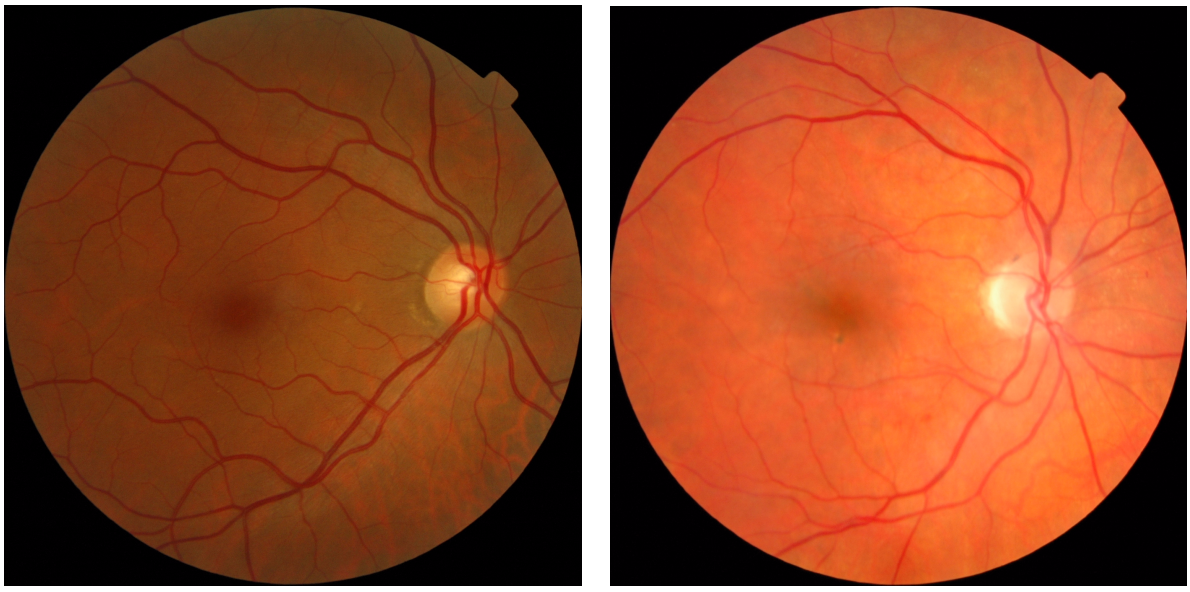


Figure 4.1: *Variability of the vascular tree in color appearance: (a) Deep red vascular tree. (b) Orange-red vascular tree.*

Martinez-Perez et al., 2002), and computer-assisted laser surgery (Becker et al., 1998; Shen et al., 2001).

4.2 State of art

Many methods for retinal vessel segmentation have been reported. These can be divided into two groups: *rule-based* methods and *supervised* methods. In the first group, methods using *vessel tracking*, *mathematical morphology*, *matched filtering*, *model-based locally adaptive thresholding* or *deformable models* are highlighted. On the other hand, *supervised* methods are those based on pixel classification (implementing some kind of classifier).

Regarding rule-based methods, vessel tracking methods (Can et al., 1999; Chutatape et al., 1998; Gagnon et al., 2001; Liu and Sun, 1993; Tolias and Panas, 1998; Zhou et al., 1994) attempt to obtain the vasculature structure by following vessel center lines. Starting from an initial set of points established automatically or by manual labeling, vessels are traced by deciding from local information the most appropriate candidate pixel from those close to that currently under evaluation. Other methods use mathematical morphology (Heneghan et al., 2002; Mendonça and Campilho, 2006; Walter and Klein, 2001; Zana and Klein, 2001) to benefit from a priori-known vasculature shape features, such as being piecewise linear and

connected. Then, by applying morphological operators, the vasculature is filtered from the background for final segmentation. Matched filtering techniques (Al-Rawi and Karajeh, 2007; Al-Rawi et al., 2007; Chaudhuri et al., 1989b; Cinsdikici and Aydın, 2009; Gang et al., 2002; Hoover et al., 2000) usually use a 2-dimensional linear structural element with a Gaussian cross-profile section, extruded or rotated into 3 dimensions for blood vessel cross-profile identification (typically a Gaussian or Gaussian-derivative profile). The kernel is rotated into many different orientations (usually 8 or 12) to fit into vessels of different configuration. The image is then thresholded to extract the vessel silhouette from the background. Regarding model-based locally adaptive thresholding, a general framework based on a verification-based multithreshold probing scheme was presented by Jiang and Mojon (2003). These authors enriched this generic methodology by incorporating relevant information related to retinal vessels into the verification process with the aim of enabling its application to retinal images. On the other hand, deformable models or *snakes* have been also used (Espona et al., 2007; McInerney and Terzopoulos, 2000). A snake is an active contour model that, once placed on the image near the contour of interest, can evolve to fit the shape of the desired structure by an iterative adaption. Other rule-based methods for retinal blood vessel segmentation were reported in Martinez-Perez et al. (1999, 2007) and Lam and Yan (2008). Martinez-Perez et al. (1999) proposed a method where features derived from image derivatives obtained at multiple scales are used in a two-stage region growing procedure which segments progressively the retinal vessels. Subsequently, Martinez-Perez et al. (2007) proposed another method based upon multiscale feature extraction. The local maxima over scales of the gradient magnitude and the maximum principal curvature of the Hessian tensor were used in a multiple pass region growing procedure. Growth progressively segmented the blood vessels by using both feature and spatial information. In the method presented by Lam and Yan (2008), blood vessel-like objects were extracted by using the Laplacian operator and noisy objects were pruned according to centerlines, detected by means of the normalized gradient vector field.

On the other hand, supervised methods are based on pixel classification, which consists on classifying each pixel into two classes, vessel and non-vessel. Classifiers are trained by supervised learning with data from manually-labeled images. Gardner et al. (1996) proposed a back propagation multilayer neural network (NN) for vascular tree segmentation. After histogram equalization, smoothing and edge detection, the image was divided into 20×20 pixel squares (400 input neurons). The NN was then fed with the values of these pixel

windows for classifying each pixel into vessel or not. Sinthanayothin et al. (1999) also used a multilayer perceptron NN. Each pixel in the image was classified by using the first principal component, and the edge strength values from a 10×10 pixel subimage centered on the pixel under evaluation, as input data. Niemeijer et al. (2004) implemented a K-nearest neighbor (kNN) classifier. A 31-component pixel feature vector was constructed with the Gaussian and its derivatives up to order 2 at 5 different scales, augmented with the gray-level from the green channel of the original image. The assumption that vessels are elongated structures is the basis for the supervised ridge-based vessel detection method presented by Staal et al. (2004). Ridges were extracted from the image and used as primitives to form line elements. Each pixel was then assigned to its nearest line element, the image thus being partitioned into patches. For every pixel, 27 features were firstly computed and those obtaining the best class separability were finally selected. Feature vectors were classified by using a kNN-classifier and sequential forward feature selection. Soares et al. (2006) used a Gaussian mixture model Bayesian classifier. Multiscale analysis was performed on the image by using the Gabor wavelet transform. The gray-level of the inverted green channel and the maximum Gabor transform response over angles at 4 different scales were considered as pixel features. Finally, Ricci and Perfetti (2007) used a support vector machine (SVM) for pixel classification as vessel or non-vessel. They used two orthogonal line detectors along with the gray-level of the target pixel to construct the feature vector.

4.3 Methodology

This chapter proposes a new supervised approach for blood vessel detection based on a NN for pixel classification. The necessary feature vector is computed from preprocessed retinal images in the neighborhood of the pixel under consideration. The following process stages may be identified: (A) original fundus image preprocessing for gray-level homogenization and blood vessel enhancement, (C) application of a classifier to label the pixel as vessel or non-vessel, and (D) postprocessing for filling pixel gaps in detected blood vessels and removing falsely-detected isolated vessel pixels.

Input images are monochrome and obtained by extracting the green band from original RGB retinal images. The green channel provides the best vessel-background contrast of the RGB-representation, while the red channel is the brightest color channel and has low

contrast, and the blue one offers poor dynamic range. Thus, blood containing elements in the retinal layer (such as vessels) are best represented and reach higher contrast in the green channel (Walter et al., 2007).

The methodology was tested on the retinal images in the DRIVE and STARE databases (see their description in appendix A). All parameters described below were set by experiments carried out on DRIVE images with the aim of contributing the best segmentation performance on this database (performance was evaluated in terms of *average accuracy* - a detailed description is provided in Sections 4.4.1 and 4.4.2). Therefore, they refer to retinas of approximately 540 pixels in diameter. The application of the methodology to retinas of different size (i.e., the diameter in pixels of STARE database retinas is approximately 650 pixels) demands either resizing input images to fulfil this condition or adapting proportionately the whole set of used parameters to this new retina size.

4.3.1 Preprocessing

Color fundus images often show important lighting variations, poor contrast and noise. In order to reduce these imperfections and generate images more suitable for extracting the pixel features demanded in the classification step, a preprocessing comprising the following steps is applied: (1) vessel central light reflex removal, (2) background homogenization, and (3) vessel enhancement. Next, a description of the procedure, illustrated through its application to a STARE database fundus image (Figure 4.2), is detailed.

Vessel central light reflex removal

Since retinal blood vessels have lower reflectance when compared to other retinal surfaces, they appear darker than the background. Although the typical vessel cross-sectional gray-level profile can be approximated by a Gaussian shaped curve (inner vessel pixels are darker than the outermost ones), some blood vessels include a light streak (known as a light reflex) which runs down the central length of the blood vessel.

To remove this brighter strip, the green plane of the image is filtered by applying a morphological opening using a 3-pixel diameter disc, defined in a square grid by using 8-connectivity, as structuring element. Disc diameter was fixed to the possible minimum value to reduce the risk of merging close vessels. I_γ denotes the resultant image for future references.

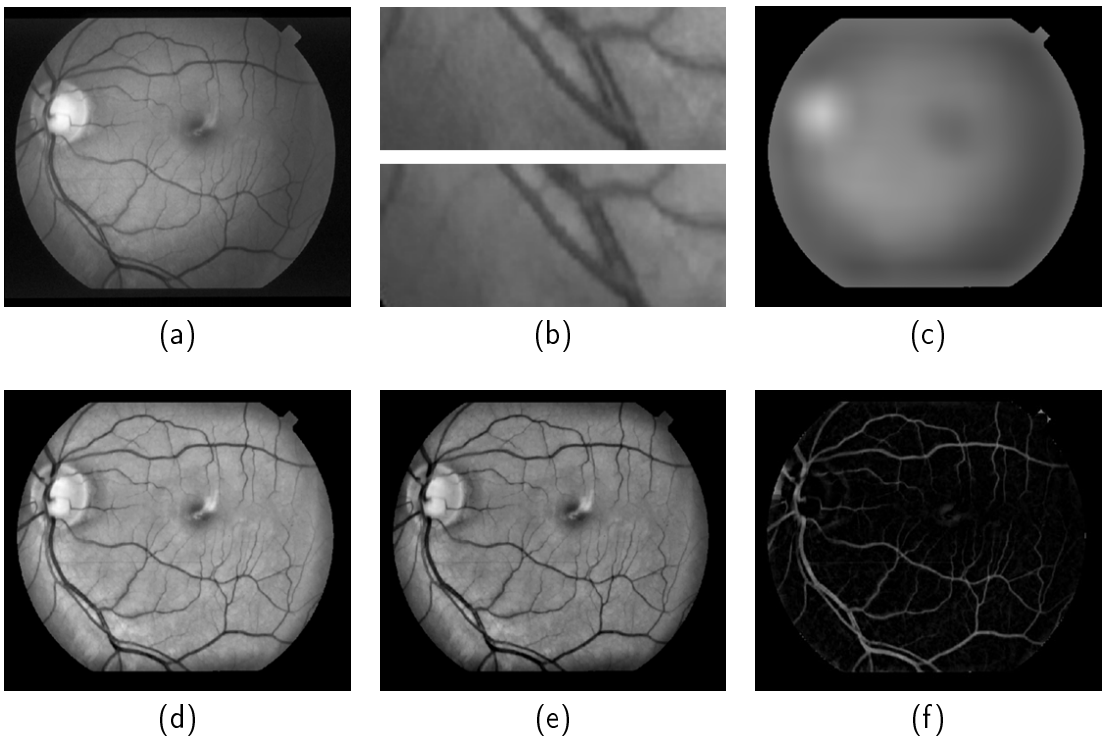


Figure 4.2: *Illustration of the preprocessing process: (a) Green channel of the original image. (b) The upper image is a fragment of the original image containing a vessel with central light reflex, while the bottom image shows the effect of reflex removal. (c) Background image. (d) Shade-corrected image. (e) Homogenized image. (f) Vessel-enhanced image.*

An example of vessel central light reflex and its removal from a fundus image by means of opening filtering operation is shown in Figure 4.2 (see images (a)-(b)).

Background homogenization

Fundus images often contain background intensity variation due to non-uniform illumination. Consequently, background pixels may have different intensity for the same image and, although their gray-levels are usually higher than those of vessel pixels (in relation to green channel images), the intensity values of some background pixels is comparable to that of brighter vessel pixels. Since the feature vector used to represent a pixel in the classification stage is formed by gray-scale values, this effect may worsen the performance of the vessel segmentation methodology. With the purpose of removing these background lightening variations, a shade-corrected image is accomplished from a background estimate. This image is the result of a filtering operation with a large arithmetic mean kernel, as described below.

Firstly, a 3×3 mean filter is applied to smooth occasional salt-and-pepper noise. Further noise smoothing is performed by convolving the resultant image with a Gaussian kernel of dimensions $m \times m = 9 \times 9$, mean $\mu = 0$ and variance $\sigma^2 = 1.8^2$, $G^m_{\mu,\sigma^2} = G^9_{0,1.8^2}$. Secondly, a background image I_B, is produced by applying a 69×69 mean filter (Figure 4.2, image (c)). When this filter is applied to the pixels in the field-of-view (FOV) near the border, the results are strongly biased by the external dark region. To overcome this problem, only the in-of-the FOV gray-levels are considered to calculate the mean value. Then, the difference D between I_γ and I_B is calculated for every pixel:

$$D(x,y) = I_\gamma(x,y) - I_B(x,y) \tag{4.1}$$

To this respect, literature reports shade-correction methods based on the subtraction of the background image from the original image (Frame et al., 1998; Niemeijer et al., 2005; Zhang et al., 2010) or the division of the latter by the former (Jagoe et al., 1992; Streeter and Cree, 2003). Both procedures rendered similar results upon testing. Moreover, none of them showed to contribute any appreciable advantage relative to the other. The subtractive approach in eq. (4.1) was used in the present work.

Finally, a shade-corrected image I_{SC} is obtained by transforming linearly RD values into integers covering the whole range of possible gray-levels ($[0 - 255]$, referred to 8-bit images). Figure 4.2 (image (d)) shows the I_{SC} corresponding to a non-uniformly illuminated image. The proposed shade-correction algorithm is observed to reduce background intensity variations and enhance contrast in relation to the original green channel image.

Besides the background intensity variations in images, intensities can reveal significant variations between images due to different illumination conditions in the acquisition process. In order to reduce this influence, a homogenized image I_H (Figure 4.2, image (e)) is produced as follows: the histogram of I_{SC} is displaced toward the middle of the gray-scale by modifying pixel intensities according to the following gray-level global transformation function:

$$g_{Output} = \begin{cases} 0 & \text{if } g < 0, \\ 255 & \text{if } g > 255, \\ \text{g} & \text{otherwise.} \end{cases} \tag{4.2}$$

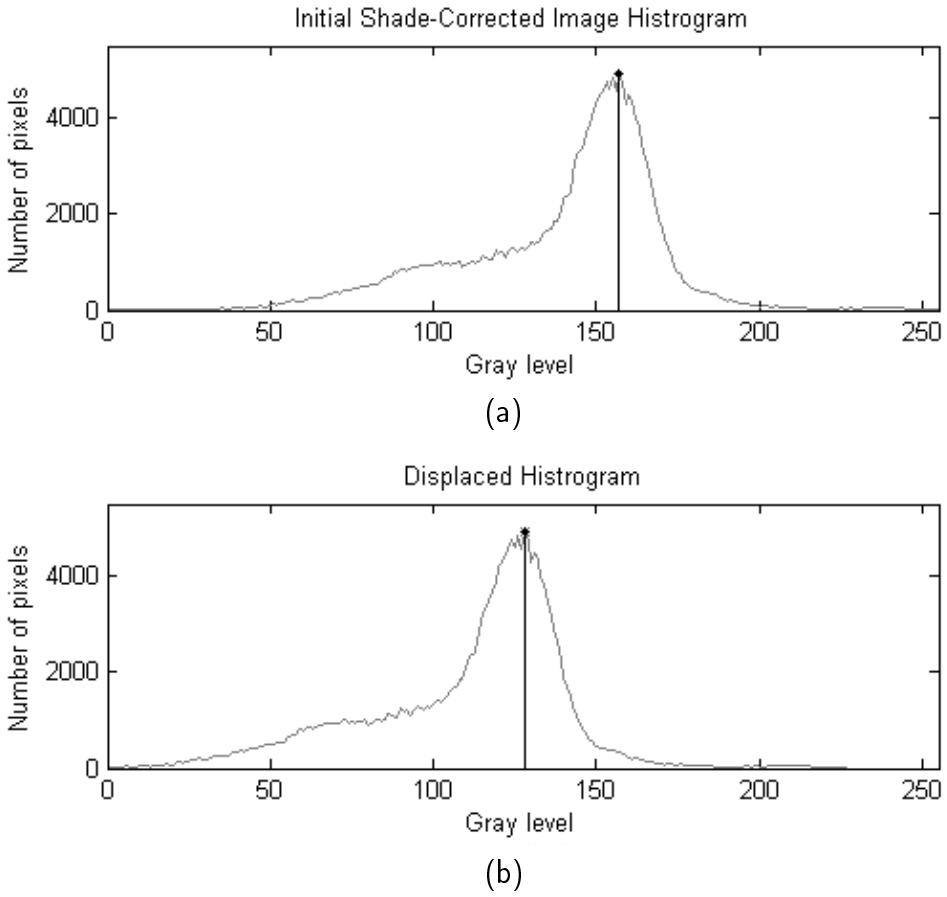


Figure 4.3: *(a) Histogram of the shade-corrected image. (b) Histogram of the homogenized image centered in the middle of the domain.*

where

$$g = g_{Input} + 128 - g_{Input\_Max} \tag{4.3}$$

g_{Input} and g_{Output} are the gray-level variables of input and output images (I_{SC} and I_{H}, respectively). The variable denoted by g_{Input_Max} defines the gray-level presenting the highest number of pixels in I_{SC}. By means of this operation, pixels with gray-level g_{Input_Max}, which are observed to correspond to the background of the retina, are set to 128 for 8-bit images. Thus, background pixels in images with different illumination conditions will standardize their intensity around this value. As an example, the histogram of the shade corrected image shown in Figure 4.2-(d), and the histogram of the homogenized image shown in Figure 4.2-(e) obtained by applying (4.2) are represented in Figure 4.3. Figure 4.4 (images (a), (b) and (d) (e)), shows the effect obtained with the described background homogenization operation on two fundus images in the STARE database.

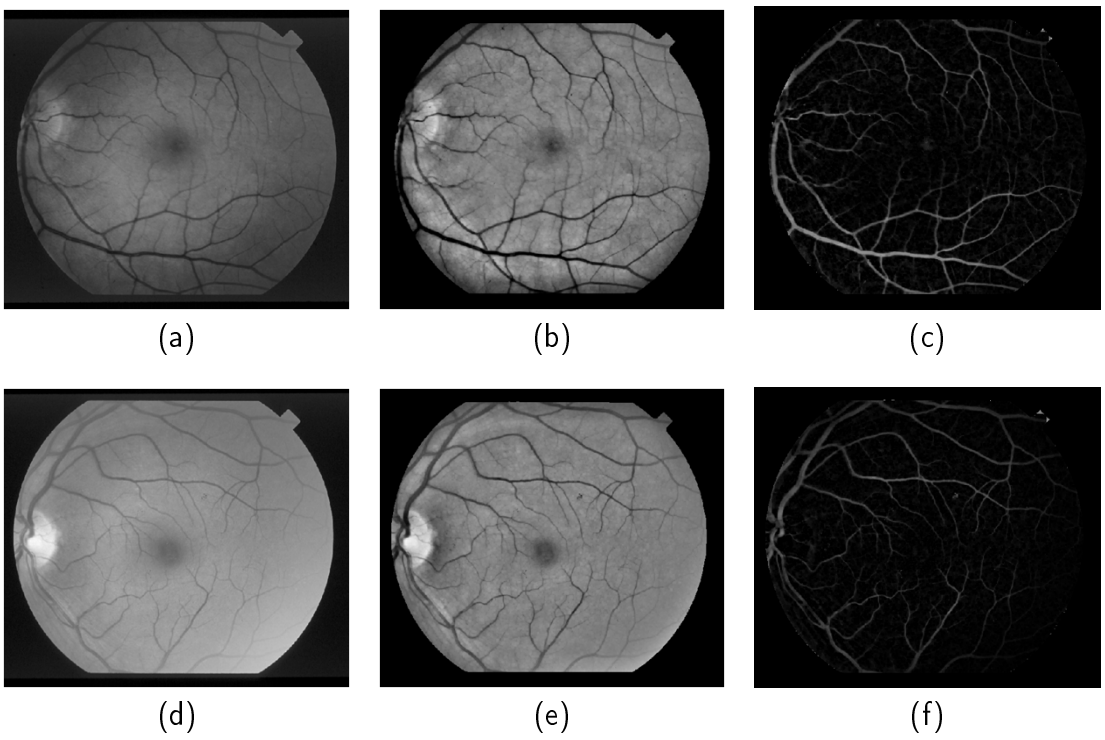


Figure 4.4: *Two examples of application of the preprocessing on two images with different illumination conditions. (a), (d) Green channel of the original images. (b), (e) Homogenized images. (c), (f) Vessel-enhanced images.*

Vessel enhancement

The final preprocessing step consists on generating a new vessel-enhanced image (I_{VE}), which proves more suitable for further extraction of moment invariants-based features (see section 4.3.2).

Vessel enhancement is performed by estimating the complementary image of the homogenized image I_H, I_H^c, and subsequently applying the morphological *Top-Hat transformation* (Figure 4.2, image (f)):

$$I_{VE} = I_H^c - \gamma\left(I_H^c\right) \tag{4.4}$$

where γ is a morphological opening operation using a disc of 8 pixels in radius. Thus, while bright retinal structures are removed (i.e., optic disc, possible presence of exudates or reflection artifacts), the darker structures remaining after the opening operation become enhanced (i.e., blood vessels, fovea, possible presence of microaneurysms or haemorrhages). Samples of vessel enhancement operation results are shown in Figure 4.4 (images (c) and (f)) for two fundus images with variable illumination conditions.

4.3.2 Feature extraction

The aim of the feature extraction stage is pixel characterization by means of a feature vector, a pixel representation in terms of some quantifiable measurements which may be easily used in the classification stage to decide whether pixels belong to a real blood vessel or not. In this work, the following sets of features were selected:

- *Gray-level-based features*: features based on the differences between the gray-level in the candidate pixel and a statistical value representative of its surroundings.

- *Moment invariants-based features*: features based on moment invariants for describing small image regions formed by the gray-scale values of a window centered on the represented pixels.

Gray-level-based features

Since blood vessels are always darker than their surroundings, features based on describing gray-level variation in the surroundings of candidate pixels seem a good choice. A set of gray-level-based descriptors taking this information into account were derived from homogenized images I_H considering only a small pixel region centered on the described pixel (x, y). $S^w_{x,y}$ stands for the set of coordinates in a $w \times w$ sized square window centered on point (x, y). Then, these descriptors can be expressed as:

$$f_1(x,y) = I_H(x,y) - \min_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} \tag{4.5}$$

$$f_2(x,y) = \max_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} - I_H(x,y) \tag{4.6}$$

$$f_3(x,y) = I_H(x,y) - \operatorname*{mean}_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} \tag{4.7}$$

$$f_4(x,y) = \operatorname*{std}_{(s,t)\in S^9_{x,y}} \{I_H(s,t)\} \tag{4.8}$$

$$f_5(x,y) = I_H(x,y) \tag{4.9}$$

Moment invariants-based features

The vasculature in retinal images is known to be piecewise linear and can be approximated by many connected line segments. For detecting these quasi-linear shapes, which are not all equally wide and may be oriented at any angle, shape descriptors invariant to translation, rotation and scale change may play an important role. Within this context, moment invariants proposed by Hu (1962) provide an attractive solution and are included in the feature vector. In this work, they are computed as follows.

Given a pixel (x, y) of the vessel-enhanced image I_{VE}, a subimage is generated by taking the region defined by $S^{17}_{x,y}$. The size of this region was fixed to 17×17 so that, considering that the region is centered on the middle of a "wide" vessel (8-9-pixel wide and referred to retinas of approximately 540 pixels in diameter), the subimage includes an approximately equal number of vessel and non-vessel pixels. For this subimage, denoted by $I_{VE}^{S^{17}_{x,y}}$, the 2-dimensional moment of order $(p+q)$ is defined as:

$$m_{pq} = \sum_i \sum_j i^p \times j^q \times I_{VE}^{S^{17}_{x,y}}(i,j) \quad p,q = 0,1,2,\ldots \tag{4.10}$$

where summations are over the values of the spatial coordinates i and j spanning the subimage, and $I_{VE}^{S^{17}_{x,y}}(i,j)$ is the gray-level at point (i,j).

The corresponding *central moment* is defined as

$$\mu_{pq} = \sum_i \sum_j (i-\bar{i})^p \times (j-\bar{j})^q \times I_{VE}^{S^{17}_{x,y}}(i,j) \tag{4.11}$$

where

$$\bar{i} = \frac{m_{10}}{m_{00}} \quad , \quad \bar{j} = \frac{m_{01}}{m_{00}} \tag{4.12}$$

are the coordinates of the center of gravity of the subimage. The normalized central moment of order $(p+q)$ is defined as

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma} \quad p,q = 0,1,2,\ldots \tag{4.13}$$

where

$$\gamma = \frac{p+q}{2} + 1 \; ; \quad (p+q) = 2,3,... \tag{4.14}$$

A set of seven moment invariants under size, translation and rotation, known as *Hu moment invariants*, can be derived from combinations of regular moments:

$$\phi_1 = \eta_{20} + \eta_{02} \tag{4.15}$$

$$\phi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \tag{4.16}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{4.17}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{4.18}$$

$$\begin{aligned}\phi_5 = {} & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \times \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right]\end{aligned} \tag{4.19}$$

$$\phi_6 = (\eta_{20} - \eta_{02}) \times \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \tag{4.20}$$

$$\begin{aligned}\phi_7 = {} & (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12}) \times \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right]\end{aligned} \tag{4.21}$$

Among these seven moment invariants, tests made revealed that moments ϕ_1 and ϕ_2 constitute the combination providing optimal performance in terms of average accuracy (see section 4.4.2). The inclusion of the remainder moments result in decreasing classification performance and increasing computation needed for classification. Moreover, the module of the logarithm was used instead of its values themselves. Using the logarithm reduces the dynamic range and the module prevents from having to deal with the complex numbers resulting from computing the logarithm of negative moment invariants. Figure 4.5 shows several samples of pixels, marked with solid white dots on an I_{VE} image (Figure 4.5, image (a)), and the subimages $I_{VE}^{S_{x,y}^{17}}$ generated around them (Figure 4.5, images (b)-(i)). Pairs of pixels were selected for different vessels: one inside and the other outside the vessel but near enough so that both subimages contain the vessel. Table 4.1 shows the moment values corresponding to each subimage. It can be checked that numbers are close, thus indicating a high degree of invariance to size, rotation and translation. Moments computed as mentioned above characterize numerically a vessel independently of its width, orientation and location in the subimage. However, they are not useful to describe the central pixel of the subimage in terms of vessel or non-vessel, as their values do not distinguish between these two situations.

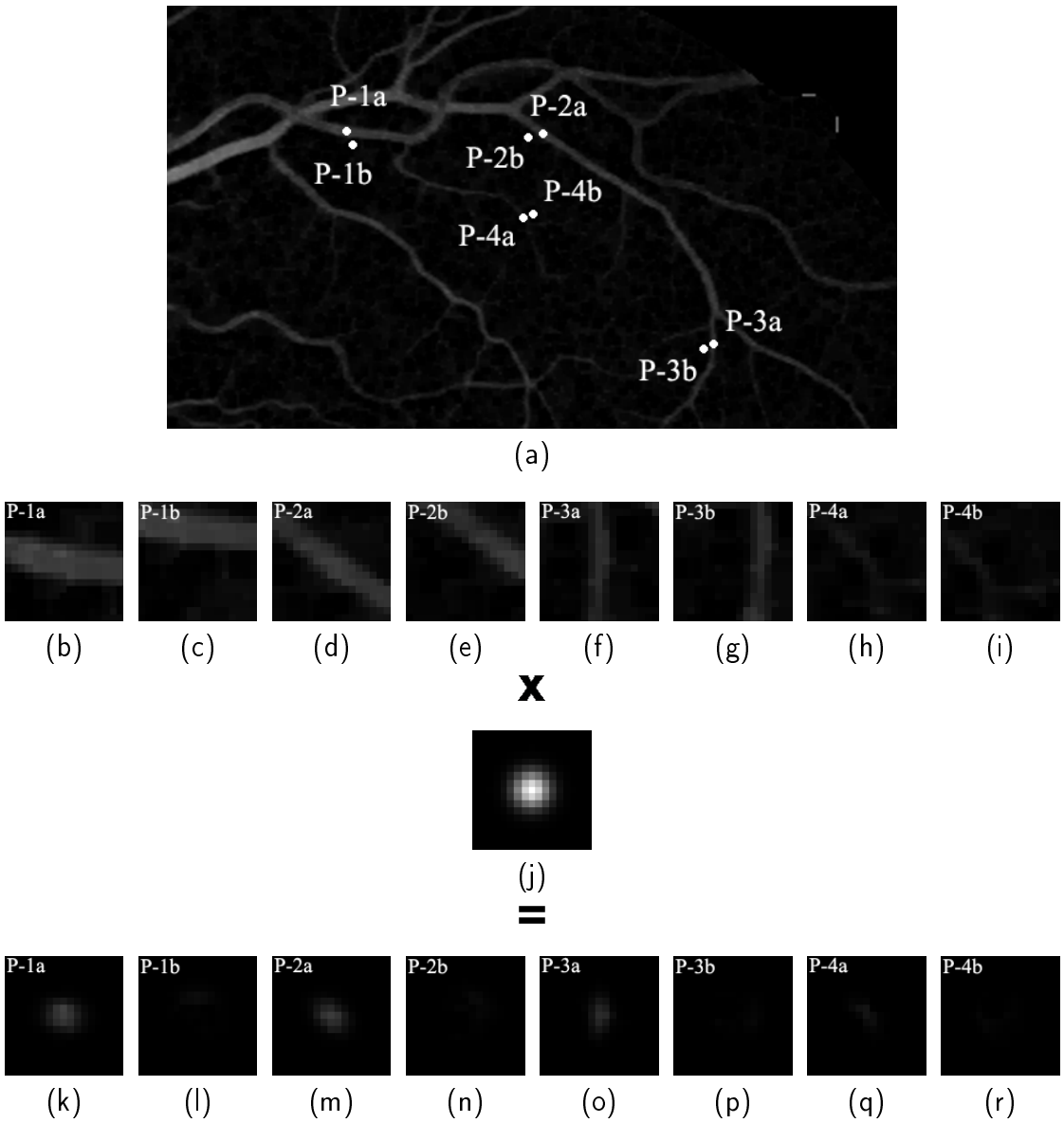


Figure 4.5: *Examples of obtaining pixel environments for moment invariants calculation. (a) Vessel enhanced subimage. Four pairs of pixels are marked with white dots: P-ka and P-kb with* $k = 1, 2, 3, 4$*; P-ka are vessel pixels and P-kb are background pixels close to their corresponding pair. (b)-(i) From left to right, extracted subimages* $I_{VE}^{S_{P\text{-}k\delta}^{17}}$ *with* $k = 1, 2, 3, 4$ *and* $\delta =$*a,b. (j) Image representation of the Gaussian matrix* $G_{0,1.7^2}^{17}$*. (k)-(r) Subimages* I_{Hu} *result of multiplying the original ones in (b)-(i) by the Gaussian matrix.*

To overcome this problem, moments are computed on new subimages I_{Hu} produced by multiplying the original ones, $I_{VE}^{S_{x,y}^{17}}$, by an equal-dimension matrix (17×17) of Gaussian values, whose mean is 0 and variance is 1.7^2, $G_{0,1.7^2}^{17}$. That is, for every point of coordinates (i, j):

$$I_{Hu}(i,j) = I_{VE}^{S_{x,y}^{17}}(i,j) \times G_{0,1.7^2}^{17}(i,j) \tag{4.22}$$

	P-1a	P-1b	P-2a	P-2b	P-3a	P-3b	P-4a	P-4b
$\lvert\log(\phi_1)\rvert$	5.26	4.73	4.87	5.02	4.36	4.23	3.96	3.92
$\lvert\log(\phi_2)\rvert$	11.70	11.29	10.71	11.81	10.92	10.90	10.59	12.11

Table 4.1: *Module of the ϕ_1 and ϕ_2 moments logarithm calculated from the subimages $I_{VE}^{S_{x,y}^{17}}$ shown in Figure 4.5 images (b)-(i).*

	P-1a	P-1b	P-2a	P-2b	P-3a	P-3b	P-4a	P-4b
$\lvert\log(\phi_1)\rvert$	5.34	2.89	5.16	3.13	4.79	2.34	4.12	2.21
$\lvert\log(\phi_2)\rvert$	13.57	9.16	12.85	10.11	11.19	8.31	10.82	7.79

Table 4.2: *Module of the ϕ_1 and ϕ_2 moments logarithm calculated from the subimages I_{Hu} shown in Figure 4.5 images (j)-(q).*

With this choice of parameters, the 9×9 central values in $G_{0,1.7^2}^{17}$ contain the 97% of the area of the represented Gaussian distribution, the remainder values being close to 0 (supposing that the central pixel of $I_{VE}^{S_{x,y}^{17}}$ is located on the middle of a "wide" vessel, these 9×9 central values in $G_{0,1.7^2}^{17}$ correspond to vessel pixels in $I_{VE}^{S_{x,y}^{17}}$). This operation is illustrated in Figure 4.5 (images (b)-(i), (j) and (k)-(r)). The effect of this multiplication is clearly observed in these new subimages and their associated moment values (Table 4.2). These values become sensitive for describing vessel and non-vessel central pixels, as they now reflect significant differences between them. Both ϕ_1 and ϕ_2 values, in comparison with their original ones, increase if they describe vessel pixels and decrease otherwise. In conclusion, the following descriptors were considered to be part of the feature vector of a pixel located at (x, y):

$$f_6(x,y) = \lvert\log(\phi_1)\rvert \tag{4.23}$$

$$f_7(x,y) = \lvert\log(\phi_2)\rvert \tag{4.24}$$

where ϕ_1 and ϕ_2 are the moment invariants given by equations (4.15) and (4.16) computed on the subimages I_{Hu}, generated according to equation (4.22).

4.3.3 Classification

In the feature extraction stage, each pixel from a fundus image is characterized by a vector in a 7-dimensional feature space:

$$F(x,y) = (f_1(x,y), \ldots, f_7(x,y)) \tag{4.25}$$

Now, a classification procedure assigns one of the classes $C1$ (vessel) or $C2$ (non-vessel) to each candidate pixel when its representation is known. In order to select a suitable classifier, the distribution of the training set data (described below) in the feature space was analysed. The results of this analysis showed that the class linear separability grade was not high enough for the accuracy level required for vasculature segmentation in retinal images. Therefore, the use of a non-linear classifier was necessary. The following non-linear classifiers can be found in the existing literature on this topic: the kNN method (Sinthanayothin et al., 1999; Staal et al., 2004), support vector machines (Ricci and Perfetti, 2007), Bayesian classifier (Niemeijer et al., 2004; Soares et al., 2006), or neural networks (NN) (Gardner et al., 1996; Jiang and Mojon, 2003). A multilayer feedforward NN was selected in this work.

Two classification stages can be distinguished: a design stage, in which the NN configuration is decided and the NN is trained, and an application stage, in which the trained NN is used to classify each pixel as vessel or non-vessel to obtain a vessel binary image.

Neural network design

- Neural-network architecture: NN architectures have been discussed in detail in the literature (see Hagan et al., 1996 and Haykin, 1999, as examples). A multilayer feedforward network consisting of an input layer, three hidden layers and an output layer, was used in this work. Since all input data to the network are propagated through the layers from the input to the hidden layers and, finally, to the output layer without any feedback, a major advantage of the feedforward architecture is that it requires relatively low computing time during training. This is especially valuable when numerous simulations are necessary to evaluate a neuronal network performance.

  The input layer is composed of a number of neurons equal to the dimension of the feature vector (7 neurons). Regarding the hidden layers, several topologies with different numbers of neurons were tested. A number of 3 hidden layers, each containing 15 neurons, provided optimal NN configuration. The output layer contains a single neuron and is attached, as the remainder units, to a non-linear logistic sigmoid activation function. By this way, its output values fall within the real-valued interval $[0, 1]$. This choice was grounded on the fact of interpreting NN output as posterior probabilities.

Thus, the output of the ith neuron, Oi, is characterized by

$$O_i = \Psi(x_i) \tag{4.26}$$

where Ψ is the neuron's sigmoid function given by

$$\Psi(x_i) = \frac{1}{1 + exp\left[-\left(x_i + \theta_i\right)/\theta_0\right]} \tag{4.27}$$

and x_i is the sum of the weighted inputs of neuron i:

$$x_i = \sum_{j=1}^{N_i} w_{ij} \times O_j \tag{4.28}$$

In equation (4.27), θ_0 controls the non-linearity of the neuron's sigmoidal transformation and θ_i is the internal threshold for node i. In equation (4.28), O_j denotes the output from jth neuron of the previous layer, w_{ij} is the weight from neuron j to neuron i, and N_i is the number of inputs to neuron i.

- Neural-network training: NN must be trained to map certain input patterns to output profiles. A network becomes trained by adjusting the weights of the connections once these patterns are established through error interpretation. There are a variety of training algorithms that can be used to adjust the weights. The back-propagation algorithm (Rumelhart et al., 1987) was employed in this work. Basically, this algorithm is an iterative learning rule which updates layer-to-later weights based on the deviation of the actual neural network outputs from their desired response or target outputs. Thus, after one complete cycle or epoch of training pattern presentations, the calculated error in the outputs is backpropagated through the network and the weights in the connecting links are updated so that the output error is minimized.

  The training set, S_T, is composed of a set of N candidates for which the feature vector (F - eq. (4.25)), and the classification result (C_1 or C_2: vessel or non-vessel) are known:

$$S_T = \left\{ \left(F^{(n)}, C_k^{(n)} \right) | n = 1, \ldots, N; k \in \{1, 2\} \right\} \tag{4.29}$$

  The samples forming S_T were collected from manually labeled non-vessel and vessel pixels in the DRIVE training images. Specifically, around 30000 pixel samples, fairly

divided into vessel and non-vessel pixels, were used (as a reference, this number represents 0.65% of the total number of the DRIVE test image pixels that will be classified later on). Unlike other authors (Ricci and Perfetti, 2007; Soares et al., 2006), who selected their training set by random pixel-sample extraction from available manual segmentations of DRIVE and STARE images, the used training set was produced by hand [1]. As discussed in literature, gold-standard images may contain errors (see Bioux et al., 2007, for a comprehensive discussion on this issue) due to the considerable difficulty involved by the creation of these hand-made images. To reduce the risk of introducing errors in S_T and, therefore, of introducing noise in the NN, careful selection of specific training samples covering all possible vessel, background and noise patterns was carried out. Moreover, it should be pointed out that the network trained with the just defined S_T, in spite of taking information from DRIVE images only, was applied to compute method performance with both DRIVE and STARE databases.

Since the features f_i of F have very different ranges and values, each of these features is normalized to zero mean and unit variance independently by applying

$$\bar{f}_i = \frac{f_i - \mu_i}{\sigma_i} \tag{4.30}$$

where μ_i and σ_i stand for the average and standard deviation of the ith feature calculated over S_T.

Neural network application

At this stage, the trained NN is applied to an "unseen" fundus image to generate a binary image in which blood vessels are identified from retinal background: pixels' mathematical descriptions are individually passed through the NN. In the present case, the NN input units receive the set of features provided by equations (4.5)-(4.9) and (4.23)-(4.24), normalized according to eq. (4.30). Since a logistic sigmoidal activation function was selected for the single neuron of the output layer, the NN decision determines a classification value between 0 and 1. Thus, a vessel probability map indicating the probability for the pixel to be part of a vessel is produced. Illustratively, the resultant probability map corresponding to a DRIVE database fundus image (Figure 4.6, image (a)) is shown as an image in Figure 4.6,

[1] The training set is available online at http://www.uhu.es/retinopathy/eng/bd.php

image (b). The bright pixels in this image indicate higher probability of being vessel pixel. In order to obtain a vessel binary segmentation, a thresholding scheme on the probability map is used to decide whether a particular pixel is part of a vessel or not. Therefore, the classification procedure assigns one of the classes C_1 (vessel) or C_2 (non-vessel) to each candidate pixel, depending on if its associated probability is greater than a threshold Th. Thus, a classification output image I_{CO} (Figure 4.6, image (c)), is obtained by associating classes C_1 and C_2 to the gray level values 255 and 0, respectively. Mathematically:

$$I_{CO}(x,y) = \begin{cases} 255\ (\equiv C1) & \text{if } p(C_1|F(x,y)) \geq Th, \\ 0\ (\equiv C2) & \text{otherwise.} \end{cases} \tag{4.31}$$

where $p(C_1|F(x,y))$ denotes the probability of a pixel (x,y) described by feature vector $F(x,y)$ to belong to class C_1. The optimal Th value is discussed in section 4.4.2.

4.3.4 Postprocessing

Classifier performance is enhanced by the inclusion of a 2-step postprocessing stage: the first step is aimed at filling pixel gaps in detected blood vessels, while the second step is aimed at removing falsely detected isolated vessel pixels.

From visual inspection of the NN output, vessels may have gaps (i.e., pixels completely surrounded by vessel points, but not labeled as vessel pixels). To overcome this problem, an iterative filling operation is performed by considering that pixels with at least six neighbors classified as vessel points must also be vessel pixels. Besides, small isolated regions misclassified as blood vessel pixels are also observed. In order to remove these artifacts, the pixel area in each connected region is measured. In artifact removal, each region connected to an area below 25 is reclassified as non-vessel. An example of the final vessel segmented image after this further processing stage is shown in Figure 4.6 (image (d)).

4.4 Experimental results

This section presents the results obtained with the presented vascular tree segmentation methodology. In the first subsection, the metrics used for performance evaluation and comparison are defined. The second subsection presents the obtained results and also offers a

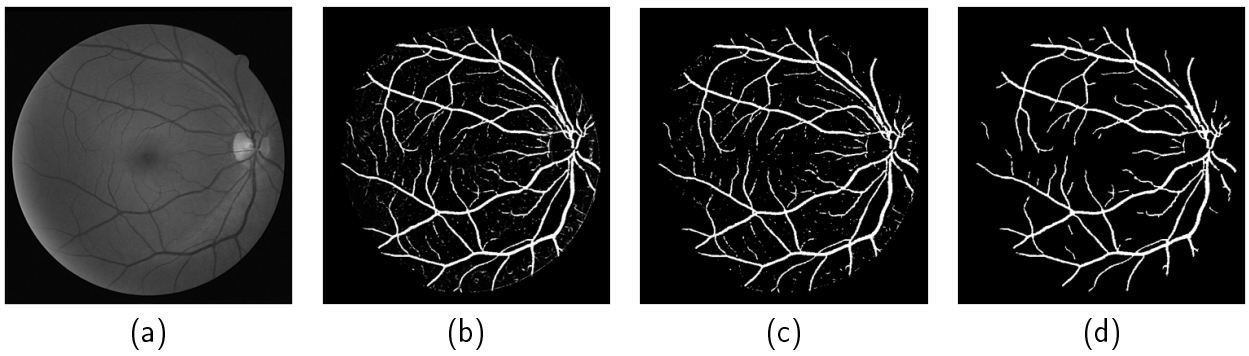


Figure 4.6: *(a) Green channel of the original image. (b) Obtained probability map represented as an image. (c) Thresholded image. (d) Postprocessed image.*

study on spatial distribution of the classification errors produced by the segmentation algorithm. The third subsection compares the obtained results with those presented in the most relevant published works.

4.4.1 Performance measures

In order to quantify the algorithmic performance of the proposed method on a fundus image, the resulting segmentation is compared to its corresponding gold-standard image. This image is obtained by manual creation of a vessel mask in which all vessel pixels are set to one and all non-vessel pixels are set to zero. Thus, automated vessel segmentation performance can be assessed.

The algorithm was evaluated in terms of *sensitivity* (Se), *specificity* (Sp), *positive predictive value* (Ppv), *negative predictive value* (Npv) and *accuracy* (Acc) (see appendix B for a comprehensive description of these metrics). Table 4.3 shows the contingency table for vessel detection. Taking into account this table, these metrics express:

- Se: the ratio of well-classified vessel pixels.
- Sp: the ratio of well-classified vessel non-vessel pixels.
- Ppv: the ratio of pixels classified as vessel pixel that are correctly classified.
- Npv: the ratio of pixels classified as non-vessel pixel that are correctly classified.
- Acc: the ratio of total well-classified as vessel or non-vessel pixel.

	Vessel present	Vessel absent
Vessel detected	True Positive (TP)	False Positive (FP)
Vessel not detected	False Negative (FN)	True Negative (TN)

Table 4.3: *Contingency table for vessel classification.*

In addition, algorithm performance was also evaluated with receiver operating characteristic (ROC) curve (see appendix B). The area under ROC curve (AUC) was used as a single measure to quantify this performance.

4.4.2 Proposed method evaluation

The method was evaluated on the DRIVE and STARE databases described in appendix A. These databases have been widely used by other researchers to test their vessel segmentation methodologies since, apart from being public, they provide manual segmentations for performance evaluation. Since the images' dark background outside the FOV is easily detected, Se, Sp, Ppv, Npv and Acc values were computed for each image considering FOV pixels only. Since FOV masks are not provided for STARE images, they were generated with an approximate diameter of 650×550. The results are listed in Table 4.4 and Table 4.5. The last row of the tables shows average Se, Sp, Ppv, Npv and Acc values (denoted as $\overline{Acc}$ for future references), for the 20 images in each database.

The performance results shown in Table 4.4 and 4.5 were obtained considering the same Th threshold value for all the images in the same database (0.63 and 0.91 for DRIVE and STARE images, respectively). These values were set to provide *maximum average accuracy* (MAA) in each database in the following way. For a given Th value, one Acc value is obtained for each of the 20 images selected for testing on a given database. These 20 Acc values are then averaged to obtain a single performance measure, $\overline{Acc}$, linked to the selected Th value. Several $\overline{Acc}$ values are obtained at certain thresholds applying these operations for different Th values. The final Th threshold value selected for a given database is that providing the maximum $\overline{Acc}$ value, MAA. Figure 4.7, image (a), shows the $\overline{Acc}$ values calculated for Th values from 0 to 1 (step of 0.02). The results for both DRIVE and STARE databases are shown. The MAA values and their corresponding Th values are marked for every database in this figure. It is worth mentioning that $\overline{Acc}$ variation shows no significant dependence on Th. As it can be observed in image(a) of Figure 4.7, although different optimum Th values are reached depending on the database on which performance is computed, a wide

Image	*Se*	*Sp*	*Ppv*	*Npv*	*Acc*
1	0.7778	0.9734	0.8149	0.9667	0.9477
2	0.7665	0.9793	0.8671	0.9597	0.9474
3	0.7225	0.9711	0.8099	0.9535	0.9349
4	0.7032	0.9855	0.8818	0.9557	0.9479
5	0.6815	0.9872	0.8931	0.9518	0.9457
6	0.6296	0.9893	0.9060	0.9420	0.9385
7	0.6780	0.9861	0.8817	0.9525	0.9453
8	0.5704	**0.9917**	**0.9083**	0.9414	0.9388
9	0.6530	0.9882	0.8809	0.9554	0.9489
10	0.6967	0.9846	0.8599	0.9599	0.9502
11	0.7164	0.9745	0.8068	0.9585	0.9410
12	0.7399	0.9785	0.8312	0.9634	0.9486
13	0.6522	0.9866	0.8896	0.9450	0.9392
14	0.7773	0.9702	0.7769	0.9702	0.9474
15	**0.8139**	0.9444	0.6290	**0.9777**	0.9308
16	0.7327	0.9816	0.8573	0.9606	0.9491
17	0.6629	0.9851	0.8623	0.9541	0.9454
18	0.7106	0.9781	0.8078	0.9630	0.9473
19	0.8058	0.9813	0.8552	0.9736	**0.9602**
20	0.6436	0.9861	0.8463	0.9586	0.9495
Average	0.7067	0.9801	0.8433	0.9582	0.9452

Table 4.4: *Performance results on DRIVE database images.*

Image	*Se*	*Sp*	*Ppv*	*Npv*	*Acc*
1	0.5997	0.9844	0.8245	0.9527	0.9425
2	0.5074	0.9931	0.8798	0.9527	0.9489
3	0.6534	0.9892	0.8422	0.9699	0.9619
4	0.4159	0.9953	0.9094	0.9379	0.9365
5	0.5884	0.9857	0.8508	0.9452	0.9372
6	0.7958	0.9712	0.7250	0.9803	0.9559
7	0.8183	0.9771	0.8142	0.9772	0.9593
8	0.8682	0.9673	0.7512	0.9848	0.9572
9	0.7729	0.9818	0.8355	0.9731	0.9595
10	0.6670	0.9809	0.8110	0.9601	0.9467
11	0.8109	0.9724	0.7601	0.9795	0.9567
12	**0.8781**	0.9675	0.7603	**0.9854**	0.9581
13	0.7796	0.9770	0.8246	0.9697	0.9530
14	0.7765	0.9787	0.8374	0.9687	0.9537
15	0.6910	0.9850	0.8598	0.9599	0.9504
16	0.6802	0.9825	0.8639	0.9497	0.9402
17	0.7039	0.9882	0.8926	0.9599	0.9534
18	0.5840	**0.9961**	**0.9172**	0.9698	0.9675
19	0.6776	0.9872	0.7694	0.9799	**0.9689**
20	0.6225	0.9763	0.7245	0.9628	0.9441
Average	0.6944	0.9819	0.8227	0.9659	0.9526

Table 4.5: *Performance results on STARE database images.*

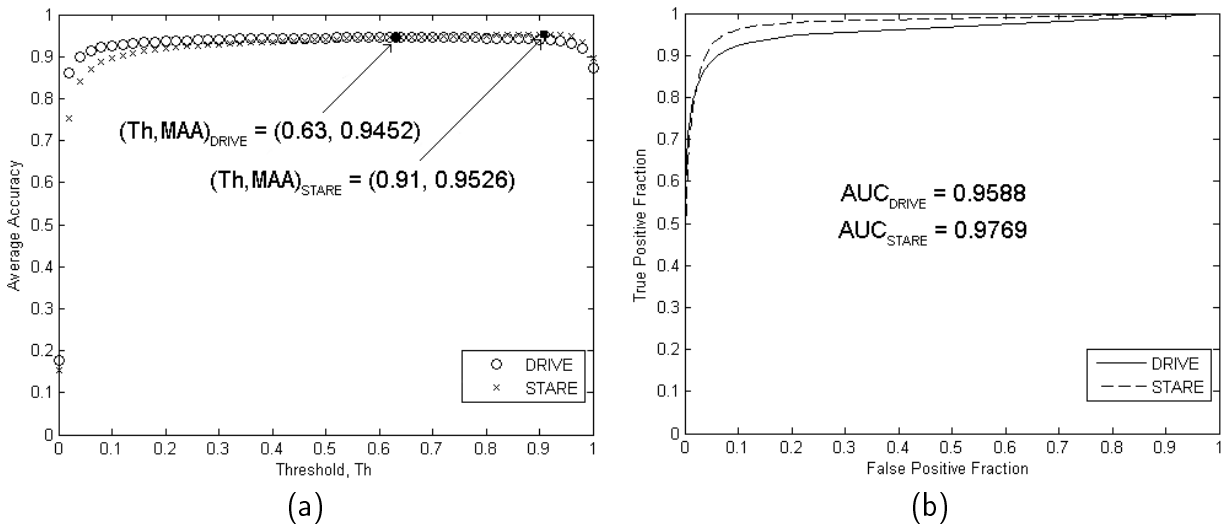


Figure 4.7: *(a) Acc of the segmentation algorithm as a function of the threshold parameter Th for the DRIVE and STARE databases. (b) ROC curves for the DRIVE and STARE databases. Measured AUC values are given.*

range of Th values provides $\overline{Acc}$ values very close to MAA. Therefore, Th values can be concluded not to be a critical method to assess performance in terms of Acc, since it slowly varies according to it. This influence of Th on system performance is also visible in the ROC curves for the two databases shown in Figure 4.7, image (b). These curves were produced by calculating the true and the false positive fraction on all test images through Th-threshold variations. The AUC measured for both curves was 0.9588 and 0.9769 for the DRIVE and STARE databases, respectively.

On the other hand, the spatial distribution of the classification errors produced by the segmentation algorithm, FN and FP, was studied. The following four situations were considered: FN produced in thin and non-thin vessel pixels, and FP produced in pixels near to and far from vessel borders. For that, thin and non-thin vessels were separated in each gold standard image (as an example, see Figure 4.8, image (b)). To perform this separation, two different approaches were firstly considered. One approach was to considerate as non-thin vessels those born in the optic disc and their direct ramifications, while the remaining were considered as thin vessels. This approach is based on a medical criterion which establishes that a retinography has enough quality to be diagnosed if the thin vessels, defined by this way, are visible. The second considered approach is based on size, and establishes that a vessel is thin if its width is lower than 50% of the width of the widest optic disc vessel. Otherwise the vessel is considered non-thin. Previous tests revealed that both approaches

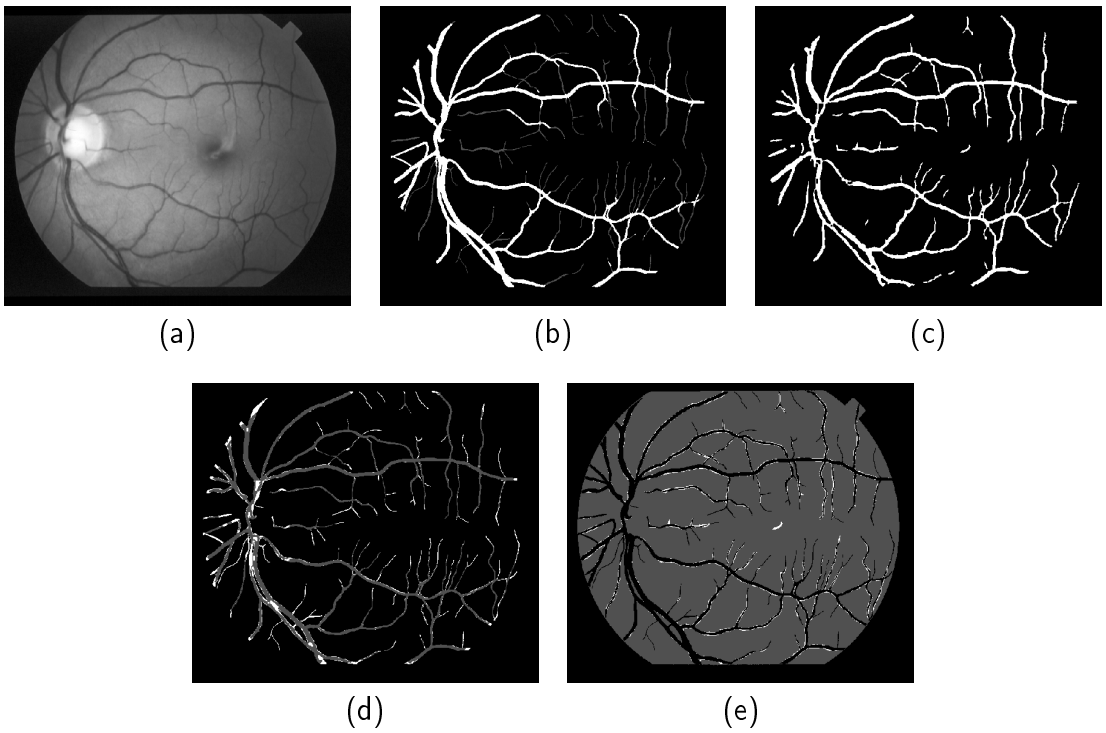


Figure 4.8: *Illustration of the spatial location of classification errors on a segmentation of a STARE image. (a) Green channel of the original image. (b) Thin and non-thin blood vessels extracted from the manual segmentation in white and dark-gray colors, respectively. (c) Segmentation of (a) generated by the presented algorithm. (d) FN and TP obtained by the proposed algorithm represented in white and dark-gray colors, respectively. (e) FP and TN obtained by the proposed algorithm represented in white and dark-gray colors, respectively.*

render similar results. The size approach was finally selected, since it allows performing automatic separation in absence of subjectivity. On the other hand, a FP is considered to be far from a vessel border if the distance from its nearest vessel border pixel in the gold-standard is over 2 pixels. Otherwise, the FP is considered to be near. Table 4.6 summarizes the results of this study. This table shows the average ratio of FN and FP provided by the segmentation algorithm for the 20 test images in the DRIVE and STARE databases. The average percent of FN and FP corresponding to the different spatial locations considered are also shown. For both databases, the percent of FN produced in non-thin vessel pixels was higher than that in thin vessel pixels. However, taking into account that thin vessels are composed by a considerably lower number of pixels than non-thin ones, the value obtained for thin vessels bears a more negative impact. This can be observed in Figure 4.6, image (d), and Figure 4.8, image (d). While FN in non-thin vessels involve no degradation in the segmented structure, FN in thin vessels produce disconnections in some of them. Regarding

	FN		FP	
Database	FN average rate		FP average rate	
DRIVE	0.2933		0.3056	
STARE	0.0199		0.0181	
	Thin Vessels	*Non-thin Vessels*	*Near to Vessels*	*Far from Vessels*
DRIVE	46.08%	53.92%	86.25%	13.75%
STARE	34.60%	65.40%	83.55%	16.45%

Table 4.6: *Study of FN and FP spatial location on the DRIVE and STARE databases.*

FP distribution, FPs tend to be near vessel borders. As it can be checked in Figure 4.8, image (e), this means that most FPs produced by the segmentation algorithm tend to slightly enlarge the vessels and not to introduce meaningless isolated noise.

4.4.3 Comparison to other methods

In order to compare the described approach to other retinal vessel segmentation algorithms, $\overline{Acc}$ and AUC were used as measures of method performance. Since these measurements were performed by other authors, this choice facilities comparing the results. Tables 4.7 and 4.8 show performance comparison results in terms of $\overline{Acc}$ and AUC respectively between the presented method (referenced in tables as *Gray-Hu Method*) and the following published methods: Chaudhuri et al. (1989b), Cinsdikici and Aydın (2009), Hoover et al. (2000), Jiang and Mojon (2003), Martinez-Perez et al. (2007), Mendonça and Campilho (2006), Niemeijer et al. (2004), Ricci and Perfetti (2007), Soares et al. (2006), and Staal et al. (2004). All these supervised or rule-based methods have been briefly commented in section 4.2. The values shown in both tables are presented for each database as reported by their authors. If they are not available for a specific database or were not calculated for the 20 images selected for testing, they were not included in the tables, thus appearing as gaps. The values in the last column of each table indicate the overall $\overline{Acc}$ and AUC when both databases are taken into account.

An overview of the segmentation results on DRIVE images shows the proposed method reaches better performance than most of the other methods, being comparable to or approximating the performance of other detection techniques. The $\overline{Acc}$ value achieved with the new approach is outperformed only by Soares et al. (2006), Mendonça and Campilho (2006) and Ricci and Perfetti (2007). Regarding the approaches by Soares et al. (2006) and Mendonça and Campilho (2006), it is important to point out that the presented method clearly

Method Type	Method	DRIVE	STARE	DRIVE + STARE
Supervised	Staal et al. (2004)	0.9441	-	-
	Niemeijer et al. (2004)	0.9417	-	-
	Soares et al. (2006)	0.9466	0.9480	0.9473
	Ricci and Perfetti (2007)	**0.9595**	**0.9646**	**0.9621**
	Gray-Hu Method (this work)	0.9452	0.9526	0.9489
Ruled-Based	Chaudhuri et al. (1989b)	0.8773	-	-
	Hoover et al. (2000)	-	0.9275	-
	Jiang and Mojon (2003)	0.8911	0.9009	0.8960
	Mendonça and Campilho (2006)	0.9463	0.9479	0.9471
	Martinez-Perez et al. (2007)	0.9344	0.9410	-
	Cinsdikici and Aydın (2009)	0.9293	0.9377	-

Table 4.7: *Performance results compared to other methods on the STARE and DRIVE databases in terms of average accuracy.*

Method Type	Method	DRIVE	STARE	DRIVE + STARE
Supervised	Staal et al. (2004)	0.9520	-	-
	Niemeijer et al. (2004)	0.9294	-	-
	Soares et al. (2006)	0.9614	0.9671	0.9642
	Ricci and Perfetti (2007)	**0.9633**	0.9680	0.9656
	Gray-Hu Method (this work)	0.9588	**0.9769**	**0.9678**
Ruled-Based	Chaudhuri et al. (1989b)	0.7878	-	-
	Hoover et al. (2000)	-	0.7590	-
	Jiang and Mojon (2003)	0.9327	0.9298	0.9312
	Mendonça and Campilho (2006)	-	-	-
	Martinez-Perez et al. (2007)	-	-	-
	Cinsdikici and Aydın (2009)	0.9407	-	-

Table 4.8: *Performance results compared to other methods on the STARE and DRIVE databases in terms of area under ROC curve.*

outperforms the $\overline{Acc}$ these authors reported on STARE images. Therefore, it renders better overall $\overline{Acc}$ for both databases than theirs. The same conclusions are drawn when these methods are compared in terms of AUC. On DRIVE database images, the AUC value provided by the new proposal is only lower than those reported by Soares et al. (2006) and Ricci and Perfetti (2007) (Mendonça and Campilho (2006) did not report AUC values). However, due to the excellent AUC result on the STARE database, the presented approach reaches the highest average AUC when both databases are considered.

The proposed method proves especially useful for vessel detection in STARE images. Its application to this database resulted in the second highest accuracy score among all experiments (only behind the approach by Ricci and Perfetti, 2007) and the first when AUC is the reference measurement. This result gains more importance by the fact that the used classifier

Method	DRIVE (training on STARE)	STARE (training on DRIVE)
Ricci and Perfetti (2007)	0.9266	0.9452
Gray-Level&Hu-MomentFeatures (this work)	**0.9448**	**0.9526**

Table 4.9: *Performance results compared to Ricci and Perfetti's method with cross trainning in terms of average accuracy.*

was trained only on DRIVE images, unlike the other supervised approaches presented in Table 4.7 and 4.8. For instance, since there are no available labeled training images for STARE images, Soares et al. (2006) performed leave-one-out tests on this database (i.e., every image is classified by using samples from the other 19 images), while Ricci and Perfetti (2007) built its classifier by using a training set comprising samples randomly extracted from test images. With the purpose of using one and the same trained classifier for testing the method on the 20 STARE images and including no sample belonging to the test set in the training, in this work the training set was formed by collecting pixels from DRIVE training images. Thus, the method's suitability for being applied on any fundus image can be checked in a more realistic way. It should be also mentioned that these good results with respect to other existing approaches were obtained on images containing pathological artifacts. The STARE database contains ten images with pathologies, while the test of DRIVE only contains four. Moreover, abnormal regions are wider in STARE.

Regarding performance comparison in terms of $\overline{Acc}$ when results are jointly analysed for DRIVE and STARE images (Table 4.7, last column), the new proposal renders greater accuracy than others authors' algorithms, being outperformed only by Ricci and Perfetti's proposal (Ricci and Perfetti, 2007). However, this method proved very dependent on the training set. Ricci and Perfetti (2007), to research the dependence of their classification method on the dataset, carried out an experiment based on, firstly, training the classifier on each of the DRIVE and STARE databases, and then, testing it on the other. Their maximum accuracy values are shown in Table 4.9. It can be observed that performance is worse now, since $\overline{Acc}$ strongly decreases from 0.9595 to 0.9266 on DRIVE and 0.9646 to 0.9452 on STARE database images. Therefore, as assumed by these authors, classifier retraining is necessary before applying their methodology on a new database. To verify dependence of the proposed method on the training set, the same experiment was completed. Thus, performance was computed on the DRIVE database, training the classifier with STARE images (as previously mentioned, the Acc values on STARE reported above was already obtained by training on

DRIVE images). The resulting $\overline{Acc}$ values are shown in Table 4.9 to facilitate comparisons between both methods under identical conditions. In this case, it is clearly observed that estimated performance of the new proposal in terms of method accuracy is higher, thus proving higher training set robustness.

4.5 Discussion and conclusion

Previous methods for blood vessel detection in retinal images can be classified into rule-based and supervised methods. This study proposes a method within the latter category. This method is based on a NN scheme for pixel classification, being the feature vector representing each pixel composed of gray-level and moment invariants-based features. To the best of our knowledge, although moment invariants have been widely used over the years as features for pattern recognition in many areas of image analysis (typical examples include optical character recognition and shape identification), they have never been applied within this framework.

The experiments aimed at evaluating the efficiency of the applied descriptors prove this method is capable of rendering accurate results, even when these types of features are used independently. Table 4.10 shows its accuracy performance when a feature vector composed of the five gray-level-based features, on one hand, and the two moment invariants-based features, on the other hand, is considered. According to these results, both sets of descriptors are suitable for pixel classification as vessel or non-vessel. Even when the feature vector is built as mentioned above, the new algorithm outperforms other segmentation proposals (e.g. Jiang and Mojon, 2003 or Martinez-Perez et al., 2007). However, these pixel representations are not optimal by themselves. Discriminative power increases when both sets of different feature types are jointly considered. Thus, accuracy improves up to 0.9452 and 0.9526 for the 20 test images in the DRIVE and STARE databases, respectively. Therefore, the method finally adopts a 7-dimensional feature vector composed by the five gray-level and the two moment invariants-based features.

The proposed method uses a NN for pixel classification as vessel or non-vessel. This classifier was selected after method accuracy assessment by means of a kNN and a SVM (used as in Staal et al., 2004, and Ricci and Perfetti, 2007, respectively), instead of a NN. Table 4.11 shows this performance comparison in terms of $\overline{Acc}$ measured on DRIVE and STARE test

Feature Vector Composition	DRIVE	STARE
G-L.	0.9407	0.9514
M-I.	0.9398	0.9410

Table 4.10: *Average accuracy on the DRIVE and STARE databases using one type of feature: Gray-Level (G-L) or Moment Invariants-Based Features (M-I).*

	Before postprocessing		After postprocessing	
Classifier	DRIVE	STARE	DRIVE	STARE
NN	0.9439	0.9517	0.9452	0.9526
kNN	0.9383	0.9500	0.9415	0.9524
SVM	0.9393	0.9498	0.9427	0.9519

Table 4.11: *Average accuracy on the DRIVE and STARE databases of the proposed method when NN, kNN and SVM classifiers are used.*

images. Results before and after the postprocessing stage application are presented. NN showed better accuracy than kNN and SVM for all cases.

Table 4.7 and 4.8 show a performance overview on previously-published vessel segmentation methods in terms of $\overline{Acc}$ and AUC, respectively, for test DRIVE and STARE images. Among them, those providing best performance results include the rule-based proposal by Mendonça and Campilho (2006) (AUC is not available for this method), and the supervised methods by Soares et al. (2006) and Ricci and Perfetti (2007). When the results on both databases are jointly analysed, the approach proposed in this chapter provides the highest average AUC. Regarding $\overline{Acc}$, the value reached in this work is slightly higher than that obtained by Mendonça and Campilho (2006) and Soares et al. (2006). Although the approach by Ricci and Perfetti (2007) seems to offer more accurate behaviour, the results in Table 4.9 indicate that their method shows strong performance loss when confronted to an "unseen" database. Those authors obtain excellent results when both training and testing are performed on the same database but accuracy significantly worsens when the method is trained and tested on a different database. A possible explanation to this fact is that authors avoid image preprocessing to preserve vessel structure at most. However, since the different images are not homogenized, the method proves very sensitive to the training set. This is an important disadvantage for practical application, since a blood vessel detection tool must work on retinal images from multiple origins and be used by different operators working with different equipment. On the contrary, the presented proposal proves robust regarding the dataset used: even when trained on the DRIVE database, its application to the STARE database results in the second highest accuracy score among all experiments, and

the first if the results of Ricci and Perfetti (2007) are recorded under the same conditions. Therefore, the training set robustness shown by the presented method allows its automated application to images taken under different conditions. To this respect, it should be pointed out that the method, in its different stages (preprocessing, feature extraction, classification, and postprocessing), uses a set of parameters fixed to provide the best accuracy on DRIVE test images. That is, parameter values were optimized to retinas of 540 pixels in diameter. It is important to point out that, since the features used for pixel representation are invariant to scale change, the method is also applicable for processing images with different resolution without retraining. To make the method applicable for processing images of any resolution, its implementation should include either resampling the image to fulfil the mentioned size condition, or adapting all parameter to the new retina size.

In addition, method simplicity should also be highlighted. Its pixel classification procedure is based on computing only seven features for each pixel, thus needing shorter computational time. The total time required to process a single image is less than approximately one minute and thirty seconds, running on a PC with an Intel Core2Duo CPU at 2.13 GHz and 2 GB of RAM. Since the method implementation is experimental, this performance might still be improved.

The demonstrated effectiveness and robustness, together with its simplicity and fast implementation, make this proposed automated blood vessel segmentation method a suitable tool for being integrated into a complete prescreening system for early DR detection.

Chapter 5

Macular Zone Identification

MACULAR zone identification is of vital importance in the development of systems for the automated diagnosis of diabetic retinopathy. Since the macula is responsible for central vision, the distance at which lesions are located from it influences their clinical relevance and, therefore, this distance is used to assess the risk of vision loss associated to the illness.

This chapter proposes the identification of the macular zone by taking advantage of its known approximate center position with respect to other fundus features and its visual appearance. The chapter contains five sections. The **first section** illustrates features of the macular zone in eye fundus color images and describes the role of its identification within the frame of the automated diagnosis of diabetic retinopathy.

The **second section** reviews the state of art of the automated location of the macular zone in eye fundus images. The most relevant works related to this topic are categorized and summarized.

The **third section** describes the methodology proposed to identify the macular zone in color retinal images. This methodology, firstly, estimates the position of the macula center, i.e. the fovea, by exploiting its a priori-known approximate location with respect to the optic disc and the vascular tree. To this effect, the automated optic disc and vascular tree segmentation methodologies presented in chapters 3 and 4 are used. Then, the estimated fovea location is refined if this is detectable using morphological processing. Finally, the macular zone is identified using a polar coordinate system centered on the obtained fovea center location.

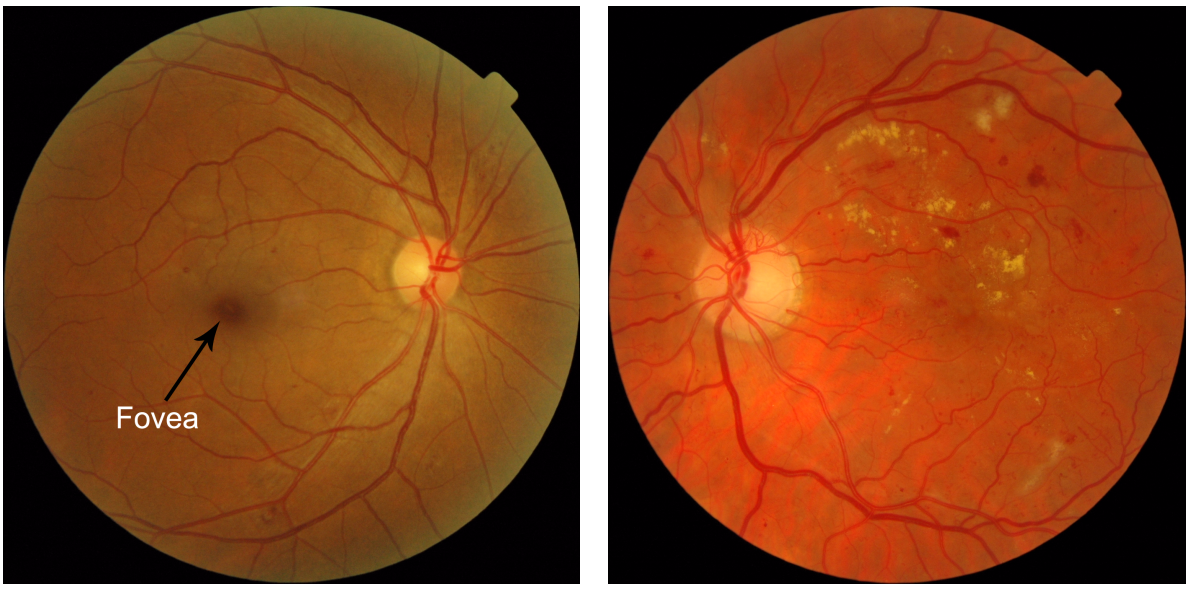


Figure 5.1: *Fovea appearance in eye fundus images: (a) In this image, the fovea is easily recognizable as a round dark area. (b) This image shows a fovea not visible due to the presence of lesions.*

In the **fourth section**, the results obtained with the presented methodology are shown. These results were evaluated using a criterion based on the distance between the estimated and the true fovea center. Since the border of the fovea region is not well defined, this criterion has been widely used and accepted in the literature. The obtained results are also compared to results of other works present in the literature in this section.

Finally, the **fifth section** ends the chapter with a discussion on the presented work and a set of key conclusions.

5.1 Introduction

The whole macula is not generally distinguishable in retinal color images. It exhibits non-specific structure and varies greatly across individuals due to variations in the levels of pigment associated with factors such as ethnicity, age, diet, and disease states. In spite of it, the macula center, the fovea, is often recognizable as a round region darker than its surrounding retinal tissue (Figure 5.1, image (a)). Nevertheless, this mark may not be visible due to, for example, the presence of lesions (Figure 5.1, image (b)).

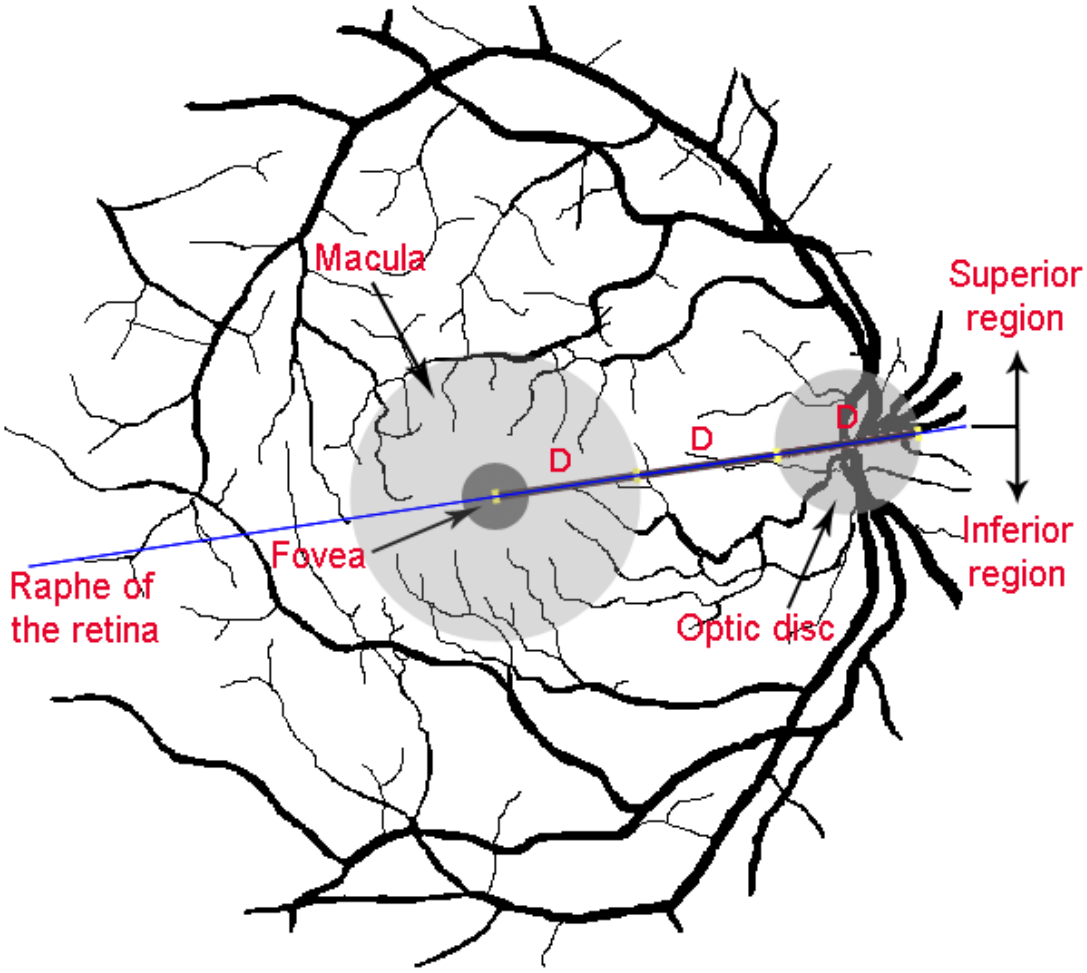


Figure 5.2: *Diagram illustrating macula features in the eye fundus.*

Figure 5.2 shows a diagram illustrating macula features in the eye fundus. Anatomically, the fovea center is located in average at 2.5 optic disc (OD) diameters from the OD center (Larsen, 1976). The fovea location follows the horizontal raphe of the retina, which is a line roughly passing through the OD and the fovea, or more generally, separating the superior and inferior regions of the retina determined by the superior and inferior vessels of the vascular tree. This raphe is rotated a few degrees with respect to an imaginary horizontal raster passing through the OD center. The fovea radius is between one third and one fourth of the macula radius, and the macula radius is approximately equal to one OD diameter (Schwiegerling, 2004).

The macula is responsible for central vision and, concretely, the fovea is in charge of sharp central vision. Because of this important function of the macula in vision, the distance at which lesions are located from this area influences their clinical relevance (this issue

was commented in section 2). That is why automated location and identification of the macular zone is unavoidable requirement in a system for the automated diagnosis of diabetic retinopathy.

5.2 State of art

The literature is sparse regarding automated methods for localizing the macula region in eye fundus images, likely due to its lack of stable visual appearance. Two types of criteria for macula identification can be clearly recognizable when reviewing the literature. These are *visual feature-based* and *anatomical feature-based* criteria. The former criterion includes techniques that attempt to find the macula by exploiting the visual appearance of the fovea. On the other hand, the latter criterion comprises techniques to obtain an estimate of the fovea center location making use of its known anatomical features regarding its position on the retinal surface. Methods using one of these criteria exclusively (Niemeijer et al., 2007; Tobin et al., 2007) or both (Fleming et al., 2007; Gagnon et al., 2001; Li and Chutatape, 2004; Niemeijer et al., 2009; Sinthanayothin et al., 1999) can be found in the literature.

Sinthanayothin et al. (1999) presented a method for fovea recognition in which this was assumed to be the darkest area of the fundus image, with approximately the same intensity as the blood vessels. Firstly, they correlated the fovea to a template of intensities, being the template chosen to approximate a typical fovea. Then, the location of the maximum response was selected as the location of the fovea if it was at approximately 2.5 times the diameter of the OD from the OD center. Gagnon et al. (2001) used a similar approach to detect the macula center. They first generated a coarse resolution image from the original image. Secondly, they selected the darkest pixel in the coarse resolution image following the cited distance criterion as the region occupied by the macula in the original image. Finally, the exact center of the macula was found by searching in the vicinity for this darkest pixel on the original fine resolution image. On the other hand, Li and Chutatape (2004) presented a method for fovea location that also used its darkness and the distance criterion, but they additionally proposed the use of the vascular arch to constrain the search area. To this effect, they presented a model based approach in which an active shape model was used to extract the main course of the vasculature based on the location of the OD. This course and the distance criterion of the fovea was used to decide a region of interest. Finally, the fovea

center was obtained by applying a thresholding scheme to the region of interest. Later, Tobin et al. (2007) proposed the automated location of the fovea using the distance criterion and the vascular arch exclusively. The method started by locating the OD and estimating the vascular arch by means of a parabolic model. Based on these two anatomical landmarks, the location of the fovea was ultimately inferred. A similar method for fovea recognition was presented by Fleming et al. (2007). The method started with the parabolic modeling of the vascular arch. Then, using this elliptical modeling of the major retinal blood vessels, the approximate locations of the OD and the fovea were obtained. Finally, the fovea position was refined based on its local darkening. On the other hand, Niemeijer et al. (2007) presented a method to model the distribution of all retinal features. The method used an optimization method to fit a point distribution model to the fundus image. After fitting, the points of the model indicated the location of the normal anatomy. These same authors presented later another work for the detection of the fovea (Niemeijer et al., 2009). They first found the OD and, based on its location, determined an area of interest for the fovea. Then, after blurring the image, the fovea location was decided as the pixel with the lowest value within the restricted area.

5.3 Methodology

This chapter describes a new methodology to identify the macular zone in fundus images. The exact location of the macula center position is not crucial for a system for the early automated detection of DR. What is especially important for such a system is the obtention of a good estimate of this position in a high percent of cases. This motivation along with the previously described lack of stable appearance of the fovea in eye fundus images justifies the philosophy of the presented methodology: to obtain an acceptable estimate of the macula center location for all cases, replacing it with an accurate location when is safe and possible. The methodology is divided into three main stages: (A) obtention of an estimate of the macula center location by using its known positional features with respect to the optic disc and the vascular tree, (B) obtention of an accurate macula center location by segmenting the fovea if this is detectable, and (C) final decision of the macula center location and identification of the macular zone.

5.3.1 Obtaining a fovea center location estimate

The primary approach for estimating the macula center is to determine the horizontal raphe of the retina, or in other words, the line delimiting the superior and inferior vessels. The fovea center is assumed to reside at a fixed distance along this line at 2.5 OD diameters from the OD center (Larsen, 1976). Therefore, to obtain the retinal raphe estimate and ultimately estimate the fovea location, the segmentations of the OD and vascular tree are required. To this effect, the methodologies presented in chapters 3 and 4 for the segmentation of these landmarks are used. D_{OD} and (x_{OD}, y_{OD}) will denote the diameter and the coordinates of the OD center, respectively, and $b_t(x, y)$ will stand for the segmentation of the vasculature structure skeletonized using an homotopic skeletonization method (Serra, 1982) (see Figure 5.3, images (a) and (b)).

To estimate the horizontal raphe of the retina, a parabolic model with origin in the OD center is applied to $b_t(x, y)$. Thus, the angle between the horizontal raster and the line of symmetry separating the superior and inferior retinal regions is identified (see illustration in Figure 5.3). To this effect, the model proposed by Tobin et al. (2007) is used.

Firstly, let's define the parabolic model of form

$$a \times y^2 = x \tag{5.1}$$

This model is applied to the statistical distribution of points defined by the segmented and morphologically skeletonized vasculature structure $b_t(x, y)$. These points form a set S defined as

$$S = \{(x, y) : b_t(x, y) \neq 0\} \tag{5.2}$$

So, points in S are those belonging to the skeleton of the segmented vascular tree and are used to estimate the required parabola. Next, to obtain a parabola with origin in the OD center, the parabolic expression (5.1) is shifted to this position to give

$$a \times (y - y_{OD})^2 = |x - x_{OD}| \tag{5.3}$$

As it was mentioned, the aim is to find the angle defining the horizontal raphe of the retina. For the dataset used for testing this methodology, the MESSIDOR database (see appendix A for database description), the raphe is distributed through a range of angles, β, from $+15°$

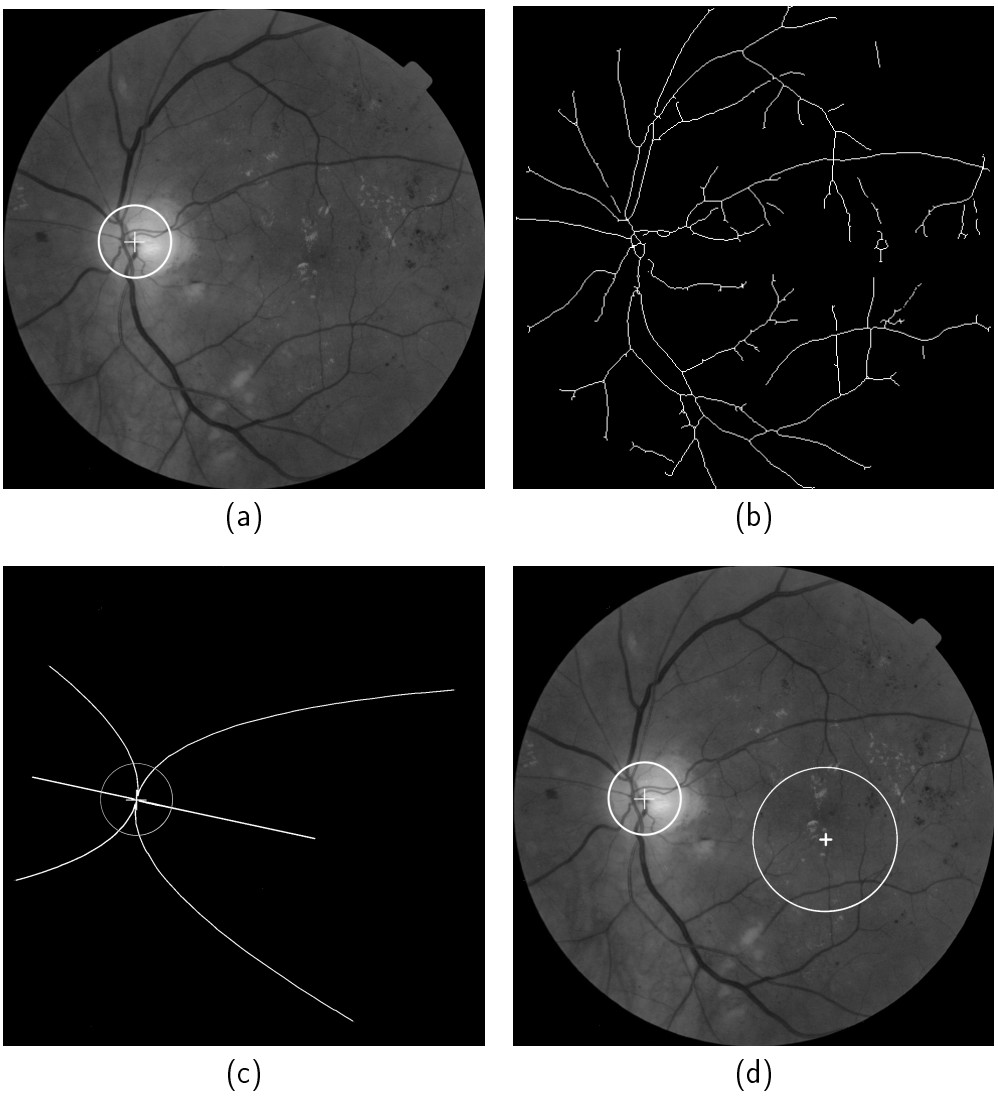


Figure 5.3: *(a) OD segmentation with center in* $(x_{OD}, y_{OD}) = (375, 665)$ *and diameter* $D_{OD} = 210$ *pixels. Skeletonized segmentation of the vascular tree* $b_t(x, y)$*. (c) Parabolic model centered on* (x_{OD}, y_{OD}) *and applied to* $b_t(x, y)$*. (d) Macula center estimate. The macula center is estimated by using the known OD center in (a), applying a geometric model to the skeletonized segmented vascular tree in (b) to determine the approximate horizontal raphe in (c) and positioning the macula center along the raphe accordingly in (d).*

to $-15\,^{\circ}$. Thus, this angle must be measured in each image to accurately place the estimate of the macula location. With this purpose, the parabolic model (5.3) is modified to include an axis rotation β (see Figure 5.4), by applying the coordinate transform:

$$x = x' \times cos\beta - y' \times sin\beta;\ y = x' \times sin\beta + y' \times cos\beta \tag{5.4}$$

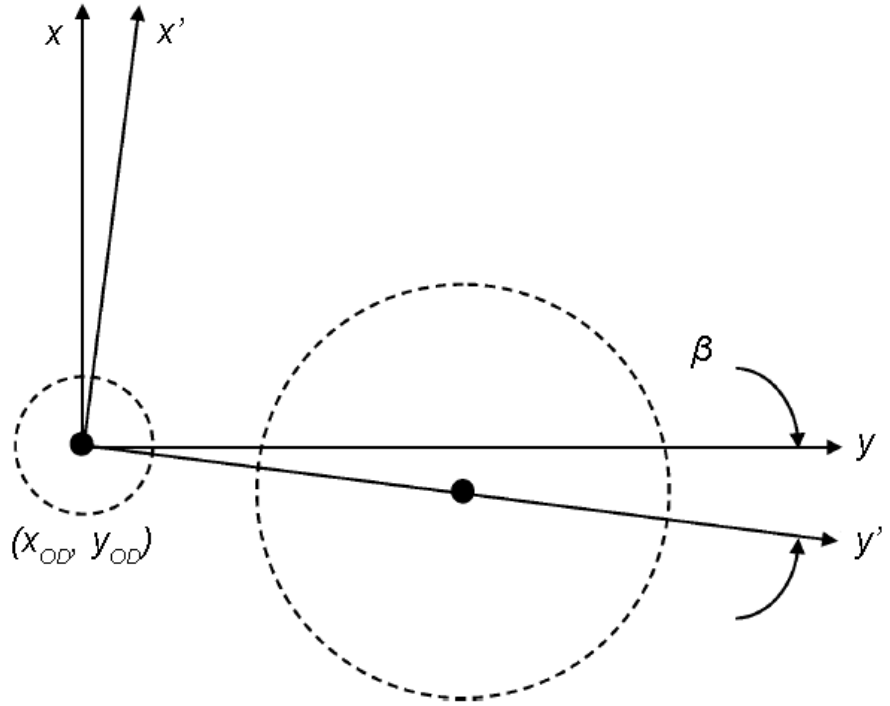


Figure 5.4: *Coordinate system used to describe the horizontal raphe of the retina.*

The modified parabolic model including this coordinate transform is therefore defined as

$$a \times [x' \times sin\beta + y' \times cos\beta - y_{OD}]^2 = |x' \times cos\beta - y' \times sin\beta - x_{OD}| \tag{5.5}$$

Finally, this can be expressed fully in terms of the transformed coordinate system by noting

$$x_{OD} = x'_{OD} \times cos\beta - y'_{OD} \times sin\beta;\ y = x'_{OD} \times sin\beta + y'_{OD} \times cos\beta \tag{5.6}$$

to give the following final expression:

$$a \times \left[\left(x' - x'_{OD}\right) \times sin\beta + \left(y' - y'_{OD}\right) \times cos\beta\right]^2 = \left|\left(x' - x'_{OD}\right) \times cos\beta - \left(y' - y'_{OD}\right) \times sin\beta\right| \tag{5.7}$$

where (x', y') are the coordinates of the axis rotated by an amount β, and aligning with the horizontal raphe. This equation is addressed by using the standard non-linear least squares method of Marquardt (Marquardt, 1963) to estimate the parameters (a, β) by iteratively minimizing the criterion function $J(a, \beta)$ on the set of points $S = \{(x', y') : b_t(x', y') \neq 0\}$

defined by

$$J(a,\beta) = \sum_{(x',y')\in S} a \times \left[\left(x' - x'_{OD}\right) \times sin\beta + \left(y' - y'_{OD}\right) \times cos\beta\right]^2 - \\ - \left|\left(x' - x'_{OD}\right) \times cos\beta - \left(y' - y'_{OD}\right) \times sin\beta\right| \tag{5.8}$$

The values used to initiate this function during iterative optimization are $a = 0.0032$ and $\beta = 0$ as Tobin et al. (2007) proposed.

Once the raphe of the retina is obtained, most of the works present in the literature estimate that the fovea center is at $2.5 * D_{OD}$ pixels from the OD center (x_{OD}, y_{OD}) following the raphe. This is true in general but, in our dataset of 1200 images were identified certain cases of very big or very small OD sizes in which this distance was not appropriated. Indeed, this is a consequence of, whereas $2.5 * D_{OD}$ pixels is an average distance, the OD size varies substantially from one tenth to one fifth of the retina (Li and Chutatape, 2001). To solve this difficulty, a criterion to detect and discard OD diameters, D_{OD}, in the extremes of the interval of OD size values was introduced. To define this criterion, firstly, the maximum OD diameter was measured in each of the manual OD segmentations created for the MESSIDOR database (this gold standard set was created to objectively evaluate the OD segmentation algorithm presented in chapter 3; this is described in section 3.4.3). Then, the ratios between these OD diameters and the diameters of their corresponding field-of-view masks (FOVs) were calculated. Thus, the distribution of the relative OD size measured in the whole set independently of image resolution was obtained. The distribution of these calculated ratios normalized to the interval $[0, 1]$ is presented in Figure 5.5. As it can be checked, the figure shows that normalized ratios follow a Gaussian distribution, whose average and standard deviation are 0.1382 and 0.0149, respectively (note that the obtained distribution corroborates that OD size is between one tenth and one fifth of the retina). Therefore, the criterion to discard an OD diameter, D_{OD}, is to consider this as extreme if its value is out of the average of the distribution plus/minus 2 times the standard deviation (i.e., out of the 95% of the distribution approximately). In this case, the discarded OD diameter value is replaced by the average value of the normal distribution.

Figure 5.6 shows some examples of the application of the described methodology for estimating the macula center. The first row shows results on images from patients affected by DR but with no risk of ME, the images in the second row are from patients affected by DR

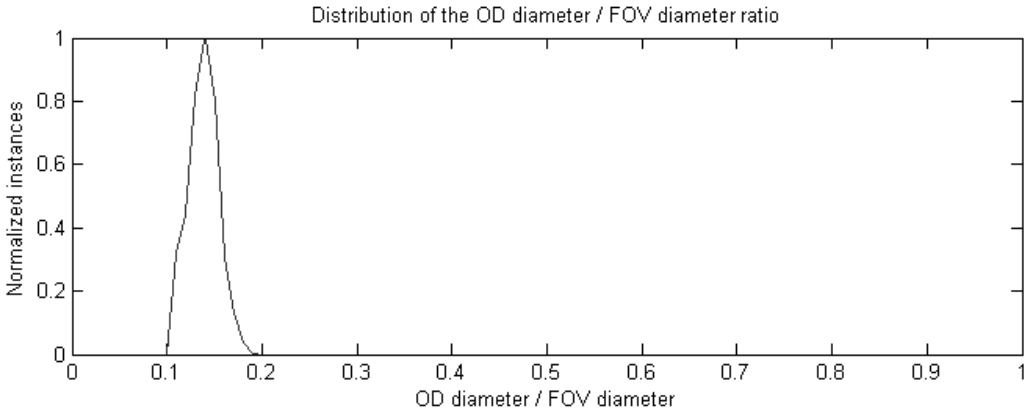


Figure 5.5: *Distribution of the relative OD size in eye fundus images. This relative OD size is measured as the ratio between the maximum OD diameter and the diameter of its corresponding FOV; data is shown normalized to the interval* $[0, 1]$. *As it can be observed, the relative OD size follows a Gaussian distribution, whose average and standard deviation are* 0.1382 *and* 0.0149*, respectively.*

with risk of ME and the last row shows results on images from healthy patients. Overlaid on the images is the parabolic fit to the vasculature with curvature a and raphe orientation β. The overlay on the images also includes the position of the located OD center, the segmented OD boundary and the placement of the estimated macula center along the raphe.

5.3.2 Obtaining an accurate fovea center location

The aim of this stage is, taking advandge of the fovea location estimate obtained in the previous step, to attempt to decide an accurate fovea center location. The main idea is, since the fovea is not always present, to detect if exists a dark object in the image with a certain depth and decide, attending to its shape and size, if it matches with a typical fovea. If exists this object and is validated, a new accurate fovea center location is decided as the centroid of this object. Otherwise, the fovea is considered as undetectable and the result of the present stage is null.

The processing described in this section works on an RGB subimage decided from the original retinography using the fovea location estimate obtained in the previous stage. Centered on this location, a $2 * D_{OD} \times 2 * D_{OD}$ fovea-containing subimage is extracted to be processed.

Although it is known that the green channel of an RGB retinography is the one with highest contrast (Walter and Klein, 2005), observing our imagery it was found out that when the fovea is visible in the color image it is also often present in the red channel. Therefore, the

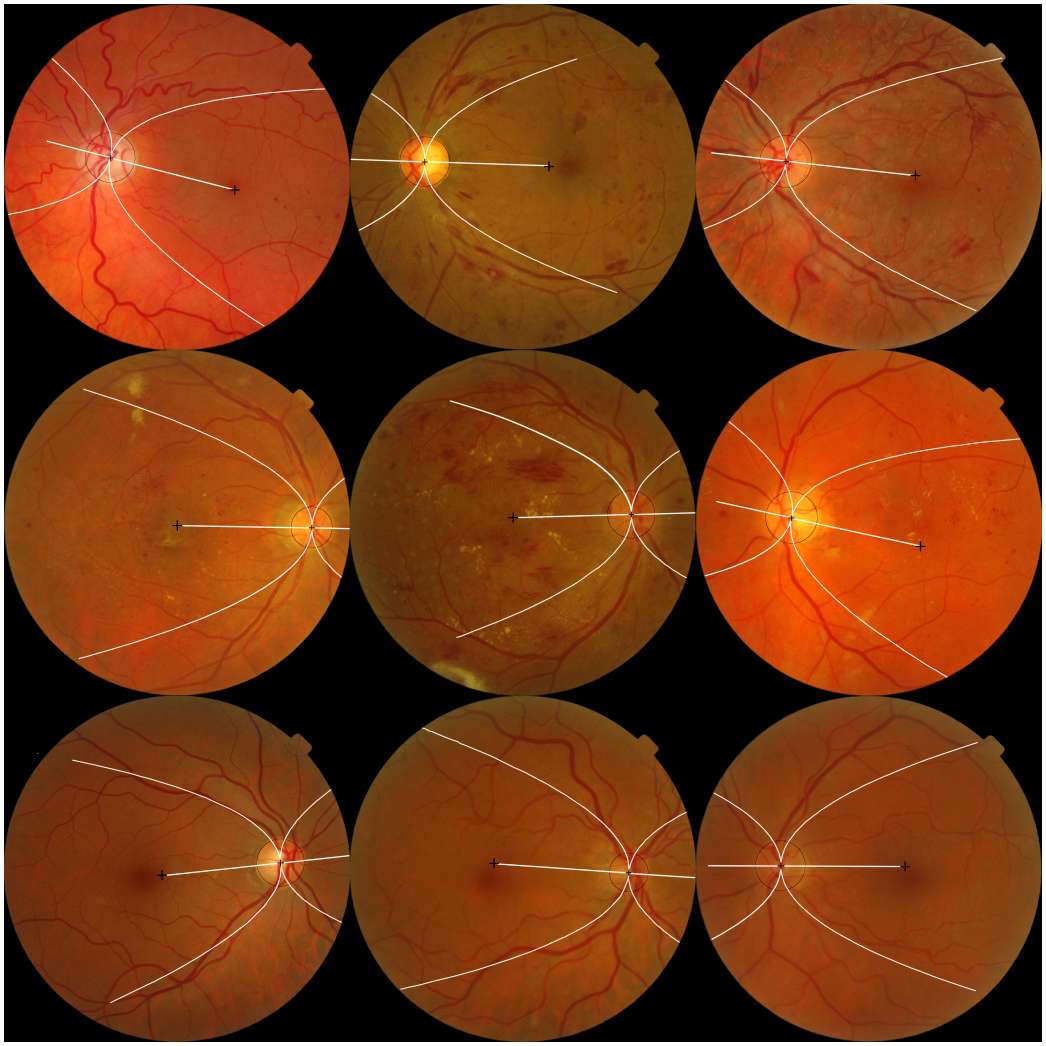


Figure 5.6: *Examples of the macula center estimate. The first row shows three retinographies from patients with DR and with no risk of ME, the second is composed of three retinographies from patients affected by DR and with risk of ME and the last three images are from healthy retinas.*

processing described in this section is performed in parallel on the red and green components and the "better" of the two results is ultimately selected (a similar approach was used for the OD segmentation in chapter 3).

Obtaining fovea candidates

When the fovea is present, it is visible as a round dark region delimited by a soft frontier. This region is usually darker in its center and gradually brighter as its frontier is reached. So, conceptually, the fovea can be seen as a soft valley which is discernable if it is deep enough.

Let I be the fovea-containing subimage from the red or the green field (Figure 5.7, images R and G). In order to filtrate from I those minima that are too shallow to be considered the

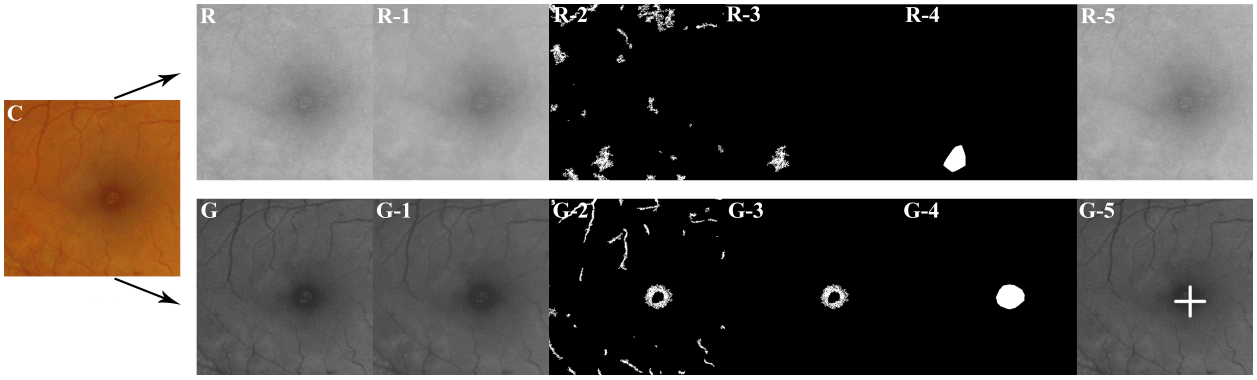


Figure 5.7: *Illustration of the process for accurate fovea center location: (C) Initial RBG fovea-containing subimage. On the right, the top row shows the process performed on the red channel, whereas the bottom corresponds to the process applied to the green component. (R) and (G) Sub-images extracted from the red and green channels of (C), respectively. (R-1) and (G-1) Result of the H-minima transform. (R-2) and (G-2) Regional minima image. (R-3) and (G-3) Image containing the fovea candidate. (R-4) and (G-4) Convex hull of the fovea candidate. (R-5) and (G-5) accurate fovea center location. The candidate in image (R-5) is discarded since it has a circularity value of* 0.3695. *The candidate in (G-5) has a correct size and a circularity value of* 0.0969, *therefore it is validated and the accurate fovea center location is the location of its centroid.*

fovea, the *H-minima* transform (Soille, 1999) is used. This transform suppresses all minima in the image whose depth is less than a scalar h (Figure 5.7, images R-1 and G-1):

$$I_{_F} = Hmin(I, h) \tag{5.9}$$

The choice of the value for parameter h is not especially critical. Tests made with our imagery indicate that a good h value is 15, although values slightly higher or lower are also valid.

Once irrelevant minina in I have been suppressed, surviving regional minima in $I_{_F}$ are segmented. For that, a regional minima is considered to be a connected component of pixels with the same intensity value, whose external boundary pixels all have a higher value. Thus, assigning the value 255 (for 8-bit images) to those pixels in a regional minima and the value 0 to the rest of pixels in $I_{_F}$, a binary image $I_{_{RM}}$ of relevant regional minima is obtained (Figure 5.7, images R-2 and G-2):

$$I_{_{RM}}(x, y) = \begin{cases} 255 & \text{if } (x, y) \in RM, \\ 0 & \text{otherwise.} \end{cases} \tag{5.10}$$

where RM is a regional minima. It is possible that I_{RM} does not contain any segmented connected component. This is due to two possible causes. The first cause is that there was not any valley in image I with a depth higher or equal to h and, therefore, there was not any surviving valley in I_F. The second cause is that, although there were surviving valleys in I_F, they were not regional minima and, therefore, they were not segmented. Anyway, in the case that I_{RM} does not contain connected components, it is considered that the fovea is not present or, at least, not contrasted enough to segment it safely and the rest of the processing is aborted. On the other hand, I_{RM} can contain more than one segmented connected component if there were more than one regional minima in I_F. In this case, the largest connected component in I_{RM} has the highest probability of being the fovea and therefore is selected as the fovea candidate, resulting from this operation the image I_{FC} (Figure 5.7, images R-3 and G-3):

$$I_{FC}(x,y) = \begin{cases} 255 & \text{if } (x,y) \in CC_{MAX}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.11}$$

where CC_{MAX} is the connected component in I_{RM} with the maximum number of elements.

Summarizing, at this point, I_{FC} contains a connected component candidate to be validated as fovea. This validation is performed in the following section by analyzing its size and shape. With respect to this last feature, the circularity of the candidate will be concretely analysed since the fovea has a circular shape. This analysis supposes a problem if it is directly performed on the connected component in I_{FC}, since this component tends to have an irregular contour. To solve this problem, a new connected component composed of the inner points of a polyhedron surrounding the candidate connected component is finally calculated. To obtain this polyhedron, firstly, the centroid of the component is calculated. Then, the hyperplanes supporting the component defined by the normal vectors taken each 5° apart from 0° to 360° are calculated. Finally, inner points of the calculated polyhedron are extracted (Figure 5.7, images R-4 and G-4). This process is illustrated in Figure 5.8.

Validation of fovea candidates and obtention of an accurate fovea center location

This stage is reached if at least one fovea candidate of the two possible, one from each of the two considered channels, has resulted from the previous stage. The aim here is to analyse the

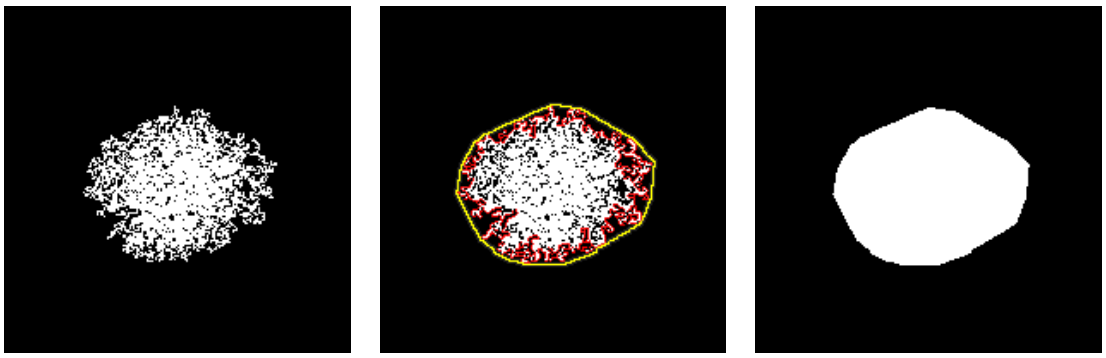


Figure 5.8: *Illustration of the process for the surrounding polyhedron obtention of a segmented object: (a) Segmented object. (b) The boundary of (a) is marked in red color and the convex hull of that boundary is traced in yellow color. The convex hull is obtained by linearly interpolating points of the object contour taken each* 5 *degrees apart from* 0 *to* 360 *degrees. (c) The obtained surrounding polyhedron.*

size and circularity of the obtained candidates to validate or discard them. If a candidate is validated, i.e. if it satisfies the size and circularity criteria, an accurate fovea center location is calculated; otherwise, the fovea is concluded as undetectable (Figure 5.7, images R-5 and G-5).

The following set of rules precisely describes the fovea candidate validation process:

- If two fovea candidates were obtained: the candidates from the red and green channel are analysed in terms of size and circularity. There are three possible outcomes:
  - The two candidates are validated: a new accurate fovea center location is calculated as the centroid of the candidate with higher circularity.
  - Only one candidate is validated: a new accurate fovea center location is calculated as the centroid of the validated candidate.
  - No candidate is validated: the fovea is concluded as undetectable.
- If only one fovea candidate was obtained: the candidate is analysed in terms of size and circularity. If it is validated, a new accurate fovea center location is calculated as the centroid of the candidate. Otherwise, the fovea is concluded as undetectable.
- If no candidate was obtained: the fovea is concluded as undetectable.

The size and circularity criteria used for validating candidates is described below.

The size evaluation criterion decides valid a candidate if its maximum radius is between one half and one fourth of the OD diameter. Anatomically, the fovea radius is between one third and one fourth of the OD diameter, however, the applied processing can increase its size and that is why the radius upper limit is increased from one third to one half of the OD diameter.

With respect to the circularity evaluation criterion, this is based on comparing areas of the candidate object and a circle centered on the centroid of the candidate with a radius equal to the mean radius of the candidate. Concretely, the quotient between non-overlapping area of both objects and the area of the candidate object is calculated. If C denotes the described circle and FC the fovea candidate object, this operation can be mathematically described as:

$$Circ = \frac{\#(C \cup FC) - \#(C \cap FC)}{\#(FC)} \tag{5.12}$$

where # denotes the cardinal operation, i.e., the area of an object in the image. Once the circularity criterion is defined, a threshold to decide when an object is circular enough has to be established. For that, 100 object samples were subjectively classified as being circular or not and their circularity values were calculate according to equation (5.12). Then, these samples were represented in the feature space to be analysed. From this analysis resulted that a circularity value of 0.25 provided a great class separability. Therefore, an object was definitely decided as circular if its circularity value according to equation (5.12) was lower or equal to 0.25, and not circular otherwise. All the described process to decide the circularity of an object is illustrated in Figure 5.9.

5.3.3 Final macular zone identification

The final stage consists on identifying the macular zone making use of the obtained fovea center location. A fovea center location estimate was obtained in section 5.3.1 and an accurate fovea center location could also have been obtained in section 5.3.2. If an accurate fovea center location was obtained, this is the finally selected fovea center location. Otherwise, the finally selected fovea center location is the estimate obtained in section 5.3.1.

To identify the macular zone, a polar coordinate system centered on the fovea center location is used, since this is the system used by ophthalmologists. The coordinate system is set up based on the Early Treatment Diabetic Retinopathy Study Report Number 10 (Early Treatment Diabetic Retinopathy Study Research Group, 1991b). A retinal image is divided

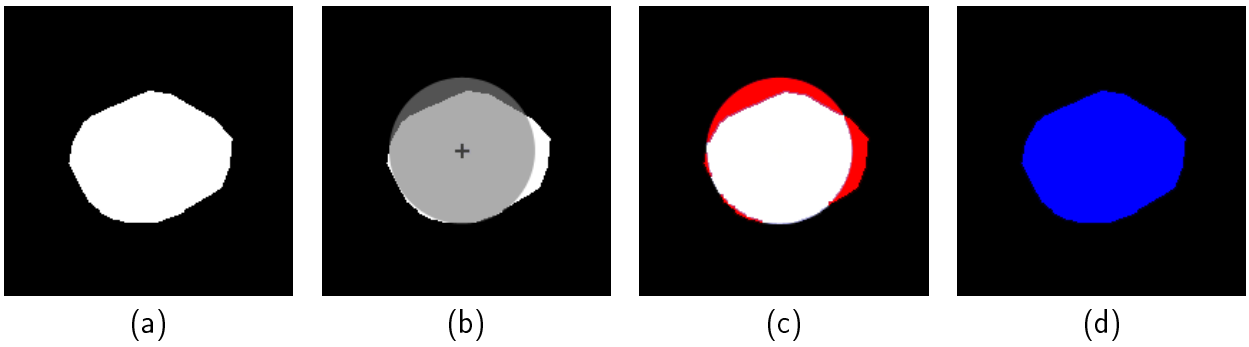


Figure 5.9: *Process to decide the circularity of an object: (a) Object to be analysed. (b) Circle centered on the centroid of (a) with radius equal to the mean radius of (a). (c) The non-overlapping area between (a) and the circle in (b) is represented in red color. (d) The area of (a) is represented in blue color. The circularity of (a) is calculated as the quotient of the red area in (c) and the blue area in (d). In this case, (a) is validated as circular, since its circularity value is* 0.1149.

into ten subfields as presented in Figure 5.10. The radii of the three fovea-centered circles from the innermost to the outermost correspond to (1/3) OD, 1 OD, and 2 OD, respectively. The ten subfields are defined as:

1. Central subfield: area within the inner circle.

2. Four inner subfields (superior, nasal, inferior, and temporal): areas between the inner and middle circles.

3. Four outer subfields (superior, nasal, inferior, and temporal): areas between the middle and outer circles.

4. Far temporal subfield: area temporal to the outer circle and between 7 : 30 and 10 : 30 meridians for the right eyes or between 1 : 30 and 4 : 30 meridians for the left eyes.

The described system is used by ophthalmologists to assess illness condition by labeling lesions in each subfield. The illness diagnosis is made by counting detected lesions, taking into account that their relevance varies according to where they are located. This relevance decreases from the most relevant location, the central subfield (threat of central vision loss), to the less relevant location, the far temporal subfield (threat of peripheral vision loss).

Figure 5.11 shows application examples of the presented methodology for identifying the macular zone. The first row shows results obtained on images affected by DR but with no risk of ME, the second row shows results on images from patients affected by DR with risk of ME and the last row shows results on healthy images.

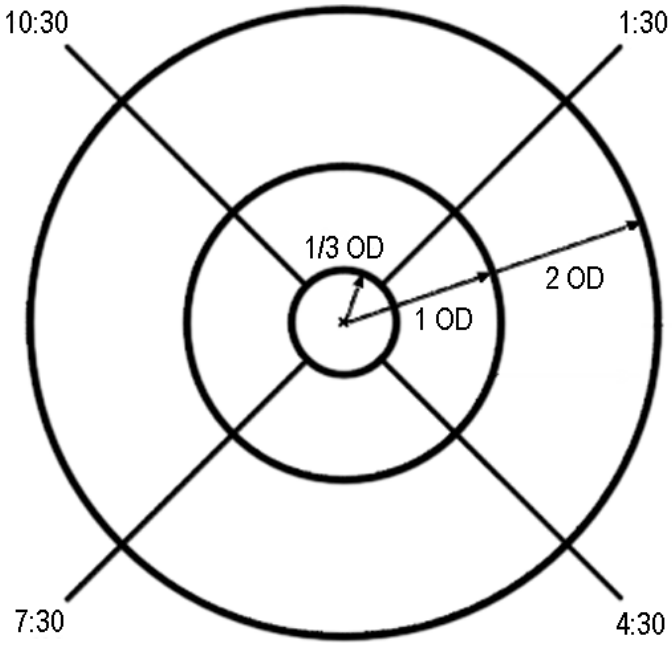


Figure 5.10: *Polar coordinate system centered on the fovea.*

5.4 Experimental results

This section presents the results obtained with the presented methodology for macular zone identification. In the first subsection, the measures for performance evaluation and comparison are defined. The second subsection presents the results obtained with the proposed methodology. Finally, the third subsection compares the obtained results with those presented in other relevant works present in the literature.

5.4.1 Performance measures

To determine an evaluation criterion, the literature was revised to check criteria previously used by other authors. Most of the published works has used the distance between the obtained fovea center location and the real fovea center location as an evaluation measure (Fleming et al., 2007; Niemeijer et al., 2007, 2009; Tobin et al., 2007). This has been justified by the lack of border definition of the fovea region. Therefore, this distance-based criterion was selected to evaluate the presented work, although some changes were introduced. Most of publications have established that an obtained fovea center location is correct if the distance between this one and the real fovea center location is less than 1 times

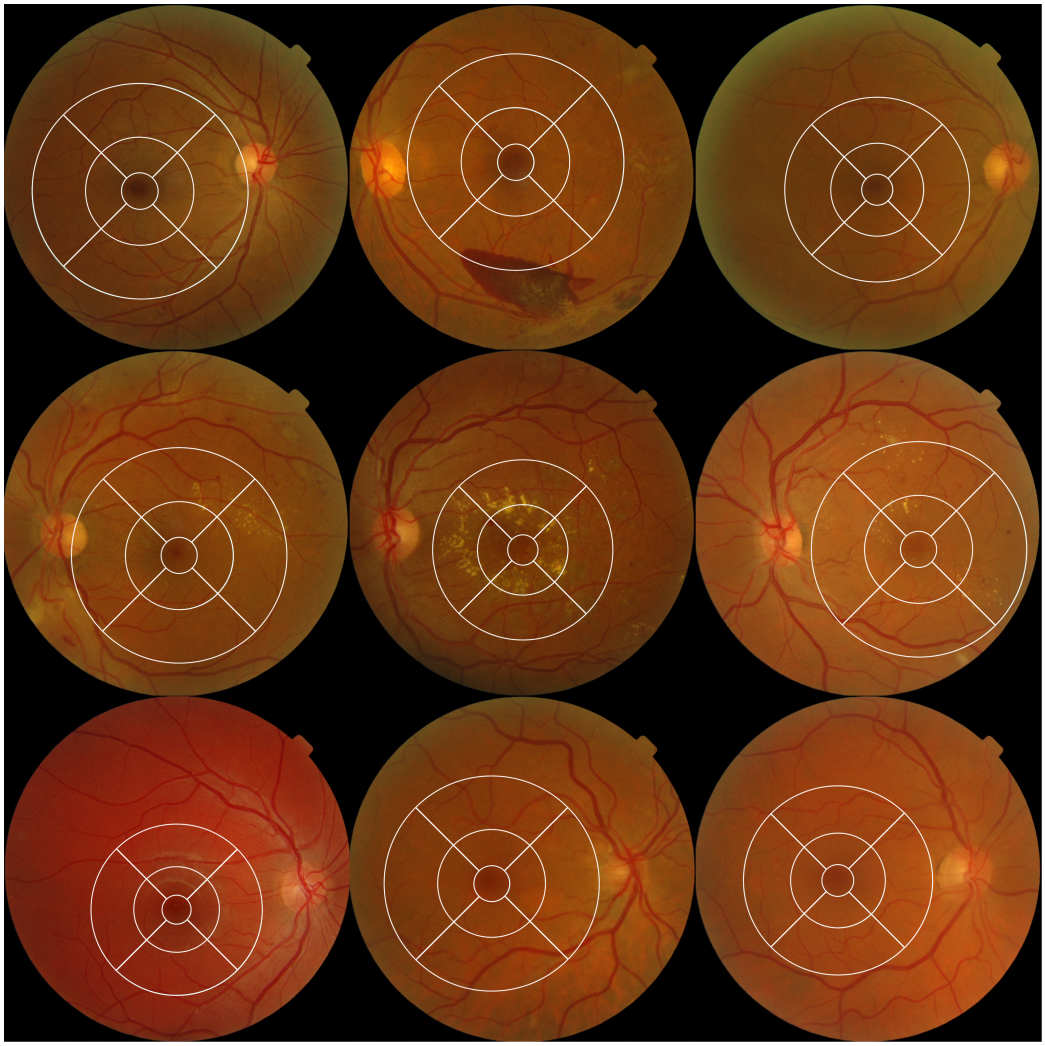


Figure 5.11: *Examples of the macular zone identification. The first row shows three retinographies from patients with DR and with no risk of ME, the second is composed of three retinographies from patients affected by DR and with risk of ME and the last three images are from healthy retinas.*

the OD radius (1R criterion). This distance is also accepted here but, in order to perform a more precise evaluation, two more distances were considered. These distances are 0.5 times the OD radius (0.5R criterion) and 0.25 times the OD radius (0.25R criterion). Therefore, these distances are used to create a quality scale of obtained fovea center location as follows:

- Excellent: an obtained fovea center location is "Excellent" if it satisfies the 0.25R criterion, i.e., if it is at a distance less or equal than 0.25 times the OD radius from the real fovea center location.

- Good: an obtained fovea center location is "Good" if it does not satisfy the 0.25R criterion but satisfies the 0.5R criterion. This is, if it is at a distance greater than 0.25 and less or equal than 0.5 times the OD radius from the real fovea center location.

- Fair: an obtained fovea center location is "Fair" if it does not satisfy the 0.5R criterion but satisfies the 1R criterion. This is, if it is at a distance greater than 0.5 and less or equal than 1 times the OD radius from the real fovea center location.

- Poor: an obtained fovea center location is "Poor" if it does not satisfy any of the previous criteria. This is, if it is at a distance greater than 1 times the OD radius from the real fovea center location.

Thus, the interval "Excellent - Fair" is equivalent to the 1R criterion and can be used for comparison to other methods. In addition, the study of all categories provides a more precise method evaluation than that provided by the 1R criterion on its own.

It is important to point out that, since the OD size presents a substantial variation, using the OD radius measured in each image to apply the distance criteria described above can distort results. To minimize this difficulty, the mean OD radius measured in section 5.3.1 was used to this effect. By this way, the OD radius in the described distance criteria goes from being a variable to a constant.

5.4.2 Proposed method evaluation

The presented methodology for identifying the macular zone was tested on the 1200 images in the MESSIDOR database described in appendix A. Results of 38 of these images were not evaluable since the fovea was completely indiscernible, therefore, final evaluated results are on 1162 images.

Table 5.1 shows the results obtained with the presented algorithm. The table shows the average percent of macular zones successfully identified according to the quality scale defined in the previous section for each diagnosis case of DR Grade - Risk of ME (cases diagnosed according to gradings described in section 2.2; DR condition from 0 to 4 and risk of ME from 0 to 2). The fovea center location was *Excellent* in 66.95% of cases, *Good* in 21.60% of cases, *Fair* in 8.87% of cases and *Poor* in the resting 2.58% of cases. These figures mean that the algorithm appropriately identified the macular zone in 97.42% of images (quality range *Excellent* - *Fair*), which means success in 1132 cases out of the whole set of 1162

DR Grade - Risk of ME	Cases	Excellent	Good	Fair	Poor	Excellent - Fair
0-0	535	71.78%	20.37%	6.36%	1.49%	98.51%
1-0	142	76.76%	17.60%	3.53%	2.11%	97.89%
1-1	5	80.0%	20.0%	0.0%	0.0%	100.0%
1-2	6	100.0%	0.0%	0.0%	0.0%	100.0%
2-0	179	67.04%	22.91%	7.82%	2.23%	97.77%
2-1	28	71.43%	17.86%	7.14%	3.57%	96.43%
2-2	37	51.35%	29.73%	18.92%	0.0%	100.0%
3-0	97	45.36%	28.86%	18.56%	7.22%	92.78%
3-1	41	65.85%	24.39%	9.76%	0.0%	100.0%
3-2	92	48.91%	22.83%	20.65%	7.61%	92.39%
Total	1162	66.95%	21.60%	8.87%	2.58%	97.42%

Table 5.1: *Results of the algorithm for macular zone identification according to the defined quality scale. Results are detailed per diagnosis case and are given as the average percent of obtained fovea center locations in each quality category.*

images. The success percent for retinas affected by DR and for healthy retinas was 96.49% and 98.51%, respectively. For images from patients affected by DR, the success rate was 96.17% for images diagnosed with risk of ME and 96.65% for those images with no risk. Attending to these results, there is not any observable pattern relating illness condition to outstanding degradation in the algorithm performance. Table 5.2 offers an extended study of the results detailed per ophthalmic center.

The presented algorithm was experimentally implemented in Matlab. Tests of computational efficiency of this implementation were performed using a PC equipped with an Intel Core2Duo CPU at 2.13 GHz and 2 GB of RAM capacity. For 1162 executions of the algorithm, the measured average computational time was 10.88 sec. with a standard deviation of 0.19 sec.

5.4.3 Comparison to other methods

In this section, the results of the presented methodology are compared to those of other relevant published works. Concretely, the considered works are those of: Fleming et al. (2007), Niemeijer et al. (2007, 2009) and Tobin et al. (2007). Fleming et al. (2007) evaluated their results according to the 1R criterion. Niemeijer et al. (2007, 2009) used a distance in number of pixels to decide when an obtained fovea center was correct. In spite of it, the

	Hôpital Lariboisière Paris		Faculté de Médecine St. Etienne		LaTIM - CHU de Brest	
DR Grade - Risk of ME	Cases	Excellent - Fair	Cases	Excellent - Fair	Cases	Excellent - Fair
0-0	151	99.34%	181	100.0%	203	96.55%
1-0	26	96.15%	68	98.53%	48	97.92%
1-1	2	100.0%	1	100.0%	2	100%
1-2	2	100%	2	100.0%	2	100%
2-0	49	97.96%	62	96.77%	68	98.53%
2-1	13	92.31%	9	100.0%	6	100.0%
2-2	8	100%	20	100.0%	9	100.0%
3-0	46	89.13%	18	94.44%	33	96.97%
3-1	25	100.0%	7	100.0%	9	100.0%
3-2	60	90.67%	20	95.0%	12	75.0%
Total	382	96.86%	388	98.71%	392	96.68%

Table 5.2: *Extended results of the algorithm for macular zone identification according to the defined quality scale. Results are detailed per diagnosis case and ophthalmic center and are given as the average percent of obtained Excellent - Fair fovea center locations.*

distance used by those authors, attending to the image resolution and FOV size used by them, is approximately equal to the distance established by the 1R criterion. Therefore, the results of these works are comparable with those of our methodology using this criterion. On the other hand, Tobin et al. (2007) considered valid an obtained fovea center if it was located at a distance less or equal than 2 times the OD radius from the true fovea center (2R criterion). Therefore, our work was also evaluated using the 2R criterion with the aim of comparing both approaches.

Table 5.3 shows the results of the above mentioned methodologies compared those of the one presented in this chapter (referred to as *Morphological Method*). As it can be checked, the results obtained with the presented methodology outperforms all analysed approaches. Datasets used to test those works are not of public access, therefore comparison under identical conditions is not possible. Anyway, the image description provided by their corresponding authors (images from screening programs of DR), indicate that the nature of the images is similar as the nature of the images in the MESSIDOR database. Therefore, accepting this hypothesis, comparison in Table 5.3 shows evidences that the new methodology presented in this chapter enhances the identification of the macula region.

Method	Images	1R criterion	2R criterion
Fleming et al. (2007)	1056	95.50%	-
Niemeijer et al. (2007)	500	94.40%	-
Niemeijer et al. (2009)	500	96.80%	-
Tobin et al. (2007)	345	-	92.50%
Morphological Method (this work)	1162	**97.42%**	**98.36%**

Table 5.3: *Results of the presented algorithm for macular zone identification compared to other approaches present in the literature.*

5.5 Discussion and conclusion

Two types of criteria have been historically used for locating the macula region: anatomical feature-based and visual feature-based criteria. The methodology presented in this chapter takes advantage of both approaches. It is divided into three main stages: (A) obtention of a fovea center estimate by using anatomical feature-based criteria, (B) obtention of an accurate fovea center location if this is detectable by using visual feature-based criteria, and (C) final fovea center decision and identification of the macular zone.

The main contribution of the presented work is the processing applied in stage (B). To the best of our knowledge, all published methodologies based on visual features of the fovea perform a rigid processing in which the fovea is assumed to be always present. On the contrary, the morphological processing applied in our work discards the accurate location of the fovea if no valid candidates are found. In this case, the fovea center estimate obtained in stage (A) is used to identify the macular zone. This approach makes the whole process of macular zone identification more safe and stable.

To test the proposed methodology, a quality scale based on the distance of the obtained fovea center to the true fovea center was defined. This scale contains four categories: *Excellent*, *Good*, *Fair* and *Poor*. Table 5.1 and 5.2 summarizes the obtained results according to this quality scale. The methodology obtained valid fovea centers, those in the range *Excellent* - *Fair*, in 97.42% of the 1162 evaluable images in the MESSIDOR database. Furthermore, the methodology obtained foveas in the range *Excellent* - *Good* in 88.5% of cases. Attending to detailed results in both tables, it can be checked that the algorithm behaviour is robust in all clinic situations. On the other hand, Table 5.3 offers a performance comparison to other relevant published methods. As it can be checked, the presented methodology outperforms

all analysed approaches. On the other hand, the algorithm computational efficiency should also be mentioned since it is also related to testing. Tests made on a mid-range PC gave a mean computational time of 10.88 sec. with a standard deviation of 0.19 sec. Although these figures are already acceptable, since the method implementation is experimental, they can be easily improved.

In the context of the automated detection of DR, a robust macular zone identification is of vital importance since relevance of lesions is related to their location with respect to the macula. The reliability, robustness and efficiency obtained with the methodology presented in this chapter, show evidence that its integration in a future system for the early automated detection of DR may be successful.

Chapter 6

A Specialized Function for Quality Evaluation of Retinal Vessel Segmentations

METRICS based on contingency tables for binary classification have been widely used for evaluating the performance of algorithms for retinal blood vessel segmentation. Particularly, *accuracy* has been mostly used as a function for evaluation of global quality of retinal vessel segmentations. Therefore, this metric has been mostly used to compare the behaviour of the mentioned segmentation algorithms. However, the weaknesses shown by this metric for this purpose are identified and discussed in this chapter. These difficulties ultimately lead to poor matching with human quality perception.

This chapter proposes a new function for enhancing the evaluation of global quality of retinal vessel segmentations. This function shows better behaviour than *accuracy* in those situations in which this metric shows difficulties.

The chapter is structured in **eight sections**. The **first section** introduces the problem of quality evaluation of segmented images. Focusing on segmentation of retinal vessels, objective evaluation of these segmentations is briefly commented. The application of *accuracy* for this purpose is reviewed and its use in the literature of this topic is referenced. Finally, the identified difficulties of this metric in this use are summarized.

The **second section** of this chapter deals with the application of metrics based on contingency tables for binary classification in the field of retinal vessel segmentation, making more emphasis in the use of *accuracy*.

The **third section** presents a study consisting on the characterization of the vascular tree by means of the measurement of some relevant numerical features.

The **fourth section** represents the core of the motivation of the work presented here. This section, taking as basis the study presented in the third section, identifies the main weaknesses of *accuracy* in its use for global quality evaluation of retinal vessel segmentations.

In the **fifth section**, the new function for evaluating the global quality of retinal vessel segmentations is presented. The proposal, known as *CAL*, is based on the mathematical evaluation of three main measurable features which describe vascularity: vascular tree connectivity, area and length.

The **sixth section** consists on a behaviour comparison between *CAL* and *accuracy*. These quality evaluation functions are here evaluated on situations representative of the *accuracy* weaknesses.

The *CAL* behaviour was also compared to other generalist quality evaluation functions that, to the best of this author's knowledge, have not been used in the field of vascular tree segmentation. This comparison is described in the **seventh section** of this chapter.

Finally, this chapter concludes in the **eight section** with the author's discussion and conclusion.

6.1 Introduction

The quality evaluation of the segmentations obtained by an automated method is an important issue. As discussed by Shi and Sun (1999), subjective and objective evaluations are the two ways for assessing segmentation quality. Subjective testing quantifies the viewer's dissatisfaction in observing the segmented image in place of the original. On the other hand, objective evaluation gives distortion between the original and the segmented image mathematically. However, it may not reflect observers' visual perception of distortion as desired (Paglieroni, 2004).

This chapter focuses on the objective evaluation of retinal vascular tree segmentations. To perform this evaluation, a reference segmentation, usually called reference gold standard, and a quality evaluation function (QEF) are used in the literature. The reference segmentation is made by a medical expert, who manually labels the original image. This way of measuring quality faces the problem that human delineations of medical images are not true gold standards (see Bioux et al., 2007, for a comprehensive discussion on this issue), but it is the most objective solution. On the other hand, metrics based on contingency tables for binary

classification (see appendix B for a complete description) are the most widely used in this topic, especially over the last years, as QEFs. The most commonly-used metrics from this family are *sensitivity*, *specificity* and *accuracy*. Specifically, *accuracy* (*Acc*) has been mostly used to quantify segmentation global quality (Al-Rawi and Karajeh, 2007; Al-Rawi et al., 2007; Cinsdikici and Aydın, 2009; Espona et al., 2007; Gardner et al., 1996; Hoover et al., 2000; Jiang and Mojon, 2003; Lam and Yan, 2008; Martinez-Perez et al., 2007; Mendonça and Campilho, 2006; Niemeijer et al., 2004; Ricci and Perfetti, 2007; Sinthanayothin et al., 1999; Soares et al., 2006; Staal et al., 2004). Therefore, *Acc* has been accepted as the main metric for comparing segmentation algorithms and for providing a global idea of algorithm behaviour (in fact, *Acc* was used with these purposes in chapter 4 of this thesis).

The difficulties involved by the quality evaluation of retinal vessel segmentations have been previously pointed out and discussed by Niemeijer et al. (2004). In our opinion, *Acc* may be complemented with other more specialized QEFs for this evaluation, since it shows poor correlation with human quality perception. This poor matching is a consequence of a set of difficulties derived from the nature of this metric and the features of the vascular tree. *Acc* is based on comparing segmentations, pixel by pixel, with hand-labeled images for determining whether a pixel is well-classified or not into the two classes, vessel or background. This pixel-by-pixel comparison makes *Acc* sensitive to slight differences between segmentations and gold standards. Moreover, *Acc* presents difficulties when these two classes are considerably unbalanced in their number of instances, as it tends to be insensitive to poor classification of the minority class.

6.2 Metrics based on contingency tables application for quality assessment of retinal vessel segmentation

Metrics based on contingency tables for binary classification are widely used for performance evaluation in classification problems. Although their theoretical basis is treated in Appendix B, the use of these metrics for quality assessment of retinal vessel segmentation is studied in this section.

Automated segmentation of the vascular tree consists on automatically obtaining a binary image of vessels and background from an eye fundus image. For a numerical assessment of the quality of a segmentation generated by means of an automated process, a binary image

representing the "perfect" case is necessary for comparison purposes. This last is commonly known as reference *ground truth* or *gold standard* and is performed by a medical expert by manually labeling the original image (this person is usually called human observer). Then, taking as actual classes $\{Vessel, Background\}$ those established in the reference gold standard and as predicted classes $\{Vessel\ detected, Background\ detected\}$ those produced in the automated segmentation, metrics based on contingency tables for binary classification can be directly applied for measuring automated segmentation quality. Thus, *sensitivity*, Se, provides the ratio of well-classified $Vessel$ instances, *specificity*, Sp, provides the ratio of well-classified $Background$ instances and *accuracy*, Acc, gives a global measure of classification quality, as it is the ratio of pixel instances correctly mapped to both $Vessel\ detected$ and $Background\ detected$ classes.

Most works present in literature on automated vessel segmentation produce a membership ranking which may be represented by a gray-level image. In this image, the value of a pixel indicates the degree to which this pixel is a member of the class $Vessel\ detected$ (or, by complementarity, of the class $Background\ detected$). Thus, the final behaviour of these methods depends on a parameter for determining the value for the best $Vessel\ detected$-$Background\ detected$ class separability. To determine this value and then evaluate the general behaviour and quality of segmentation methods on a set of images, *maximum accuracy average* ($MAccA$) and, more frequently, *maximum average accuracy* (MAA) are usually used. $MAccA$ is calculated by finding for every image the value of the method's parameter which renders the maximum Acc value for each of them and taking the average of all these maximums. Thus, $MAccA$ provides a measure concerning the maximum quality reachable with the method under evaluation for the set of images if the best parameter value could be set for each image. On the other hand, the MAA is calculated as the maximum average of Acc produced by applying the same parameter value for all images. By this way, MAA gives a global idea of the quality reachable by applying a unique parameter value for all the images in the set. Additionally, average Se and Sp, denoted by $\overline{Se}$ and $\overline{Sp}$, are usually also provided. $\overline{Se}$ offers a measure concerning the amount of well-segmented vasculature and $\overline{Sp}$ indicates the amount of well-detected background.

Algorithm performance measurement is also often complemented with receiver operating characteristic (ROC) curves. The area under the curve (AUC), which is 1 for a perfect system, is a single measure also used to quantify this performance.

This chapter proposes a QEF for measuring the general quality of retinal vessel segmentations. Since *Acc* is the generalist metric currently used for this purpose, the following study focuses on this metric and the new proposal.

6.3 Numeric characterisation of the retinal vascular tree

This section focuses on the study of some relevant numerical features of the vascular tree. This work forms the basis for the identification of the main *Acc* weaknesses discussed below in section 6.4.

The experimentation developed in this section as well as in the rest of the chapter was carried out using the test set of the DRIVE database (see descriptions in Appendix A). This set, referred to as the study set hereafter, was concretely selected because it offers two sets of gold standard images made by two different observers. This material was necessary for the development of some experiments described in this chapter.

The retinal vascular tree is a capillary structure whose function is feeding the retinal cells. It is a tree-like connected structure which originates in the optic disc with some main gross blood vessels, which progressively split into many other smaller vessels for covering the retina. Thus, the vascular tree is generally considered to be composed by two kinds of vessels: main and gross vessels, and thin vessels.

To evaluate the features of the vascular tree in retinal images, experiments were carried out using the two different gold standard sets made by the two observers in the study set. Firstly, the average ratio of pixels of the whole vasculature with respect to the total pixels in the field-of-view (FOV) of a retinography was measured. With this purpose, besides the manual segmentations (see Figure 6.1, image (c)), their associated FOV masks (Figure 6.1, image (b)) available in this database were also used. Another interesting feature is the average ratio of gross and thin vessel pixels which constitute a vascular tree. To calculate it, gross and thin vessels have to be previously separated for each gold standard image in the study set. This separation was carried out using the approach defined and used in section 4.4.2 for this purpose. This is, a vessel is considered thin if its width is lower than 50% of the width of the widest optic disc vessel and, otherwise, the vessel is considered non-thin or gross. Once gross and thin vessels were separated from gold standards (see Figure 6.1, images (d) and

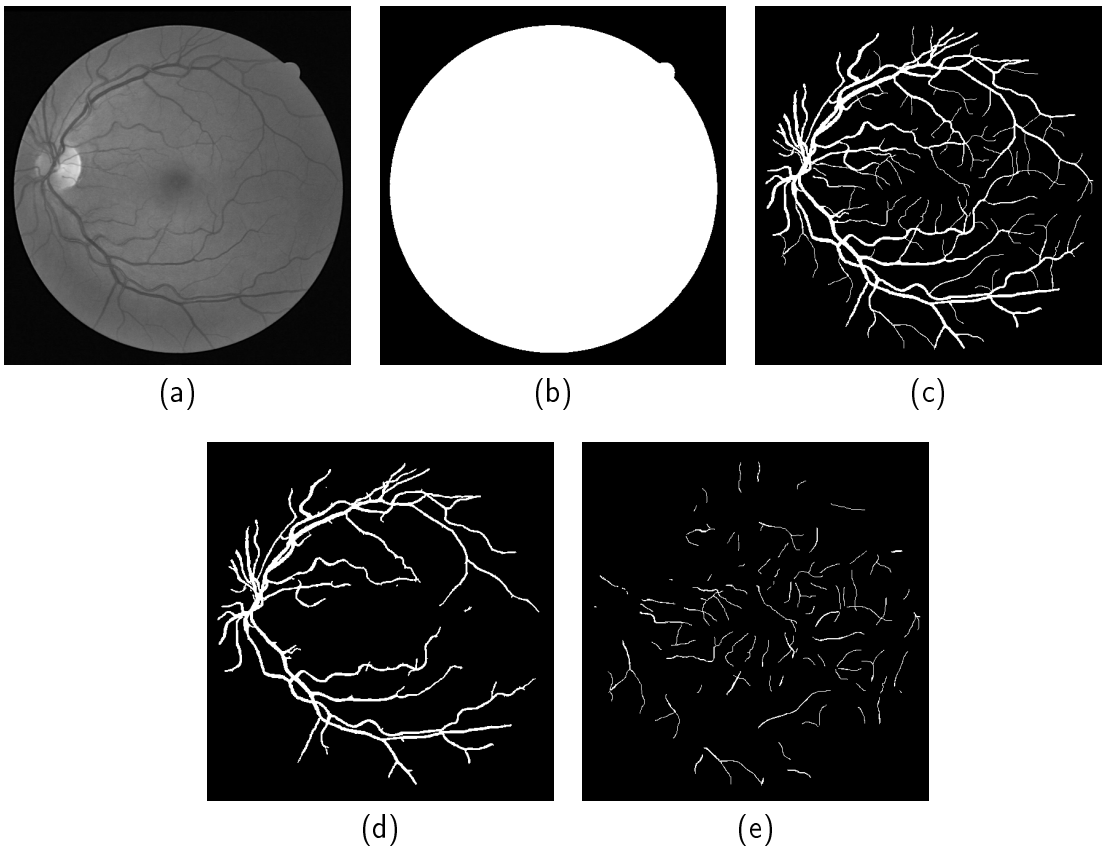


Figure 6.1: *Example of sets of thin and gross vessels. (a) Green channel of an image from the test set of the DRIVE database. (b) Mask that delimits the FOV from the background. (c) Gold Standard image by the first observer. (d) Set of gross vessels obtained from (c). (e) Set of thin vessels obtained from (c).*

(e)), the average ratio of gross and thin vessel pixels which constitute a vascular tree was calculated.

Table 6.1 summarizes the results of the experiments. As it can be observed in the first row, the vascular tree only represents, on average, around 12% of the pixels in the retinography FOV. This value is quite stable for the two sets of manual segmentations. This value indicates a great unbalance between the vessel and background classes. On the other hand, the second and third rows show up the existence of a great difference in terms of ratio of pixels between thin and gross vessels, as thin ones only constitute around 21% of the vasculature pixels. This value is also quite similar for both sets of images and implies that only 3% of the pixels in the FOV belong to thin blood vessels.

Feature	1st Observer	2nd Observer
Vascular Tree / FOV	0.1274	0.1226
Thin Vessels / FOV	0.0263	0.0270
Gross Vessels / FOV	0.1010	0.0956
Thin Vessels / Vascular Tree	0.2068	0.2202
Gross Vessels / Vascular Tree	0.7932	0.7798

Table 6.1: *Study of vascular tree features in terms of average ratio for the 20 images in the DRIVE database test set.*

6.4 Motivation: weaknesses of the accuracy

Here, the development of a new QEF for assessing the quality of retinal vessel segmentations is motivated. This motivation lies on the identification of the following weaknesses of the *Acc* for this use: low sensitivity to poor blood vessel segmentation, very high insensitivity to poor thin blood vessel segmentation, strong dependence with the gold standard and poor matching with human perception.

Next, a description of the mentioned weaknesses is presented.

6.4.1 Low sensitivity to poor blood vessel segmentation

As shown in Table 6.1, in a digital image the whole vascular tree only constitutes about 12% of the retina pixels. Specifically, and according to this value, if only 12% of the pixels in the image are vessel pixels, for a segmentation in which all pixels were classified as background, its *Acc* value would be 0.88. This low sensitivity of *Acc* to poor blood vessel detection is a consequence of a general *Acc* difficulty when the classes are considerably unbalanced in their number of instances. In this case, bad classification of the minority class has low impact on *Acc*. The degree of insensibility of *Acc* to this situation increases with the unbalance magnitude between the two classes.

6.4.2 Very high insensitivity to poor thin blood vessel segmentation

Despite the small size of thin vessels, their assessment is relevant from a medical viewpoint, since they may also suffer damages and therefore contribute diagnostic information. That is why the segmentation of these vessels is an important goal which, to the best of this author's knowledge, remains unsolved and inappropriately discussed. Attending to data in Table 6.1, thin vessels only represent about 3% of the pixels of the whole retina. This makes *Acc* almost

insensitive against the absence of thin-vessel detection. Specifically, for segmentations with no detected thin-vessel pixels, *Acc* can reach 0.97.

6.4.3 Strong dependence with the gold standard

Objective quality assessment of automated segmentations is performed by mathematically evaluating distortion between these segmentations and reference gold standards. These reference gold standards are generated by human observers with the help of a paint program by manual vessel tracing. Therefore, this objective quality evaluation includes a subjective factor introduced by the human observer in the reference gold standards. In their generation, personal aspects such as fatigue tolerance, patience and particular hand precision, among other factors, influence results. Furthermore, different subjective perception or professional criteria may also affect the results contributed by different observers. Consequently, the differences between the gold standards contributed by different observers for the same image are easily visible. Concretely, differences in vessel tracing (especially thin ones), width, length or even in painting are always present (see examples taken from the study set in Figure 6.2). That is why, although the mathematical comparison with these gold standards is the most objective way of evaluating vessel segmentation quality, these human-made references cannot be considered true gold standards (see Bioux et al., 2007). This introduces an uncertainty factor in quality evaluation which should be considered.

In view of the foregoing, a QEF for evaluating retinal vessel segmentation quality should minimize the impact of subjectivity in the reference gold standard. It is important to highlight that any QEF would be sensitive to differences in the presence and absence of vessels. However, the impact of slight differences in vessel tracing or width can be minimized by the QEF and *Acc* shows difficulties in these cases. This fact is illustrated by the example described below.

Figure 6.3 shows an illustrative and simplified study case. Image (a) is a small section of a gold standard of the study set made by the first observer, image (b) is the same section of the corresponding gold standard image created by the second observer, and image (c) is the color composition of both images. The black color in image (c) indicates coincidence in background, pixels in yellow color represent coincidences of both vessel segmentations, and red and green show vessel disagreement. The red pixels belong to the first observer, whereas the green ones belong to the second. Image (d) represents a null vessel-detection case (i.e.,

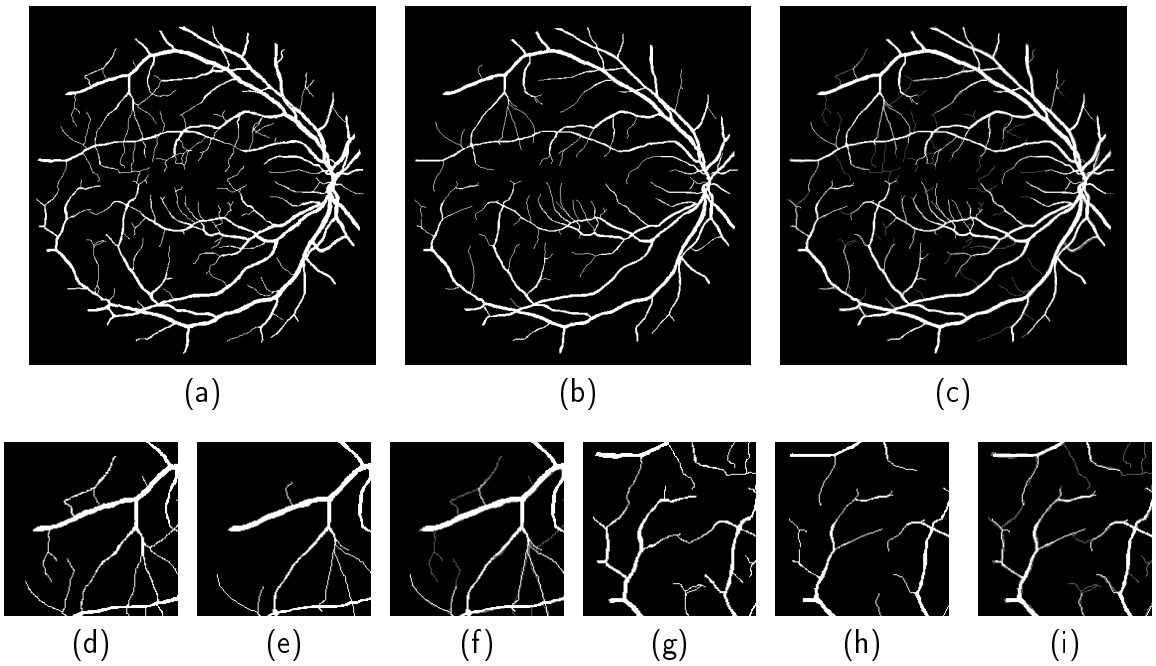


Figure 6.2: *Illustration of differences of gold standard images made by two different observers. (a) Gold standard image by the first observer. (b) Gold standard image by the second observer. (c) Union of (a) and (b); (a) is represented in dark gray color and (b) in white color. (d), (e) and (f) Close ups of the same zone of (a), (b) and (c), respectively. (g), (h) and (i) Another example showing close ups of another zone of (a), (b) and (c), respectively.*

all pixels were detected as background). Considering image (b) as a segmentation and image (a) as the reference gold standard, the *Acc* value for image (b) would then be:

- $Acc = \frac{(46+2773)}{(46+96+2773+85)} = 0.9397$

On the other hand, if the Acc value for image (d) is calculated by using image (a) again as reference gold standard, it is obtained:

- $Acc = \frac{(0+2858)}{(0+142+2858+0)} = 0.9527$

Paradoxically, *Acc* shows that image (d) is a better segmentation than image (b) when compared to image (a). That is why each vessel pixel in image (b) which does not exactly match to a vessel pixel in the reference gold standard image (a) is doubly penalized with a FP (green pixels in image (c)) and a FN (red pixels in image (c)). However, as all pixels in image (d) are background, all these pixels are penalized only with a FN.

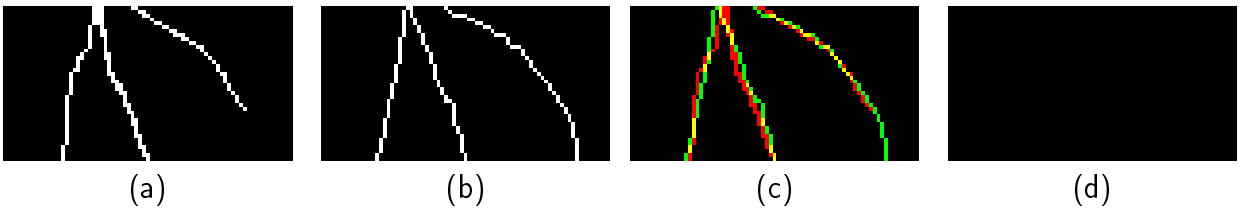


Figure 6.3: *Illustration of the impact of the reference gold standard on Acc. (a) Subimage of a reference gold standard. (b) Subimage of the gold standard by the second observer. (c) Color composition of (a) and (b). (d) Black image with the same size of (a) and (b). The Acc value is* 0.9397 *for (b) and* 0.9527 *for (d).*

6.4.4 Poor matching with human perception

All these described weaknesses of *Acc* lead to poor matching between *Acc* evaluation and the subjective perception of quality. Figure 6.4 shows an illustrative example of this mismatching. Image (a) is a reference gold standard from the study set. Images (b)-(d) are segmentations created from this image with an *Acc* value of 0.9681. These images were generated with the following criteria:

- Figure 6.4, image (b): taking gross blood vessel pixels and redistributing them at the thin blood vessels border, making them wider.

- Figure 6.4, image (c): taking thin blood vessel pixels and redistributing them at the gross blood vessels border, making them wider.

- Figure 6.4, image (d): taking thin blood vessel pixels and redistributing them in the background as random noise.

As it can be observed, although *Acc* indicates equal quality for these three segmentations, human observers' perception can be expected to differ. Since segmentation (b) preserves all the vessels, in spite of showing some slight deficiencies in the width of some vessels, it would surely be widely qualified by human observers as the best of the three. On the other hand, human observers can also be expected to decide that (c) is a more valuable segmentation than (d). Both segmentations detected a similar amount of vascularity. However, noise in image (d) degrades the visual perception of quality more than the vessel-width deficiencies in image (c).

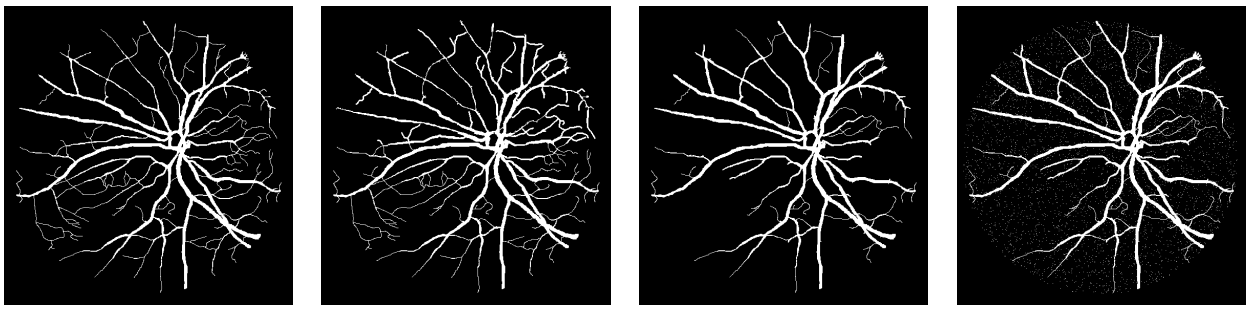


Figure 6.4: *Examples of Acc mismatching with human perception. (a) Reference gold standard. (b)-(d) Images with an Acc value of* 0.9681. *Although Acc indicates equal quality for the three segmentations, the subjectively-perceived quality would tend to rank the images.*

6.5 New function for quality assessment of retinal vessel segmentation

Considering the weaknesses of *Acc* and the structural properties of the vascular tree, a QEF based on evaluation of specific descriptive features is proposed. According to the author's perception, connectivity (C), area (A) and length (L) were identified as three measurable features providing a good description of the vascular tree. Hence, the QEF described below is based on three functions (C, A and L) which evaluate each feature. These functions are defined to take values within the $[0, 1]$ interval and are described as follows:

- Connectivity (C): this factor assesses the fragmentation degree of the segmentation relative to the gold standard. Since the vascular tree is a connected structure, a good vascular segmentation is expected to have only a few connected components (ideally one). So, this factor penalizes fragmented segmentations according to:

$$C(S, S_G) = 1 - min\left(1, \frac{|\#_C(S_G) - \#_C(S)|}{\#(S_G)}\right) \tag{6.1}$$

  where min is the minimum function, $\#_C(S_G)$ and $\#_C(S)$ stand for the number of connected components in gold standard S_G and segmentation S respectively, and $\#(S_G)$ denotes the cardinal of S_G. The min function is introduced to prevent inadequate behaviour of C in the extremes of its domain interval. Note that, for the sake of simplicity, segmentation is here (and hereafter) referred to as the set of vessel pixels exclusively, therefore excluding the set of background pixels from this term.

- Area (A): this factor, based on the Jaccard coefficient (Jaccard, 1901), evaluates the degree of overlapping areas between the segmentation and the gold standard and is defined as:

$$A(S, S_G) = \frac{\#((\delta_\alpha(S) \cap S_G) \cup (S \cap \delta_\alpha(S_G)))}{\#(S \cup S_G)} \tag{6.2}$$

  where S_G is the gold standard and S is the segmentation to be evaluated. Function δ_α is a morphological dilation using a disc of α pixels in radius. The introduction of this operator provides tolerance with slight differences in vessel width. The magnitude of this tolerance is controlled by the value of α.

- Length (L): this factor measures the degree of coincidence between the segmentation and the gold standard in terms of total length. It is formally expressed as:

$$L(S, S_G) = \frac{\#((\varphi(S) \cap \delta_\beta(S_G)) \cup (\delta_\beta(S) \cap \varphi(S_G)))}{\#(\varphi(S) \cup \varphi(S_G))} \tag{6.3}$$

  where φ is an homotopic morphological skeletonization (Serra, 1982) and δ_β is a morphological dilation with a disc of β pixels in radius for decreasing the impact of slight differences in vessel tracing. The value of β controls the sensibility degree to these differences.

A QEF based on these features is proposed to be a function f defined as

$$\begin{aligned} f &: \Re^3 \to \Re \\ (C, A, L) &\mapsto [0, 1] \subseteq \Re \end{aligned} \tag{6.4}$$

and defined increasing monotone:

$$\begin{aligned} f(x_1, x_2, x_3) &\geq f(y_1, y_2, y_3) \\ &\Updownarrow \\ x_i &\geq y_i \ , \quad \forall i = 1, 2, 3 \end{aligned} \tag{6.5}$$

Thus, f is a generic QEF, defined within the $[0, 1]$ interval, function of the set of descriptive features C, A and L, and with an increasing monotone behaviour with respect to these ones. The values 0 and 1 denote the worst and perfect cases of segmentation, respectively.

In this chapter, it is proposed as a specialized QEF for global quality assessment of retinal vessel segmentations, in the following the CAL, the product of C, A and L:

$$f(C, A, L) = C \times A \times L \equiv CAL \tag{6.6}$$

Note that any function fulfilling (6.4) and (6.5) may be considered. Since the product of C, A and L tends to equitably preserve quality in all features, this solution was considered to be an appropriate choice for the goal of this work. On the other hand, remarkable CAL interpretability regarding results must also be highlighted. To this respect, CAL offers certain information on the quality in the three main vascularity features. It allows detecting the concrete weaknesses of the algorithms in a more objective and methodical way.

6.6 Experimental results

The behaviour of CAL and Acc is studied and compared in this section. Firstly, the values of the CAL parameters have to be decided. The first subsection studies this parameter setting. Then, in the following three subsections, theoretical cases representative of Acc weaknesses described in section 6.4 are used to assess CAL in these situations. The last subsection formally studies and compares the matching degree of both QEFs with human quality perception.

6.6.1 Parameter setting

Before using CAL, the values of the α and β parameters defined in equations (6.2) and (6.3) must be fixed. Note that A and L are increasing monotone functions with respect to their parameters α and β, respectively. The low values of these parameters make A and L very rigorous with respect to differences between images, while high parameter values make functions very tolerant to these differences, entailing this last case loss of descriptive potential. In this sense, the main idea of the study was to find an optimal compromise between rigor in the similarity between images and tolerance to discrepancy between images. For this reason, it was intended to find the tolerance margin between segmentations made by human experts. This was made by means of an experimental study consisting on the comparison of the gold standards of the study set made by the two observers. The segmentations made by the first

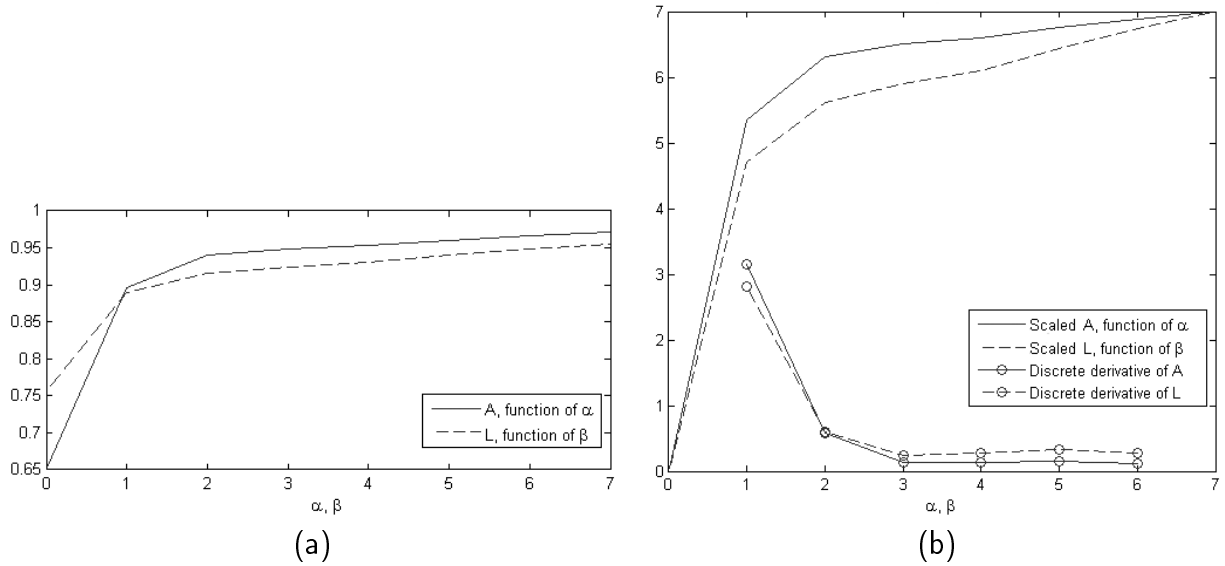


Figure 6.5: *Study for determining optimal α and β values. (a) Evolution of A and L mean values as functions of their α and β parameters, respectively, when comparing the gold standard sets by the two observers. (b) Two representations are shown: the evolution of scaled A and L to match their range and domain and the discrete derivative of A and L. As is can be checked, when α and β are equal to 2, discrete derivatives are the closer to 1.*

observer were considered as gold standard references and those made by the second observer were considered as the segmentations to be evaluated. In Figure 6.5, image (a) shows the evolution of the mean values of A and L on the gold standard images as functions of their α and β parameters, respectively. The aim is finding the minimum parameter values from which the A and L functions show no significant increase. With this purpose, the range and domain of A and L were matched (see Figure 6.5, image (b)) to check where the increase of the parameter value is closer to the increase of the function value. The selected value for each parameter was where the discrete derivative is closer to 1. As it can be checked in image (b) in Figure 6.5, it occurs for both functions for their parameter value 2. Therefore, for all the experimentation described in the rest of the chapter, the value for the parameters α and β was set to 2.

6.6.2 Sensibility to poor blood vessel segmentation

Let consider three sets of segmentations for the study set generated from the gold standards of the first observer. The first set is composed of segmentations with the 50% of the vascularity while the second set contains segmentations with the 25% of the vascular tree. For generating

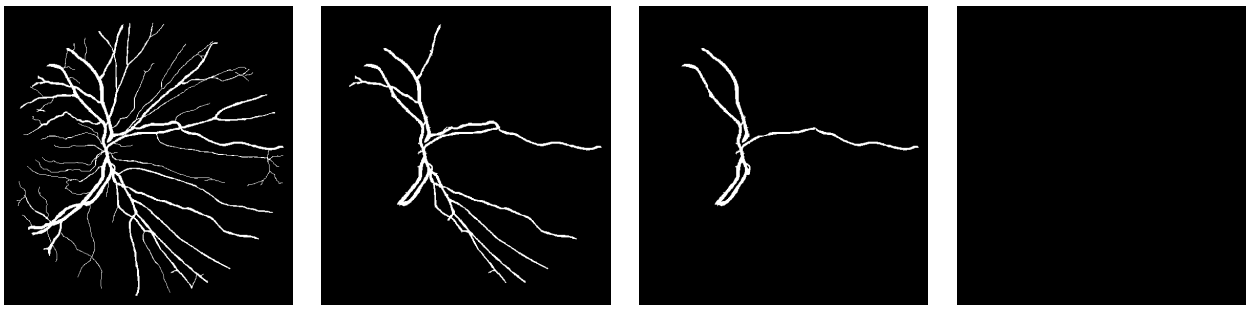


Figure 6.6: *Examples of enhanced sensitivity to poor blood vessel segmentation: (a) Reference gold standard. (b) Segmentation of (a) with* 50% *of detected vascularity:* $Acc = 0.9478$, $CAL = 0.2482$. *(c) Segmentation of (a) with* 25% *of detected vascularity:* $Acc = 0.9220$, $CAL = 0.0655$ *(d) Segmentation of (a) with* 0% *of detected vascularity:* $Acc = 0.8962$, $CAL = 0.0$.

these images the number of connected components was not increased. The last set contains null segmentations (i.e., segmentations in which all pixels were set as background). As Acc only takes into account the number of badly-classified pixels, the following average values of Acc, $\overline{Acc}$, for each set can be directly inferred using data of the first observer in Table 6.1:

- Set of 50%: $\overline{Acc} = (1 - (0.1274 \times 1/2)) = 0.9363$
- Set of 25%: $\overline{Acc} = (1 - (0.1274 \times 3/4)) = 0.9045$
- Set of 0%: $\overline{Acc} = (1 - 0.1274) = 0.8726$

On the other hand, CAL does not only evaluate the amount of classification errors but also their spatial configuration and, specifically, their impact on the three descriptive features of the vascular tree. That is why CAL values have to be calculated image by image and depend on the features of each image. Thus, the following average values were obtained measuring the quality of these segmentations using CAL with respect to their corresponding reference gold standards:

- Set of 50%: $\overline{CAL} = 1.0 \times 0.5032 \times 0.6693 = 0.3368$
- Set of 25%: $\overline{CAL} = 1.0 \times 0.2537 \times 0.3670 = 0.0931$
- Set of 0%: $\overline{CAL} = 0.9999 \times 0.0 \times 0.0 = 0.0$

Figure 6.6 shows an illustrative example. It can be observed that, although segmentations (b) and (c) compared to the reference gold standard (a) are very deficient, their obtained

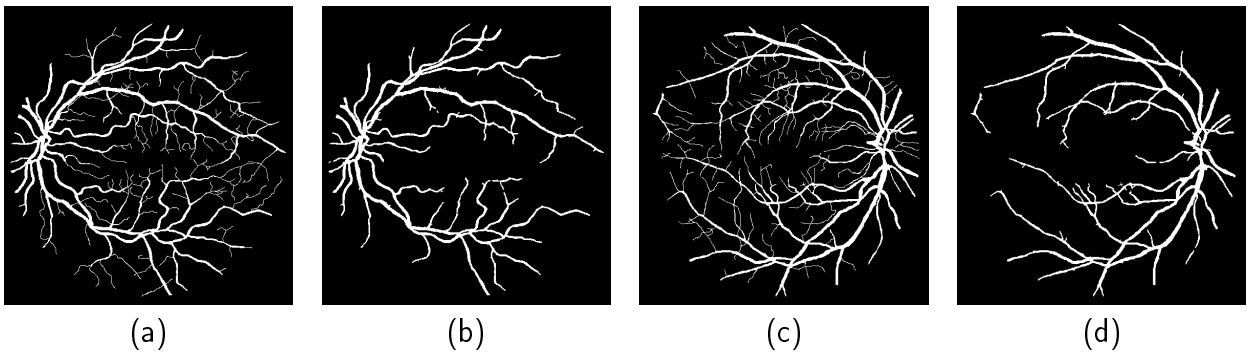


Figure 6.7: *Two examples of enhanced sensitivity to poor thin blood vessel segmentation: (a) Reference gold standard. (b) Segmentation of (a): Acc* $= 0.9749$*, CAL* $= 0.6085$*. (c) Reference gold standard. (d) Segmentation of (c): Acc* $= 0.9719$*, CAL* $= 0.5825$*.*

Acc value can be defined as excellent, as both obtained a score above 0.9. On the other hand, image (d), in spite of being a null segmentation representing an "absurd" case, obtained a quality score of almost 0.9. As results show, the behaviour of the new approach is closer to the expected behaviour, giving low values to images (b) and (c), while 0 was given to image (d).

6.6.3 Sensibility to poor thin blood vessel segmentation

To assess *Acc* and CAL behaviour against poor thin-vessel detection, let's consider a vessel segmentation set for the 20 images of the study set in which, for all images, no thin vessel was detected and all gross vessels were perfectly detected. Then, using the data of the first observer in Table 6.1, the following results of average *Acc* for this theoretical set can be inferred:

- $\overline{Acc} = (1 - 0.1274) + (0.7932 \times 0.1274) = 0.9736$

Now, using CAL, the average quality for this segmentation set, calculated image by image, was:

- $\overline{CAL} = 1.0 \times 0.8010 \times 0.7300 = 0.5847$

Note that, although no thin vessel pixel was detected, average *Acc* indicates nearly-perfect segmentation. This defect is much more penalized in the new proposal.

Figure 6.7 shows two examples of this study case. It can be checked that, although the loss of detection of all thin vessels involves an evident degradation in the visual perception of

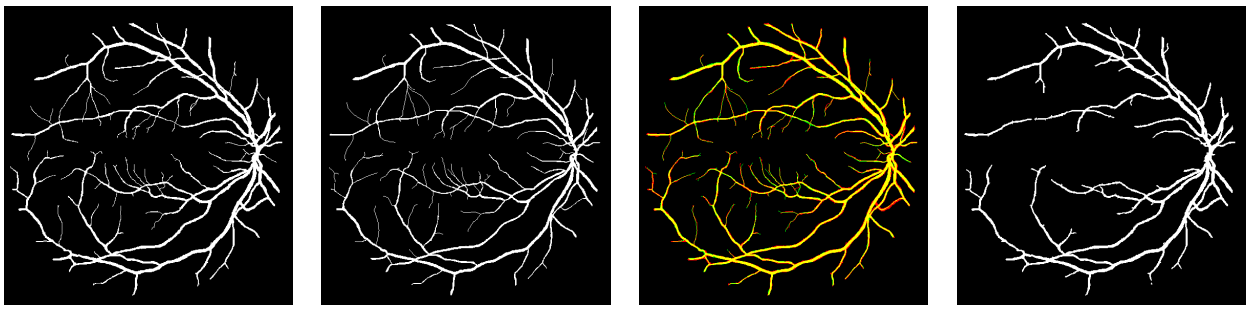


Figure 6.8: *Example of the impact of the reference gold standard used in the Acc and CAL. (a) Reference gold standard. (b) Gold standard by the second observer:* $Acc = 0.9689$, $CAL = 0.9711$. *(c) Color composition of (a) and (b). (d) Degradated version of image (a):* $Acc = 0.9888$, $CAL = 0.7297$. *Note that the Acc indicates that (d) is better segmentation than (b) due to its sensitivity to differences in vessels tracing and width. The CAL reduces this sensitivity and enhances quality evaluation.*

quality, the Acc is almost insensitive to this situation. This metric gives the following quality values for the images (b) and (d), respectively: 0.9749 and 0.9719. On the contrary, CAL behaviour is closer to what could be expected, detecting and penalizing this defect. The quality values for images (b) and (d) with the new QEF are 0.6085 and 0.5825, respectively.

6.6.4 Dependence with the gold standard

The influence of slight variations in vessel tracing or width on the Acc and CAL will be studied by means of an illustrative example. In Figure 6.8, image (a) is a reference gold standard from the study set made by the first observer, and image (b) is the gold standard for the same image made by the second observer. With the aim of studying only variations in vessel tracing or width, those vessels not originally present in both images were eliminated. Image (c) is the color composition of both images where the black color indicates coincidence in background, yellow pixels are coincidences of both segmentations in vessels, and the red and green show vessel disagreement; red pixels belong to the first observer, whereas the green ones belong to the second observer. It can be checked that all vessels are present in the images (a) and (b), the appreciable differences being due to differences in vessel tracing or width. Then, considering image (b) as a segmentation to be evaluated, and image (a) as the reference gold standard to be compared with, the following quality measures are obtained:

- $Acc = \frac{(20662+299034)}{(20662+7451+299034+2813)} = 0.9689$

- $CAL = 1.0 \times 0.9858 \times 0.9851 = 0.9711$

On the other hand, image (d) is a degradated version of image (a), generated by cutting and erasing many vessels. The quality values obtained for this image by comparison with the reference gold standard (a) are:

- $Acc = \frac{(24501+301761)}{(24501+3698+301761+0)} = 0.9888$
- $CAL = 1.0 \times 0.8763 \times 0.8327 = 0.7297$

As it can be checked, although any human observer would consider that image (b) is clearly better segmentation than image (d) when visually comparing with the reference gold standard (a), *Acc* precisely indicates the contrary. That is result of the sensitivity of this metric to the aforementioned variations. As results indicate, this sensitivity is reduced with CAL.

6.6.5 Degree of matching with human perception

To formally evaluate the consistency between both *Acc* and CAL and the subjective perception of quality, two approaches were used: the first is based on the *methodology of consistency assessment between perceived and computed matches of segmented images*, proposed by Paglieroni (2004), and the second one is a *study of statistical correlation between subjective and objective quality evaluation.*

Figure 6.9 shows the images used for this study, and contains two sets of images: the *set of synthetic images* (subfigure (a)) and the *set of real algorithm images* (subfigure (b)). The set of synthetic images is composed by the reference gold standard (G) and 5 segmentations created by modifying the reference image (images (S1)-(S5)). The set of real algorithm images is composed by the reference gold standard image (G), the gold standard image made by the second observer (image M1), and 7 segmentations generated by real algorithms for retinal vessel segmentation present in the literature (images M2-M8). Concretely, these images are from the following methods: Niemeijer et al. (2004) (image M2), Staal et al. (2004) (image M3), Zana and Klein (2001) (image M4), Soares et al. (2006) (image M5), Chaudhuri et al. (1989b) (image M6), Jiang and Mojon (2003) (image M7) and Martinez-Perez et al. (1999) (image M8). The images were downloaded from (http://www.isi.uu.nl/Research/Databases) except the images from the work by Soares et al. (2006), provided by the authors.

Subjective evaluation was carried out by asking 20 human observers to score the segmentations of both sets in terms of quality and similarity with respect to their corresponding

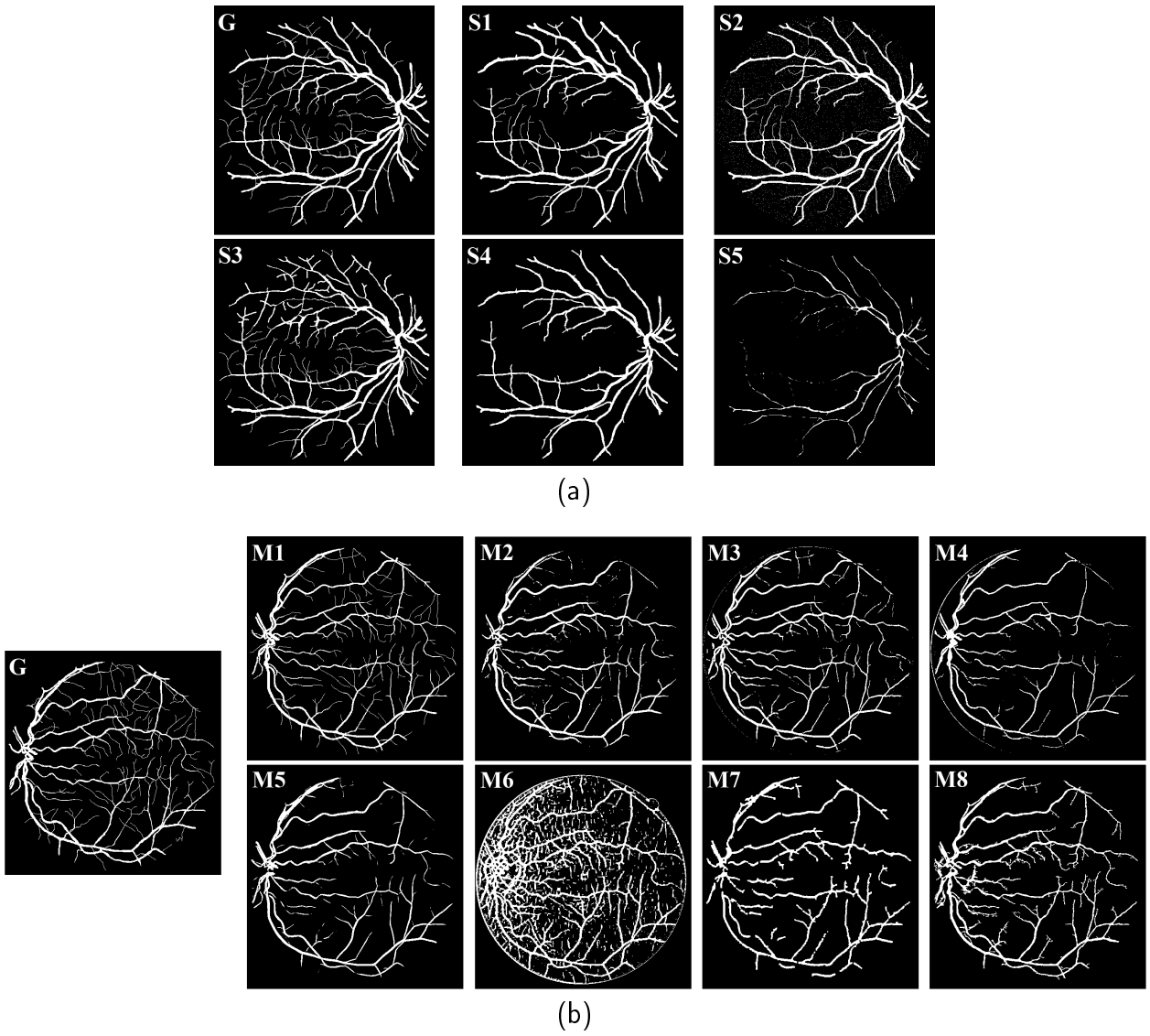


Figure 6.9: *Images used to study correspondence between subjective and objective quality evaluation. (a) Set of synthetic images: (G) is the reference gold standard of the set and images (S1)-(S5) are segmentations created by modifying (G). (b) Set of real algorithm images: (G) is the reference gold standard of the set, (M1) is the segmentation by the second observer, and (M2)-(M8) are automated segmentations generated by algorithms present in the literature.*

reference gold standard. Subjective scores are real numbers within the $[0, 10]$ interval, where 0 and 10 denote the worst and best quality cases, respectively. Assigning the same quality value for different images was not allowed. Once the observers' scores were collected, outliers were eliminated (scores with at least one component out of the average sample plus/minus 2 times the standard deviation). On the other hand, Acc and CAL values were also calculated for the segmentations of the two sets. Table 6.2 shows average of human observer scores

	Set of Synthetic Images				
	S1	S2	S3	S4	S5
Acc	0.9745	0.9746	0.9747	0.9767	0.8888
CAL	0.8240	0.6468	0.9998	0.6590	0.4075
$\overline{HO}$	0.7143	0.5357	0.8786	0.5500	0.3143
HO_σ	0.0770	0.1336	0.0469	0.1109	0.1231

	Set of Real Algorithm Images							
	M1	M2	M3	M4	M5	M6	M7	M8
Acc	0.9468	0.9389	0.9434	0.9337	0.9439	0.8039	0.9189	0.9313
CAL	0.8575	0.6638	0.7148	0.5310	0.6571	0.4009	0.6002	0.6587
$\overline{HO}$	0.8923	0.5662	0.6169	0.3685	0.5308	0.0962	0.3077	0.6231
HO_σ	0.0641	0.1003	0.1010	0.0534	0.0751	0.0776	0.0760	0.1641

Table 6.2: *Results of objective and subjective quality evaluation. Objective evaluation is in terms of Acc and CAL. Subjective evaluation is in terms of average of human observer scores ($\overline{HO}$) and standard deviation of human observer scores (HO_σ).*

($\overline{HO}$) and standard deviation of human observer scores (HO_σ). The CAL and Acc values per image are also shown.

The first methodology used in this work, based on the one proposed by Paglieroni (2004), consists on obtaining a disparity value between perceived quality (provided by human observers) and computed quality (provided by a QEF). The methodology can be summarized as follows:

- Given n human observers and c segmentations, a set of $P_i = (P_{i_1}, \ldots, P_{i_c})$ with $i = 1, \ldots, n$, vectors is built, where P_{i_j} is the score of the ith human observer for image j.

- From the P_i vectors, a new set of $r_i = (r_{i_1}, \ldots, r_{i_c})$ vectors called *individual ranking profiles* is created, where

$$r_{i_k} = \#\left\{P_{i_j} \in P_i \mid P_{i_j} \leq P_{i_k}, j = 1, \ldots, c\right\} \tag{6.7}$$

  So, in r_i vectors, components take integer values from 1 to c according to the order provided by P_i.

- From the r_i vectors, a *mean consensus ranking profile* vector $\overline{R} = (\overline{R}_1, \ldots, \overline{R}_c)$ is then obtained, where $\overline{R_i} = \frac{1}{n}\sum_{j=1}^{n} r_{j_i}$. So, $\overline{R}$ is a real-valued vector obtained by computing the average of the ranked scores of the n human observers for each of the c segmentations.

- A *standard consensus ranking profile* vector $R = (R_1, \ldots, R_c)$ is also obtained, where

$$R_i = \#\left\{R_j \in \overline{R} \mid R_j \leq \overline{R}_i, j = 1, \ldots, c\right\} \tag{6.8}$$

  So, R is a vector of components that take integer values from 1 to c according to the order provided by $\overline{R}$.

- Now, given a QEF and the c segmentations, a $\hat{D} = (\hat{D}_1, \ldots, \hat{D}_c)$ vector is constructed, where $\hat{D}_i$ is the score of the QEF for the ith image.

- From the vector $\hat{D}$, a $\tilde{D} = (\tilde{D}_1, \ldots, \tilde{D}_c)$ vector called *standard computer-generated ranking profile* is constructed, where

$$\tilde{D}_i = \left[\frac{\hat{D}_i - a}{b - a} \times (c - 1)\right] + 1 \tag{6.9}$$

  being $a = min\left\{\hat{D}_i, i = 1, \ldots, c\right\}$ and $b = max\left\{\hat{D}_i, i = 1, \ldots, c\right\}$. So, $\tilde{D}$ is a real-valued vector obtained by normalizing computed quality scores to the $[1, c]$ interval.

- A *standard computer-generated ranking profile* vector $D = (D_1, \ldots, D_c)$ is also calculated, where

$$D_i = \#\left\{D_j \in \tilde{D} \mid D_j \leq \tilde{D}_i, j = 1, \ldots, c\right\} \tag{6.10}$$

  So, D is a vector of integer-valued components that take values from 1 to c according to the order provided by $\tilde{D}$.

- Finally, two measures of inconsistency Δ_R between perceived and computed disparity are computed: the *mean inconsistency* Δ_{R1} and the *standard inconsistency* Δ_{R2}. They are defined as:

$$\Delta_{R1} \triangleq \left\|\overline{R} - \tilde{D}\right\| / c \quad \in [0, c - 1] \tag{6.11}$$

$$\Delta_{R2} \triangleq \frac{1}{c}\sum_{i=0}^{c-1} |R(i) - D(i)| \quad \in [0, c - 1] \tag{6.12}$$

  Both inconsistency measures reflect rank disparity in units which range from 0 to $c-1$ (note that value 0 denotes perfect matching between perceived and computed quality). Δ_{R2} is only sensitive to changes in the order of the compared scores. However, Δ_{R1} is also sensitive to the real distance between the compared scores. Therefore, this

	Set of Synthetic Images			Set of Real Algorithm Images		
	Δ_{R1}	Δ_{R2}	ρ	Δ_{R1}	Δ_{R2}	ρ
Acc	**0.7067**	**1.20**	**0.7442**	**0.9716**	**1.0**	**0.7688**
CAL	0.0387	0.0	0.9999	0.2656	**1.0**	0.9657

Table 6.3: *Matching degree with human perception for CAL and Acc in terms of* Δ_{R1}, Δ_{R2} *and* ρ.

measure provides more sensitivity to differences than Δ_{R2}. As the author claims in his paper (Paglieroni, 2004), "the real-valued ranking profiles used by mean inconsistency (Δ_{R1}) can capture more subtle differences between rankings than the integer-valued ranking profiles used by standard inconsistency (Δ_{R2})".

The second methodology used in this work is a study of statistical correlation between subjective and objective quality evaluation. This study was made by considering the two random variables $\overline{HO}$ and S_{QEF}. $\overline{HO}$ denotes, as defined before, the average of human observer scores and S_{QEF} denotes QEF-generated scores. For c segmentations, the c observations of each variable are obtained and the correlation of variables can be calculated as:

$$\rho = \frac{\sigma_{\overline{HO},S_{QEF}}}{\sigma_{\overline{HO}}\ \sigma_{S_{QEF}}}$$

where σ denotes de standard deviation and $\sigma_{\overline{HO},S_{QEF}}$ is the covariance of $\overline{HO}$ and S_{QEF}. Correlation ρ measures the linear relationship of both variables which is, in this case, an indicator of the similarity degree of the behaviour of subjective and objective evaluation.

Table 6.3 shows the results of this experiment for *Acc* and *CAL*. As it can be checked, *CAL* provides the best correspondence with human perception. It is remarkable that the ρ value for *CAL* reaches 0.9999 in the case of the set of synthetic images, thus indicating almost complete linear relationship between the behaviour of *CAL* and subjective evaluation. This value is also high for the set of real algorithm images, 0.9657. The results obtained by *CAL* are better than those obtained by *Acc*, except with Δ_{R2} in the case of the set of real algorithm images. However, this is due to the fact that Δ_{R2} is only sensitive to changes in the order of the compared scores, not to the real distance between them.

6.7 Discussion on the application of other generalist quality evaluation functions

Other generalist QEFs which, to the best of the author's knowledge, have not been used in the topic of retinal vessel segmentation were evaluated. This evaluation was in terms of correspondence with human perception, so the methodology and data presented in section were also used for this study.

The QEFs evaluated were the *Jaccard coefficient* (Jaccard, 1901), the *Dice coefficient* (Dice, 1945), the *average symmetric contour distance*, the *root mean square symmetric contour distance*, and the *maximum symmetric contour distance*. The three last mentioned QEFs are 2-D adaptations of the *average symmetric surface distance* (Lamecker et al., 2004), the *root mean square symmetric surface distance* (Lamecker et al., 2004), and the *maximum symmetric surface distance* (Huttenlocher et al., 1993) respectively, as they were originally defined for quality evaluation of 3-D segmentations.

For all the cited QEFs two versions were used. The first version corresponds to the implementation of the original functions as they were proposed by their corresponding authors. The second version was made with the aim of comparing the mentioned QEFs with CAL under the same conditions. This version proposes modifications of the original formulations to introduce tolerance on spatial overlap. This tolerance was implemented following the same approach used in CAL, i.e., by performing morphological dilations on the segmentations to be compared. The description of each quality evaluation function is what follows:

1. Jaccard coefficient: the Jaccard coefficient for a segmentation S and a reference gold standard $S_{_G}$ is defined as the ratio between intersection and union of both images:

$$J(S, S_{_G}) = \frac{\#(S \cap S_{_G})}{\#(S \cup S_{_G})} \tag{6.13}$$

The modified version of this coefficient for introducing tolerance on spatial overlap was defined as:

$$J_{_\gamma}(S, S_{_G}) = \frac{\#((\delta_{_\gamma}(S) \cap S_{_G}) \cup (S \cap \delta_{_\gamma}(S_{_G})))}{\#(S \cup S_{_G})} \tag{6.14}$$

where $\delta_{_\gamma}$ is a morphological dilation using a disc of γ pixels in radius. This operator is introduced in the modified versions of all above-described QEFs.

2. Dice coefficient: the Dice coefficient for a segmentation S and a reference gold standard S_G is defined as the size of the intersection of the two images divided by their average size:

$$D(S, S_G) = \frac{\#(S \cap S_G)}{\frac{1}{2}(\#(S) + \#(S_G))} \tag{6.15}$$

The alternative version of this coefficient including tolerance on spatial overlap is:

$$D_\gamma(S, S_G) = \frac{\#((\delta_\gamma(S) \cap S_G) \cup (S \cap \delta_\gamma(S_G)))}{\frac{1}{2}(\#(S) + \#(S_G))} \tag{6.16}$$

3. Average Symmetric Contour Distance: the average symmetric contour distance is based on the frontier pixels of the segmentation S and reference gold standard S_G. Frontier pixels are defined by having at least one non-vessel pixel within their 8-neighborhood. For each frontier pixel of S, the Euclidean distance to the closest frontier pixel of S_G is calculated. In order to provide symmetry, the same process is applied from the frontier pixels of S to S_G. The average symmetric contour distance is then defined as the average of all stored distances, which is 0 for perfect segmentation.

Let $P(S)$ denote the set of frontier pixels of S. The shortest distance of an arbitrary pixel p to $P(S)$ is defined as $d(p, P(S)) = \min_{p_S \in P(S)} \|p - p_S\|$, where $\|.\|$ denotes the Euclidean distance. The average symmetric contour distance is then given by:

$$\begin{aligned} ASD(S, S_G) = & \frac{1}{|P(S)| + |P(S_G)|} \times \\ & \times \left(\sum_{p_S \in P(S)} d(p_S, P(S_G)) + \sum_{p_{S_G} \in P(S_G)} d(p_{S_G}, P(S)) \right) \end{aligned} \tag{6.17}$$

and its alternative version is:

$$\begin{aligned} ASD_\gamma(S, S_G) = & \frac{1}{|P(S)| + |P(S_G)|} \times \\ & \times \left(\sum_{p_S \in P(S)} d(p_S, P(\delta_\gamma(S_G))) + \sum_{p_{S_G} \in P(S_G)} d(p_{S_G}, P(\delta_\gamma(S))) \right) \end{aligned} \tag{6.18}$$

4. Root Mean Square Symmetric Contour Distance: the root mean square symmetric contour distance is also based on contour distances. It is calculated as the average symmetric contour distance described above, except that Euclidean distances between frontier pixels are squared before storing them. The root of averaged squared distances then yields the root mean square symmetric contour distance, which is 0 for a perfect segmentation:

$$RMSD(S, S_G) = \sqrt{\frac{1}{|P(S)| + |P(S_G)|}} \times \\ \times \sqrt{\sum_{p_S \in P(S)} d^2(p_S, P(S_G)) + \sum_{p_{S_G} \in P(S_G)} d^2(p_{S_G}, P(S))} \tag{6.19}$$

The modified version proposed for this function is:

$$RMSD_\gamma(S, S_G) = \sqrt{\frac{1}{|P(S)| + |P(S_G)|}} \times \\ \times \sqrt{\sum_{p_S \in P(S)} d^2(p_S, P(\delta_\gamma(S_G))) + \sum_{p_{S_G} \in P(S_G)} d^2(p_{S_G}, P(\delta_\gamma(S)))} \tag{6.20}$$

5. Maximum Symmetric Contour Distance: the maximum symmetric contour distance, also known as Hausdorff distance, is determined similarly to the previous two QEFs. Differences between both sets of frontier pixels are determined using Euclidean distances, and the maximum value yields the maximum symmetric contour distance. The distance for a perfect segmentation is 0. This function is formally defined as:

$$MSD(S, S_G) = \max \left\{ \max_{p_S \in P(S)} d(p_S, P(S_G)), \max_{p_{S_G} \in P(S_G)} d(p_{S_G}, P(S)) \right\} \tag{6.21}$$

and its modified version is defined as follows:

$$MSD_\gamma(S, S_G) = \max \left\{ \max_{p_S \in P(S)} d(p_S, P(\delta_\gamma(S_G))), \max_{p_{S_G} \in P(S_G)} d(p_{S_G}, P(\delta_\gamma(S))) \right\} \tag{6.22}$$

	Acc	CAL	J	J_2	D	D_2	ASD	ASD_2	$RMSD$	$RMSD_2$	MSD	MSD_2
Set of Synthetic Images												
Δ_{R1}	0.7067	**0.0387**	0.7244	0.2061	0.7246	0.9085	0.3631	0.4038	0.5412	0.5949	0.6727	0.6727
Δ_{R2}	1.20	**0.0**	0.80	0.40	0.80	1.60	0.80	0.80	0.80	0.80	0.40	0.40
ρ	0.7442	**0.9997**	0.7517	0.9753	0.7516	0.1003	0.8622	0.8662	0.8001	0.7863	0.7668	0.7668
Set of Real Algorithm Images												
Δ_{R1}	0.9716	**0.2656**	0.5943	0.4418	0.660	0.4419	0.5376	0.5942	0.6662	0.6783	0.8069	0.8069
Δ_{R2}	1.0	1.0	**0.750**	**0.750**	**0.750**	1.250	1.250	1.250	1.250	1.250	1.750	1.750
ρ	0.7688	**0.9657**	0.9117	0.9361	0.8949	0.9039	0.8759	0.8614	0.7628	0.7611	0.5283	0.5283

* Note that low Δ_{R1} and Δ_{R2} values and high ρ values denote better matching.

Table 6.4: *Matching degree with human perception for all the considered QEFs in terms of* Δ_{R1}, Δ_{R2} *and* ρ.

Table 6.4 shows the values of Δ_{R1}, Δ_{R2} and ρ for all the considered QEFs. For the modified versions of these QEFs, the γ value was fixed to 2, since this value was obtained as the most appropriate for CAL in section 6.6.1. As it can be observed, CAL provides the best correspondence with human perception in most of cases. The value of Δ_{R1} for the CAL is the lowest for the two sets of images. With respect to correlation ρ, CAL is also the best correlated with human perception of all approaches. Note also that the second best correlated QEF is J_2, which exactly corresponds to factor A of CAL. Figure 6.10 shows the representation *Image Vs Score* for the average of human observer scores $\overline{HO}$, CAL, *Acc* and J_2. The similarity of the behaviours of the $\overline{HO}$ and CAL justifies the high values of correlation obtained. Observe that the Δ_{R2} value, in the case of the set of real algorithm images, for three of the compared QEFs is lower than for CAL. However, it may be a consequence of the previously commented poor sensitivity of Δ_{R2} to the real distance between the compared scores.

Taking into account the most reliable matching degree indicators, Δ_{R1} and ρ, it is important to point out that, as shown in Table 6.4, *Acc* shows worse behaviour than, not only CAL, but also than most of the other considered generalist QEFs.

6.8 Discussion and conclusion

In the context of image segmentation, measurement of segmentation quality is usually based on the use of generalist QEFs as *accuracy*, *Jaccard coefficient*, etc. However, in the author's opinion, evaluation of segmentation quality may be complemented or even improved with

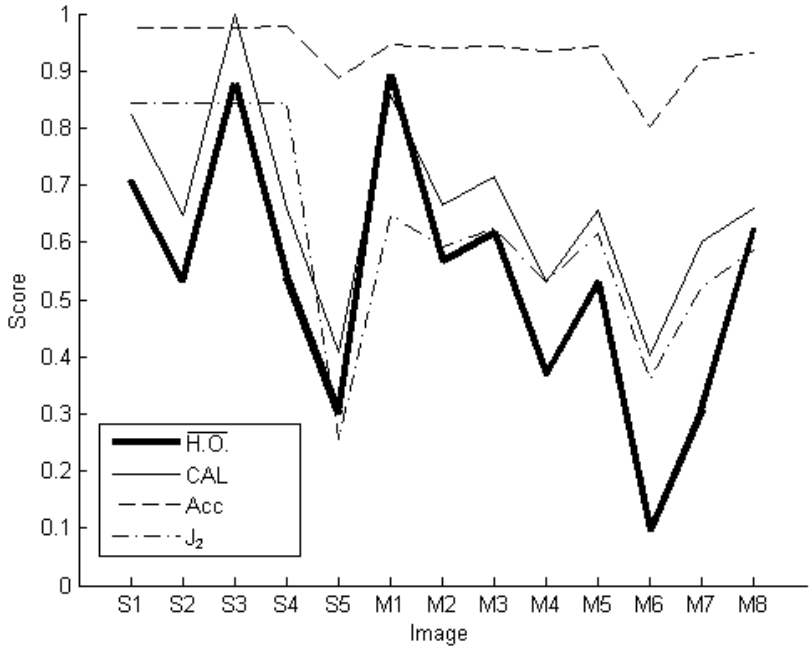


Figure 6.10: *Image Vs Score representation for the $\overline{HO}$, CAL, Acc and J_2.*

the use of specialized or ad-hoc QEFs. This is especially interesting when the target to be segmented has a characteristic structure.

This chapter presents an experimental study which shows the advantages of the use of a specialized QEF within the frame of retinal vessel segmentation. This work was motivated by the author interest in the field of retinal vessel segmentation. Working on the development of segmentation algorithms, the author realized that some perceptible improvements had very low incidence on the *Acc* values and that these values did not match what they expected according to their subjective perception. These difficulties suggested that *Acc* could be complemented with other specialized QEFs for assessing quality of retinal vasculature segmentation.

The vascular tree is a tree-like connected structure that may be characterized by its connectivity, area and total length. In this chapter, these three features have been used for designing a new specialized QEF, known as CAL. Besides providing a quality measurement, this kind of specialized QEFs provides higher result interpretability. In this case, CAL can be used to detect concrete features (connectivity, area or length) of the vascular tree inappropriately segmented.

In order to compare CAL and *Acc* behaviour, a detailed study was included in this chapter. Firstly, the weaknesses of *Acc* were characterized and a set of examples where *Acc* behaviour

is deficient was provided. However, in these cases, CAL presented a more appropriate performance. Moreover, the performance of other generalist QEFs was also studied and compared. In this sense, two sets of segmentations, generated synthetically as well as by means of segmentation algorithms in literature, were considered. These segmentations were scored by a set of human observers in terms of perceived quality and by the considered QEFs, including CAL and *Acc*. Finally, a study of correspondence between both subjective and objective evaluation was carried out. The final conclusion of this study is that CAL shows higher matching with the human quality perception than the other tested QEFs. Interestingly, although *Acc* is the most widely-used QEF in the topic of retinal vascular tree segmentation, it leads to worse performance than some of the other considered generalist QEFs. Therefore, due to the foregoing and in our opinion, the CAL may be used as a good supplement of the information provided by *Acc*.

Note that CAL has been designed for measuring segmentation quality of a concrete structure and, therefore, it is not applicable in general cases. However, in general, the applied concept of measuring descriptive features may be useful for designing other specialized QEFs aimed at enhancing segmentation quality assessment of other complex shapes.

Chapter 7

Conclusion and future work

DIABETIC retinopathy (DR) is a chronic retinal disease affecting diabetic population, being this population forecasted to rise from 171 million in 2000 to 366 million in 2030. It is estimated that about 2% of diabetic patients are blind and 10% undergo vision degradation after 15 years of diabetes due to DR. Laser photocoagulation is nowadays the main early therapeutic treatment used to slow down vision degradation in DR patients, but it is not effective if it is not applied when DR is in its early stages. Therefore, the early detection of the illness is the key factor in the preservation of vision in affected patients. Since DR patients usually perceive symptoms of vision loss when the illness is in advance stages, this is, when treatment is less effective, diabetic patients are annually examined in their public health systems. However, the examination of such a huge population involves a huge challenge for these Administrations and many patients are usually prevented from receiving adequate treatment. In this context, a system which could be used by non-experts to discriminate and filtrate with higher confidence those cases that require referral to an ophthalmologist for further investigation and treatment and those that do not, would reduce the workload of these specialists, would improve preventive protocols and therefore, would decrease costs too. Although some prototypes of such a system have been presented in the recent years, none of them have reached a quality level enough to be used in the clinical practice.

This thesis is framed within a research project for the development of a system for the automated detection of DR by means of the automated analysis of eye fundus images. The construction of such a system can be structurally divided into four modules. These modules are those for: quality assessment, segmentation of the main anatomical components, detec-

tion of lesions and diagnosis of DR. The investigations carried out to develop the module for the segmentation of the main anatomical components have been the central goal of this thesis. This module comprises the segmentation of the optic disc (OD) and vascular tree, and the identification of the macular zone. Thus, reliable and robust automated methodologies for the recognition of these retinal features have been investigated and developed.

Concerning the proposed optic disc (OD) segmentation methodology, it provides a tool for obtaining OD boundary approximations in retinographies from healthy patients as well as from patients affected by DR. This methodology is based on morphological and edge detection techniques as well as uses the Circular Hough Transform so as to obtain a circular OD boundary approximation. The segmentation algorithm starts extracting an OD-containing subimage defined by a pixel located within the OD and a surrounding region wide enough to include it completely. To this effect, a reliable and very fast OD location methodology is also proposed in this thesis. The segmentation methodology was subjectively and objectively evaluated on the MESSIDOR database containing 1200 images with available DR diagnoses for all images (see database description in appendix A). Subjective evaluation, carried out by medical experts, concluded that the methodology obtained a good OD boundary approximation in 94.58% of cases, which means segmentation success in 1135 out of 1200 images. Objective evaluation was performed by measuring the common area overlapping between automated segmentations and true OD regions hand-delimited by medical experts. The methodology rendered an average common area overlapping of 86% measured on the whole database. Attending to results partially evaluated according to the available diagnoses per image, they lead to the same conclusion for both subjective and objective evaluations: since success rate did not show a decreasing pattern related to illness presence or to illness condition, algorithm performance is neither sensitive to illness presence or absence nor dependent to illness condition when it is present. Regarding computational efficiency, the proposed segmentation methodology also constitutes a fast segmentation solution since, for a total of 1200 executions of the algorithm, the average computational time obtained was 5.69 sec. with a standard deviation of 0.54 sec.

With respect to the methodology for blood vessel segmentation, it is a new supervised approach that uses a neural network scheme for pixel classification. This methodology uses a neural network scheme for pixel classification and computes a 7-dimensional vector composed of gray-level and moment invariants-based features for pixel representation. This

feature vector is computed from preprocessed retinal images in the neighborhood of the pixel under consideration and given as input to a neural network. Classification results (real values between 0 and 1) are thresholded to classify each pixel into two classes: vessel and non-vessel. Finally, a postprocessing fills pixel gaps in detected blood vessels and removes falsely-detected isolated vessel pixels. The proposed segmentation method was evaluated and compared to the most relevant segmentation methods currently available in the literature. In order to compare the new methodology to these methods, the *accuracy* and the *area under ROC curve* (these metrics are defined in appendix B) were used as performance measures. Since these measurements were used by the authors of the considered works, this choice facilities comparing the results. The *accuracy* and *area under ROC curve* values rendered by the presented segmentation methodology calculated jointly on the DRIVE and STARE database were 0.9621 and 0.9678. These results outperform all analysed and compared segmentation approaches. With respect to computational complexity, the measured computational time obtained was less than approximately 90 seconds per image. This figure, in spite of the performed implementation is experimental, indicates that the new proposal is also an attractive solution from a computational point of view.

On the other hand, the presented macular zone identification methodology offers a choice for automatically delimiting the macula in digital color retinographies by means of the obtention of its center, i.e., the fovea center. The methodology, firstly, calculates a fovea center estimate by exploiting its a priori-known positional features with respect to the OD and vascular tree. To this effect, the proposed OD and vascular tree segmentation methodologies are used to estimate the raphe of the retina, which is a line separating the superior and inferior regions of the retina determined by the superior and inferior vessels. Then, the fovea center is estimated to be at 2.5 OD diameters from the OD center following the raphe. Once a fovea center estimate is obtained, secondly, this location is refined if the fovea is detectable using morphological processing. This processing is applied in a fovea-containing subimage centered on the estimated position in order to obtain an accurate fovea center position. By means of this morphological processing, the fovea is segmented if it is detectable and, in this case, it is evaluated in terms of circularity and size. If the segmented fovea satisfies both criteria it is validated and an accurate fovea center position is obtained. A final fovea center position is decided as the accurate position if it was possible to be obtained or, otherwise, as the fovea center estimate. Finally, the macular zone is identified using a polar coordi-

nate system centered on the final fovea center location. The methodology was tested on the MESSIDOR database using a distance criterion between the obtained fovea center and the real fovea center. The defined criterion establishes four quality categories according to the distance between the obtained fovea center and the real fovea center: *Excellent*, *Good*, *Fair* and *Poor*. The methodology obtained fovea centers in the interval *Excellent* - *Fair* (fovea centers widely accepted as valid in the literature) in 97.42% of cases. Furthermore, 88.55% of obtained fovea centers were in the interval *Excellent* - *Good*. These results outperform all compared approaches. Regarding computational complexity, the algorithm takes in average 10.88 sec for the whole process, excluding from this measure the time needed for OD and vascular tree segmentation. This figure makes the presented algorithm, besides a high accurate approach, a valid solution from a computational point of view.

In addition to these new methodologies for the segmentation of the main anatomical components, a new function for evaluating the global quality of retinal vessel segmentations has also been proposed. Metrics based on contingency tables for binary classification have been widely used for evaluating quality of these segmentations. Particularly, *accuracy* has been mostly used as a measure of global performance of vessel segmentation algorithms and, therefore, has been mostly used to compare them in the literature. However, when using this metric during researching on the presented vessel segmentation methodology, some weaknesses of this metric were identified. These difficulties lead to poor matching between the *accuracy* evaluation and the subjective perception of quality. The quality evaluation function presented in this thesis is based on the evaluation of three main measurable features which describe vascularity: vascular tree connectivity, area and length. This quality evaluation function represents a more natural approach for evaluating vascular segmentations and, since it is based on evaluating descriptive features of the vascular tree, it provides higher result interpretability to easily detect concrete features (connectivity, area or length) of the vascular tree inappropriately segmented.

Therefore, new high accurate proposals for the segmentation of the optic disc, vascular tree and macular zone have been presented. Since all of them outperform all analysed approaches present in the literature, this thesis supposes a step forward in the development of systems for the early automated detection of DR.

The future work regarding this thesis is focused on the detection of lesions produced by DR. In fact, this work is being currently carried out in our laboratory within the frame of the

second phase of the above mentioned project funded by the Health Ministry of the Andalusian Regional Government (Andalusia, Spain). Investigations for the accurate detection of microaneurysms, haemorrhages and exudates are being performed with the aim of obtaining at the end of 2012 a comprehensive system for the processing and analysis of retinal images. At that point, it is planned that an expert system for the classification of the illness will be implemented and a pilot experience will be performed in some primary centers of the Andalusian Health Service for testing and validation.

Appendix A

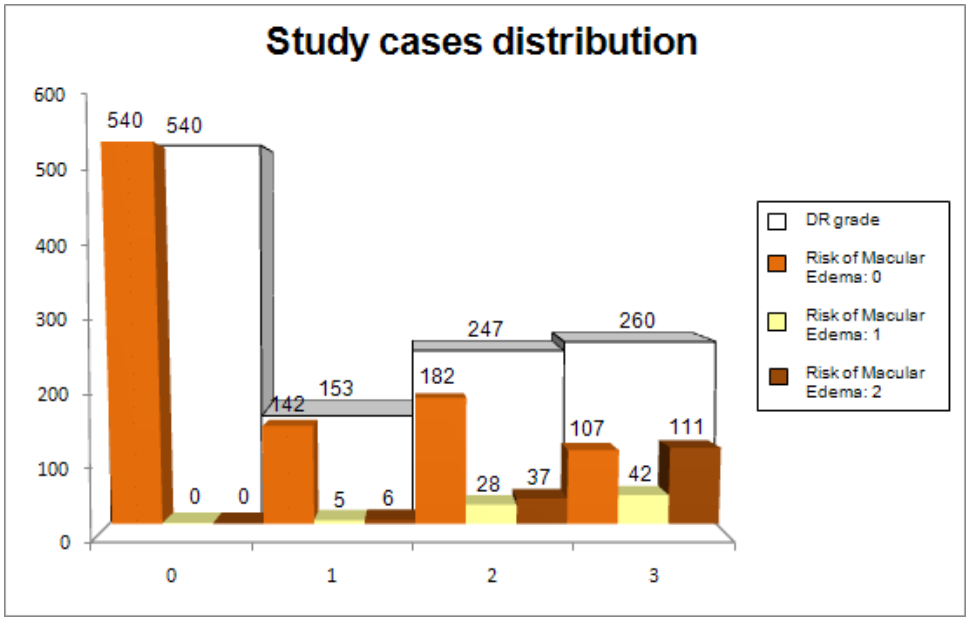
Digital Retinal Image Databases

This appendix provides the technical description of all the image databases used in this thesis. The appendix is divided into **four sections**, each of them describing a database.

A.1 The MESSIDOR database

The MESSIDOR database[1] contains 1200 eye fundus color images of the posterior pole acquired by the Hôpital Lariboisière Paris, the Faculté de Médecine St. Etienne and the LaTIM - CHU de Brest (France). 800 of these images were captured with pupil dilation (one drop of Tropicamide at 10%) and 400 without dilation, using a Topcon TRC NW6 non-mydriatic retinograph with a 45° FOV. The images are 1440 × 960, 2240 × 1488 or 2304 × 1536 pixels in size and 8 bits per color plane and are provided in TIFF format. Diabetic retinopathy (DR) grade and risk of macular edema (ME) diagnoses were provided by medical experts for each image according to the gradings described in section 2.2 (see Figure A.1 for description of study case distribution). Some examples of images in the MESSIDOR database can be checked in Figure A.2.

[1]Download images section, MESSIDOR: Digital Retinal Images, MESSIDOR TECHNO-VISION Project, France [Online]:`http://messidor.crihan.fr/download-en.php`



(a)

Figure A.1: *Distribution of the study cases grouped by grade of DR and risk of ME.*

(a) (b) (c)

Figure A.2: *Examples of retinographies in the MESSIDOR database: (a) Image from a healthy patient. (b) Image from a patient with DR grade* 3 *and no risk of ME. (c) Image from a patient with DR grade* 3 *and risk* 2 *of ME.*

A.2 The DRIVE database

The DRIVE database[2] comprises 40 retinal color images taken with a Canon CR5 non-mydriatic 3CCD camera with a 45° field-of-view (FOV). Each image was captured at 768 × 584 pixels, 8 bits per color plane and, in spite of being offered in LZW compressed TIFF

[2]Research section, digital retinal image for vessel extraction (DRIVE) database University Medical Center Utrecht, Image Sciences Institute, Utrecht, The Netherlands [Online]:`http://www.isi.uu.nl/Re-search/Databases/DRIVE`

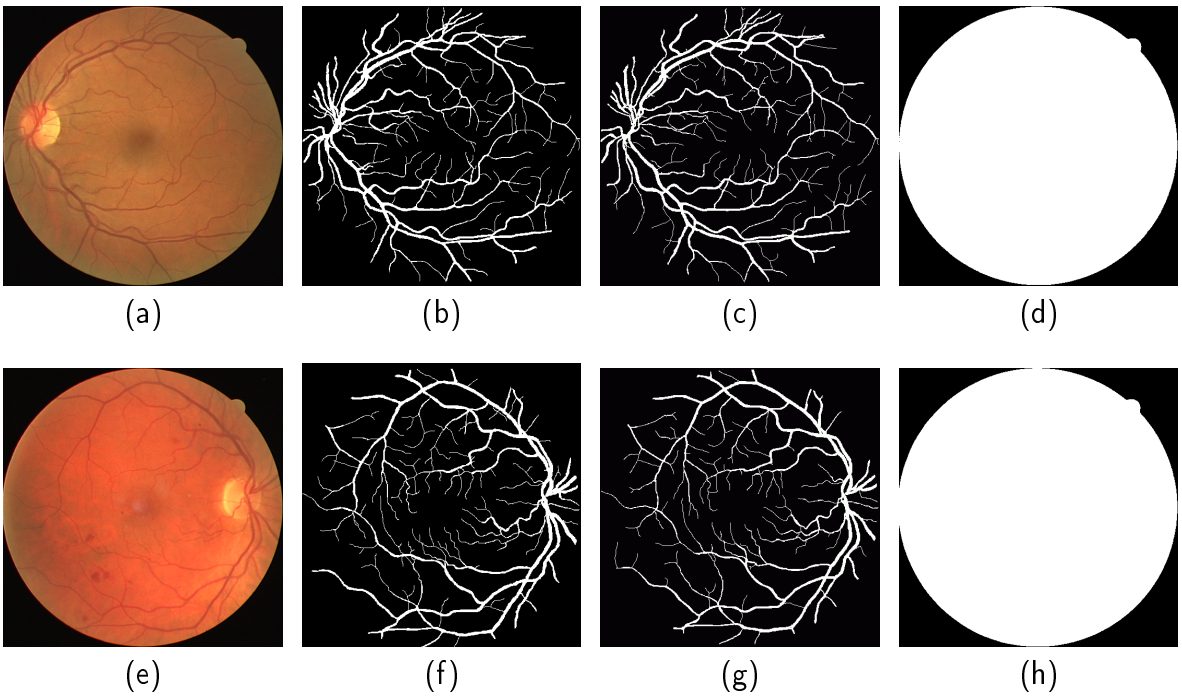


Figure A.3: *Examples of the material included in the DRIVE database: (a) Image from a healthy patient. (b) Gold standard segmentation of (a) made by the first observer. (c) Gold standard segmentation of (a) made by the second observer. (d) FOV mask for image (a). (e) Image from a patient affected by DR. (f) Gold standard segmentation of (e) made by the first observer. (g) Gold standard segmentation of (e) made by the second observer. (h) FOV mask for image (e).*

format, they were originally saved in JPEG format. The database is divided into two sets: a test set and a training set, each of them containing 20 images. The test set provides the corresponding FOV masks for the images, which are circular (approximated diameter of 540 pixels) and two manual segmentations of the vascular tree generated by two different specialists for each image. The training set also includes the FOV masks for the images and a set of manual segmentations of the vascular tree made by the first observer. Of the 40 images in the database, 7 contain pathology and the rest are from healthy patients. Examples of the material included in this database are shown in Figure A.3.

A.3 The STARE database

The STARE database[3] comprises 20 retinal color images captured with a TopCon TRV-50 fundus camera at 35° FOV. The images were digitalized to 700 × 605 pixels, 8 bits per

[3]Structured Analysis of the Retina (STARE), College of Engineering and Science, CLEMSON University, Clemson, U.S.A. [Online]. Available: `http://www.ces.clemson.edu/~ahoover/stare`

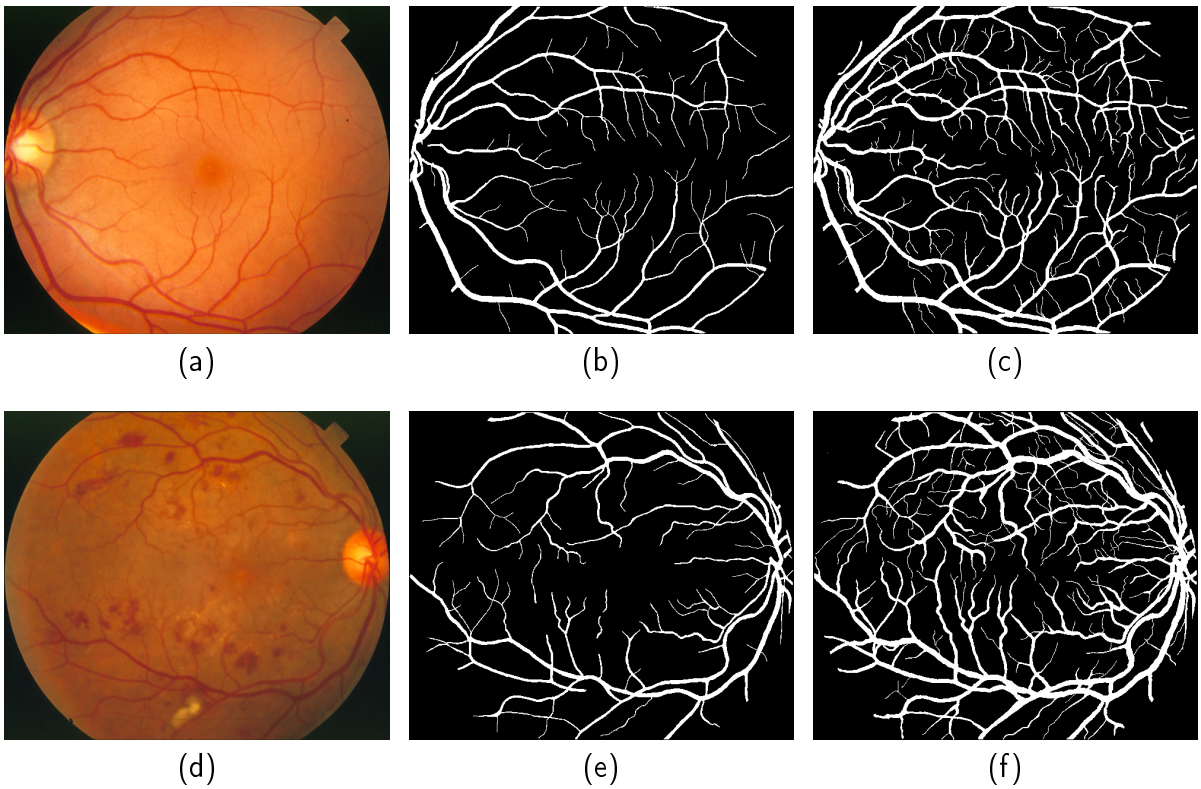


Figure A.4: *Examples of the material included in the STARE database: (a) Image from a healthy patient. (b) Gold standard segmentation of (a) made by the first observer. (c) Gold standard segmentation of (a) made by the second observer. (d) Image from a patient affected by DR. (e) Gold standard segmentation of (d) made by the first observer. (f) Gold standard segmentation of (d) made by the second observer.*

color channel and are available in PPM format. The database contains two sets of manual segmentations of the vascular tree made by two different observers. 10 of the images are of healthy patients and the other 10 images contain pathology. Figure A.4 shows examples of the described material.

A.4 The LOWELL database

The LOWELL database[4] consists of a set of 90 eye fundus images taken from 50 patients attending the diabetic retinal-screening program at City Hospital, Birmingham. The images were acquired using a Canon CR6 45MNf retinograph, with a 45° FOV at a resolution of 640×480 and 8 bits per color plane. The database provides four manual segmentations

[4]This database is not of public access. It was generously provided upon request by Dr. Andrew Hunter (contact: `ahunter@lincoln.ac.uk`)

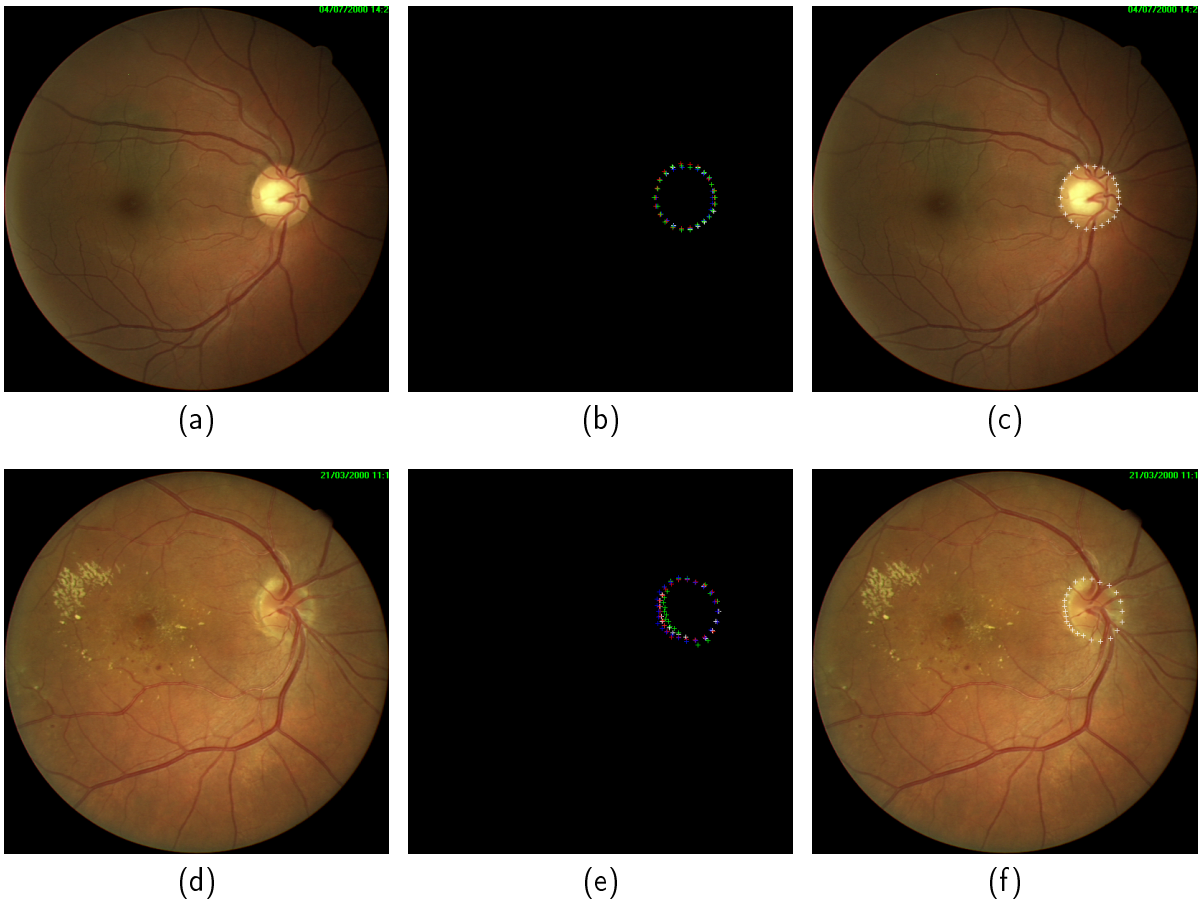


Figure A.5: *Examples of the material in the LOWELL database: (a) Image from a healthy patient. (b) The four true OD contours for image (a) represented in white, red, green and blue colors. (c) The average of the true OD contours in the gold standard image (b) represented in white color on the original image (a). (d) Image from a patient affected by DR. (e) The four true OD contours for image (d) represented in white, red, green and blue colors. (f) The average of the true OD contours in the gold standard image (e) represented in white color on the original image (d).*

of the optic disc rim made by different clinicians for each image. 19 patients were affected by type 2 diabetes mellitus, while the diabetes status was unavailable for the remaining 31. Some examples of the material in the LOWELL database can be checked in Figure A.5.

Appendix B

Performance Evaluation Based on Contigency Tables for Binary Classification

Since performance evaluation based on contigency tables for binary classification is used in several chapters of this thesis, this appendix offers the theoretical background on this kind of evaluation.

The appendix is divided into **three sections**. The **first section** briefly establishes the basis of the contingency tables for, in the **second section**, defining the metrics derived from these that concern this thesis. The **third section** is focused on *receiver operating characteristic (ROC)* graphs analysis. Here, the ROC graph is defined and foundations of this kind of analysis are explained.

B.1 Contigency tables for binary classification

Let A be a set of instances to be classified into two classes. Formally, each instance in A is mapped to one element of the set $\{p, n\}$ of positive and negative class labels. A binary classification model is a mapping from instances to two predicted classes. Some classification models produce a continuous output (e.g., an estimate of an instance's class membership probability) to which different thresholds may be applied to predict class membership. Other models produce a discrete class label indicating only the predicted class of the instance. To distinguish between the actual class and the predicted class the labels $\{P, N\}$ are used for the class predictions produced by a model.

Hypothesized class	True class: p	True class: n
P	True Positives (TP)	False Positives (FP)
N	False Negatives (FN)	True Negatives (TN)

Table B.1: *Contingency Table for Vessel Classification*

Given a classifier and an instance, there are four possible outcomes. If an instance is positive and it is classified as positive, it is counted as a *True Positive* (TP); if it is classified as negative, it is counted as a *False Negative* (FN). If an instance is negative and it is classified as negative, it is a *True Negative* (TN) and if it is classified as positive, it is a *False Positive* (FP). Thus, given a classifier and a set of instances, a two-by-two contingency table can be constructed representing the dispositions of the set of instances (Table B.1). The cases along the major diagonal represent the correct decisions made, and the cases of this diagonal represent the classificaction errors.

B.2 Metrics based on contigency tables for binary classification

The contingency table forms the basis for a family of metrics. Those that concern this thesis are *sensitivity* (Se), *specificity* (Sp), *positive predictive value* (Ppv), *negative predictive value* (Npv) and *accuracy* (Acc). These metrics fall within the $[0, 1]$ range, denoting 0 and 1 the worst and best performance cases. These metrics are formally defined as:

$$Se = \frac{TP}{TP + FN} \tag{B.1}$$

$$Sp = \frac{TN}{TN + FP} \tag{B.2}$$

$$Ppv = \frac{TP}{TP + FP} \tag{B.3}$$

$$Npv = \frac{TN}{TN + FN} \tag{B.4}$$

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \tag{B.5}$$

Thus, they can be described as:

- Se: it is the ratio of positive instances correctly mapped to the class P. So, Se provides information concerning the quality of the performed classification only on the positive class, particularly, the ratio of well-classified positive instances.
- Sp: it is the ratio of negative instances correctly mapped to the class N. Therefore, Sp assesses the quality of the performed classification only on the negative class, i.e., it is the ratio of well-classified negative instances.
- Ppv: it is the ratio of instances correctly mapped to the class P. Ppv is therefore a measure concerning the quality of the predicted positive class P, or, in other words, it can be seen as a measure of the similarity between the actual and predicted positive classes p and P, respectively.
- Npv: it is the ratio of instances correctly mapped to the class N. Consequently, Npv is a metric for measuring the quality of the predicted negative class N, i.e., it is a measure of the similarity between the actual and predicted negative classes n and N, respectively.
- Acc: it is the ratio of positive and negative instances correctly mapped to their corresponding predicted classes P and N, respectively. Thus, Acc is a generalist metric which provides information about the quality of the performed classification on both positive and negative classes, as it is the ratio of all well-classified instances.

As mentioned above, Acc takes values from the interval $[0, 1]$. The 1 value denotes the perfect global classification case, in which all instances were correctly classified to their corresponding predicted classes. On the contrary, the value 0 indicates cross classification, i.e., all instances were classified to their contrary class. Regarding this last case, note that the perfect classification result can be obtained by inverting the performed classification.

B.3 ROC analysis

A *receiver operating characteristic* (ROC) graph is a technique for visualizing and analyzing the classifiers performance. ROC graphs are two-dimensional graphs in which Se (also known

as true positive fraction, $t.p.f.$) is plotted on the Y axis and $(1 - Sp)$ (false positive fraction, $f.p.f.$) is plotted on the X axis.

Discrete classifiers, such as decision trees or rule sets, are designed to produce only a class decision, i.e., a P or N on each instance. When a discrete classifier is applied to a test set, it yields a single confusion matrix, which in turn corresponds to one ROC point. Thus, a discrete classifier produces only a single point in ROC space. The classifiers in Figure B.1 - image (a) are all discrete classifiers.

Other classifiers, such as a Naive Bayes classifier or a neural network, naturally yield an instance probability or score, a numeric value that represents the degree to which an instance is a member of the class P or, by complementarity, to the class N. These ranking or scoring classifiers can be used with a threshold to produce a discrete (binary) classifier: if the classifier output is above the threshold, the classifier produces a P, else a N. Each threshold value produces a different point in ROC space. By varying the threshold value from 0 to 1, many points in the ROC space are produced and a ROC curve can be traced through this space (Figure B.1, image (b)). The closer a curve approaches the top left corner, the better the performance of the system. The area under the curve (AUC) , which is 1 for a perfect system, is a single measure to quantify this behaviour. The ROC curve may also be interpreted as providing information about the degree of dependence of the classifier behaviour with the threshold variable. The magnitude of this performance dependence is therefore given by AUC. A ROC graph depicts relative tradeoffs between benefits (TP) and costs (FN). Several points in ROC space are important to note. The lower left point $(0, 0)$ represents the strategy of never issuing a positive classification; such a classifier commits no FP errors but also gains no TP. The opposite strategy, of unconditionally issuing positive classifications, is represented by the upper right point $(1, 1)$. The point $(0, 1)$ represents perfect classification. Informally, one point in ROC space is better than another if it is to the northwest ($t.p.f.$ is higher, $f.p.f.$ is lower, or both) of the first. Classifiers appearing on the left-hand side of an ROC graph, near the X axis, may be thought of as "conservative": they make positive classifications only with strong evidence so they make few FP errors, but they often have low TP rates as well. Classifiers on the upper right-hand side of an ROC graph may be thought of as "liberal": they make positive classifications with weak evidence so they classify nearly all positives correctly, but they often have high FP rates.

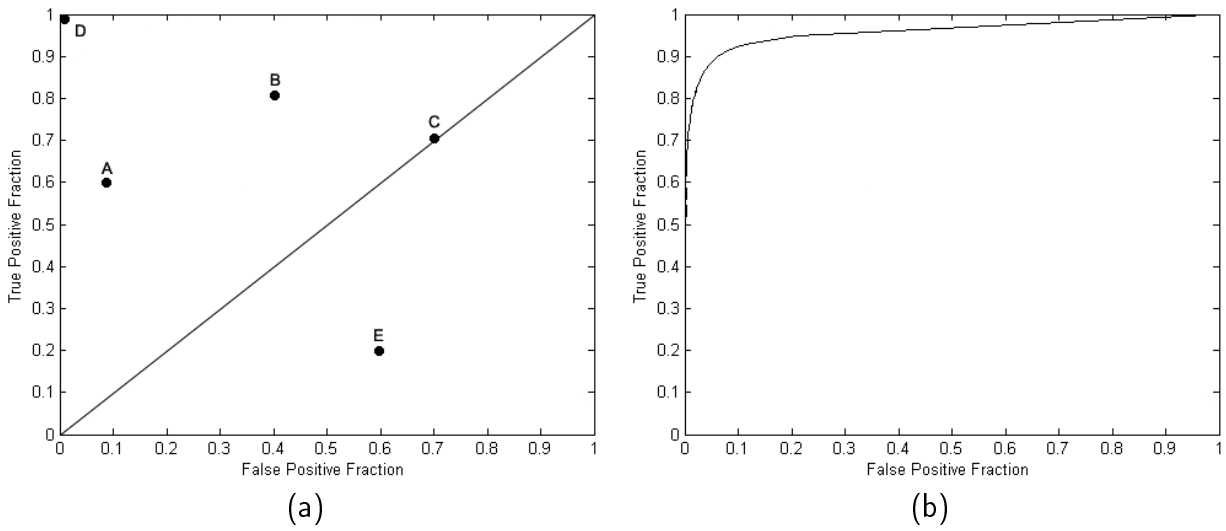


Figure B.1: *ROC grahps: (a) A basic ROC graph showing five discrete classifiers. (b) A ROC curve produced by a ranking or scoring classifier.*

Bibliography

Abràmoff, M., Alward, W., Greenlee, E., Shuba, L., Kim, C., Fingert, J., Kwon, Y., 2007. Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features. Investig. Ophthalmol. Vis. Sci. 48, 1665–1673.

Abràmoff, M., Niemeijer, M., Suttorp-Schulten, M., Viergever, M., Russell, S., 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, 193–198.

Aguilar Diosdado, M., Amo Alfonso, M., Lama Herrera, C., Mayoral Sánchez, E., 2009. II Plan Integral de Diabetes de Andalucía: 2009-2013. Junta de Andalucía, Consejería de Salud.

Akita, K., Kuga, H., 1982. A computer method of understanding ocular fundus images. Pattern Recogn. 16, 431–443.

Al-Rawi, M., Karajeh, H., 2007. Genetic algorithm matched filter optimization for automated detection of blood vessels from digital retinal images. Comput. Methods Programs Biomed. 87, 248–253.

Al-Rawi, M., Qutaishat, M., Arrar, M., 2007. An improved matched filter for blood vessel detection of digital retinal images. Comput. Biol. Med. 37, 262–267.

American Academy of Ophthalmology, 2008. Diabetic retinopathy. Preferred Practice Pattern Guidelines , http://www.aao.org/ppp.

American Diabetes Association, 2008. Economic costs of diabetes in the U.S. in 2007. Diabetes Care 31, 596–615.

Becker, D., Can, A., Turner, J., Tanenbaum, H., Roysam, B., 1998. Image processing algorithms for retinal montage, synthesis, mapping and real-time location determination. IEEE Trans. Biomed. Eng. 45, 115–118.

Bioux, S., Martin-Fernandez, M., Ungar, L., Nakamura, M., Koo, M.-S., McCarley, R., Shenton, M., 2007. On evaluating brain tissue classifiers without a ground truth. NeuroImage 36 (4), 1207–1224.

Bresnick, G., Mukamel, D., Dickinson, J., Cole, D., 2000. A screening approach to the surveillance of patients with diabetes for the presence of vision-threatening retinopathy. Opthalmology 107 (1), 19–24.

Can, A., Shen, H., Turner, J., Tanenbaum, H., Roysam, B., 1999. Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. IEEE Trans. Inform. Technol. Biomed. 3, 125–138.

Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., Goldbaum, M., 1989a. Automatic detection of the optic nerve in retinal images. In: Proceedings of the IEEE International Conference on Image Processing. Vol. 1. Singapore, pp. 1–5.

Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., Goldbaum, M., 1989b. Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Trans. Med. Imag. 8, 263–269.

Chutatape, O., Zheng, L., Krishman, S., 1998. Retinal blood vessel detection and tracking by matched gaussian and kalman filters. In: Proc. IEEE Int. Conf. Eng. Biol. Soc. Vol. 20. pp. 3144–3149.

Cinsdikici, M., Aydın, D., 2009. Detection of blood vessels in ophthalmoscope images using mf/ant (matched filter/ant colony) algorithm. Comput. Methods Programs Biomed. 96, 85–95.

Diabetic Retinopathy Screening Services in Scotland, 2003. Diabetic retinopathy screening: Annex E. Scottish diabetic retinopathy grading scheme. The Scottish Government Publications , http://www.scotland.gov.uk/Publications/2003/07/17638/23088.

Dice, L., 1945. Measures of the amount of ecologic association between species. J. Ecol. 26, 297–302.

Duda, R., Hart, P., 1972. Use of the hough transformation to detect lines and curves in picture. Commun. ACM 15, 11–15.

Early Treatment Diabetic Retinopathy Study Research Group, 1991a. Early photocoagulation for diabetic retinopathy: Etdrs report 9. Ophthalmology 98, 766–785.

Early Treatment Diabetic Retinopathy Study Research Group, 1991b. Grading diabetic retinopathy from stereoscopic color fundus photographs–an extension of the modified airlie house classification. etdrs report number 10. Ophthalmology 98, 786–806.

Espona, L., Carreira, M., Ortega, M., Penedo, M., 2007. A snake for retinal vessel segmentation. In: Pattern Recognition and Image Analysis. Vol. 4478 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg, pp. 178–185.

Fitzgibbon, A., Pilu, M., Fisher, R., 1999. Direct least square fitting of ellipses. IEEE Trans. Pattern Analysis and Machine Intelligence 21, 476–480.

Fleming, A., Goatman, K., Philip, S., Olson, J., Sharp, P., 2007. Automatic detection of retinal anatomy to assist diabetic retinopathy screening. Phys. Med. Biol. 52, 331–345.

Fong, D., Aiello, L., Gardner, T., King, G., Blankenship, G., Cavallerano, J., Ferris, F., Klein, R., 2003. Diabetic retinopathy. Diabetes Care 26, 226–229.

Foracchia, M., Grisan, E., Ruggeri, A., 2001. Extraction and quantitative description of vessel features in hypertensive retinopathy fundus images. In: Book Abstracts 2nd Int. Workshop Comput. Asst. Fundus Image Anal. p. 6.

Foracchia, M., Grisan, E., Ruggeri, A., 2004. Detection of optic disc in retinal images by means of a geometrical model of vessel structure. IEEE Trans. Med. Imag. 23, 1189–1195.

Frame, A., Undrill, P., Cree, M., Olson, J., McHardy, K., Sharp, P., Forrester, J., 1998. A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms. Comput. Biol. Med. 28 (3), 225–238.

Gagnon, L., Lalonde, M., Beaulieu, M., Boucher, M.-C., 2001. Procedure to detect anatomical structures in optical fundus images. In: Proc. SPIE Med. Imaging: Image Processing. Vol. 4322. pp. 1218–1225.

Gander, W., Golub, G., Strebel, R., 1994. Least square fitting of circles and ellipses. BIT 43, 558–578.

Gang, L., Chutatape, O., Krishnan, S., 2002. Detection and measurement of retinal vessels in fundus images using amplitude modified second-order gaussian filter. IEEE Trans. Biomed. Eng. 49, 168–172.

Gardner, G., Keating, D., Williamson, T., Elliott, A., 1996. Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool. Br. J. Ophthalmol. 80, 940–944.

Goa, X., Bharath, A., Stanton, A., Hughes, A., Chapman, N., Thom, S., 2001. A method of vessel tracking for vessel diameter measurement on retinal images. In: Proc. ICIP. pp. 881–884.

Gonzalez, R., Woods, R., 2002. Digital Image Processing 2nd ed. Addison-Wesley Publishing Company, Ch. Image Segmentation, pp. 577–581.

Grisan, E., Ruggeri, A., 2003. A divide and impera strategy for the automatic classification of retinal vessels into arteries and veins. In: Proc. 25th Int. Conf. IEEE Eng. Med. Biol. Soc. pp. 890–893.

Hagan, M., Demuth, H., Beale, M., 1996. Neural Network Design. PWS Pub. Co., Boston.

Hatanaka, Y., Fujita, H., Aoyama, M., Uchida, H., Yamamoto, T., 2004. Automated analysis of the distributions and geometries of blood vessels on retinal fundus images. In: Proc. SPIE Med. Imag. 2004: Image Process. Vol. 5370. pp. 1621–1628.

Haykin, S., 1999. Neural Networks. A Comprehensive Foundation. Prentice-Hall, Upper Saddle River, NJ.

Heneghan, C., Flynn, J., O'Keefe, M., Cahill, M., 2002. Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis. Med. Image Anal. 6, 407–429.

Hoover, A., Goldbaum, M., 1998. Fuzzy convergence. In: Proc. IEEE Computer Soc. Conf. Computer Vis. Pattern Recognit. pp. 716–721.

Hoover, A., Goldbaum, M., 2003. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. IEEE Trans. Med. Imag. 22, 951–958.

Hoover, A., Kouznetsova, V., Goldbaum, M., 2000. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Trans. Med. Imag. 19, 203–210.

Hough, P., 1962. Methods and means for recognizing complex patterns.

Hu, M., 1962. Visual pattern recognition by moment invariants. IRE Trans. Inform. Theory 8 (2), 179–187.

Huttenlocher, D., Klanderman, D., Rucklige, A., 1993. Comparing images using the hausdorff distance. IEEE Trans. Pattern Anal. Mach. Intell. 15 (9), 850–863.

Jaccard, P., 1901. Etude comparative de la distribution florale dans une portion des alpes et des jura. Bull. Soc. Vaudoise Sci. Nat. 37, 547–579.

Jagoe, R., Arnold, J., Blauth, C., Smith, P., Taylor, K., Wootton, R., 1992. Measurement of capillary dropout in retinal angiograms by computerised image analysis. Pattern Recognit. Lett. 13, 143–151.

Javitt, J., Aiello, L., Chiang, Y., Ferris, F., Canner, J., Greenfield, S., 1994. Preventative eye care in people with diabetes is cost-saving to the federal government: implications for health-care reform. Diabetes Care 17 (8), 909–917.

Jiang, X., Mojon, D., 2003. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. IEEE Trans. Pattern Analy. Mach. Intell. 254, 131–137.

Kimme, C., Ballard, D., Sklansky, J., 1975. Finding circles by an array of accumulators. Communications of the Association for Computing Machinery 18, 120–122.

Kinyoun, J., Barton, F., Fisher, M., Hubbard, L., Aiello, L., Ferris, F., 1989. Detection of diabetic macular edema. ophthalmoscopy versus photography-early treatment diabetic retinopathy study report number 5. the etdrs research group. Ophthalmology 96, 746–750.

Klein, R., Meuer, S., Moss, S., Klein, B., 1995. Retinal microaneurysm counts and 10-year progression of diabetic retinopathy. Archives of Ophthalmology 113, 1386–1391.

Lairson, D., Pugh, J., Kapadia, A., Lorimor, R., Jacobson, J., Velez, R., 1992. Cost effectiveness of alternative methods for diabetic retinopathy screening. Diabetes Care 15, 1369–1377.

Lalonde, M., Beaulieu, M., Gagnon, L., 2001. Fast and robust optic disk detection using pyramidal decomposition and hausdorff-based template matching. IEEE Trans. Med. Imag. 20, 1193–1200.

Lam, B., Yan, H., 2008. A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields. IEEE Trans. Med. Imag. 27, 237–246.

Lamecker, H., Lange, T., Seebass, M., April 2004. Segmentation of the liver using a 3d statistical shape model. Tech. rep., Zuse Institute, Berlin.

Larsen, H., 1976. The Ocular Fundus: A Color Atlas. Copenhagen, Denmark: Munksgaard.

Larsen, M., Godt, J., Larsen, N., Lund-Andersen, H., Sjø lie, A., Agardh, E., Kalm, H., Grunkin, M., Owens, D., 2003. Automated detection of fundus photographic red lesions in diabetic retinopathy. Investigat. Opht. Vis. Sci. 44 (2), 761–766.

Lee, S., McCarty, C., Taylor, H., Keeffe, J., 2001. Costs of mobile screening for diabetic retinopathy: A practical framework for rural populations. Aust. J. Rural Health 8, 186–192.

Lee, S., Wang, Y., Lee, E., 1999. A computer algorithm for automated detection and quantification of microaneurysms and haemorrhages in color retinal images. In: SPIE Conference on Image Perception and Performance. Vol. 3663. pp. 61–71.

Li, H., Chutatape, O., 2001. Automatic location of optic disc in retinal images. In: Proc. IEEE Int. Conf. on Image Proc. pp. 837–840.

Li, H., Chutatape, O., 2003. A model-based approach for automated feature extraction in fundus images. In: 9th IEEE Int. Conf. Computer Vision (ICCV'03). pp. 394–399.

Li, H., Chutatape, O., 2004. Automated feature extraction in color retinal images by a model based approach. IEEE Trans. Biomed. Eng. 51 (2), 246–254.

Liu, I., Sun, Y., 1993. Recursive tracking of vascular networks in angiograms based on the detection-deletion scheme. IEEE Trans. Med. Imag. 12, 334–341.

Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E., Kennedy, L., 2004a. Optic nerve head segmentation. IEEE Trans. Med. Imag. 23, 256–264.

Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Kennedy, R., 2004b. Measurement of retinal vessel widths from fundus images based on 2-d modeling. IEEE Trans. Med. Imag. 23 (10), 1196–1204.

Marquardt, D., 1963. An algorithm for least-squares estimation of nonlinear parameters. SIAM J. Appl. Math. 11, 431–441.

Martinez-Perez, M., Hughes, A., Stanton, A., Thom, S., Bharath, A., Parker, K., 1999. Scale-space analysis for the characterisation of retinal blood vessels. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI99. pp. 90–97.

Martinez-Perez, M., Hughes, A., Stanton, A., Thom, S., Chapman, N., Bharath, A., Parker, K., 2002. Retinal vascular tree morphology: A semiautomatic quantification. IEEE Trans. Biomed. Eng. 49 (8), 912–917.

Martinez-Perez, M., Hughes, A., Thom, S., Bharath, A., Parker, K., 2007. Segmentation of blood vessels from red-free and fluorescein retinal images. Med. Imag. Anal. 11, 47–61.

Mason, J., 2003. National screening for diabetic retinopathy: clear vision needed. Diabet. Med. 20, 959–961.

Massin, P., Angioi-Duprez, K., Bacin, F., Cathelineau, B., Cathelineau, G., Chaine, G., Coscas, G., Flament, J., Sahel, J., Turut, P., Guillausseau, P., Gaudric, A., 1996. Recommandations de l'alfediam pour le dépistage et la surveillance de la rétinopathie diabétique. Diabetes Metab. 22, 203–209.

Massin, P., Angioi-Duprez, K., Bacin, F., Cathelineau, B., Cathelineau, G., Chaine, G., Coscas, G., Flament, J., Sahel, J., Turut, P., Guillausseau, P., Gaudric, A., 1997. Recommandations de l'alfediam pour le dépistage et la surveillance de la rétinopathie diabétique. J. Fr. Ophtalmol. 20, 302–310.

Massin, P., Erginay, A., Gaudric, A., 2000. Rétinopathie Diabétique. Vol. 1. Elsevier, Paris, France.

Matsopoulos, G., Asvestas, P., Mouravliansky, N., Delibasis, K., 2004. Multimodal registration of retinal images using self organizing maps. IEEE Trans. Med. Imag. 23, 1557–1563.

McInerney, T., Terzopoulos, D., 2000. T-snakes: Topology adaptive snakes. Med. Imag. Anal. 4, 73–91.

Mendonça, A., Campilho, A., 2006. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. IEEE Trans. Med. Imag. 25, 1200–1213.

Niemeijer, M., Abràmoff, M., van Ginneken, B., 2007. Segmentation of the optic disc, macula and vascular arch in fundus photographs. IEEE Transactions on Medical Imaging 26, 116–127.

Niemeijer, M., Abràmoff, M., van Ginneken, B., 2009. Fast detection of the optic disc and fovea in color fundus photographs. Medical Image Analysis 13, 859–870.

Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., Abràmoff, M., 2004. Comparative study of retinal vessel segmentation methods on a new publicly available database. In: Fitzpatrick, J. M., Sonka, M. (Eds.), SPIE Medical Imaging. Vol. 5370. SPIE, SPIE, pp. 648–656.

Niemeijer, M., van Ginneken, B., Staal, J., Suttorp-Schulten, M., Abràmoff, M., 2005. Automatic detection of red lesions in digital color fundus photographs. IEEE Trans. Med. Imag. 24 (5), 584–592.

Osareh, A., Mirmehdi, M., Thomas, B., Markham, R., 2002. Comparison of colour spaces for optic disc localisation in retinal images. In: 16th Int. Conf. Pattern Recognition. pp. 743–746.

Osareh, A., Mirmehdi, M., Thomas, B., Markham, R., 2003. Automated identification of diabetic retinal exudates in digital colour images. Br. J. Ophthalmol. 87, 1220–1223.

Otsu, N., 1978. A threshold selection method from gray-scale histogram. IEEE Trans. Syst. Man Cybern. 8, 62–66.

Paglieroni, D., 2004. Design considerations for image segmentation quality assessment measures. Pattern Recognition 37, 1607–1617.

Philip, S., Fleming, A., Goatman, K., Fonseca, S., McNamee, P., Scotland, G., Prescott, G., Sharp, P., Olson, J., 2007. The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme. Br J Ophthalmol. 91, 1512–1517.

Quigley, H., Broman, A., 2006. The number of people with glaucoma worldwide in 2010 and 2020. Br. J. Ophthalmol. 90, 262–267.

Rassam, S., Patel, V., Brinchmann-Hansen, O., Engvold, O., Kohner, E., 1994. Accurate vessel width measurement from fundus photographs: A new concept. Br. J. Ophthalmol. 78, 24–29.

Reza, A., Eswaran, C., Hati, S., 2008. Automatic tracing of optic disc and exudates from color fundus images using fixed and variable thresholds. J. Med. Syst. 33, 73–80.

Ricci, E., Perfetti, R., 2007. Retinal blood vessel segmentation using line operators and support vector classification. IEEE Trans. Med. Imag. 26, 1357–1365.

Rumelhart, D., McClelland, J., PDP Research Group, t., 1987. Parallel distributed processing, explorations in the micro structure of cognition, Vol. 1: Foundations. MIT Press Cambridge, MA, USA.

Schwiegerling, J., 2004. Field Guide to Visual and Ophthalmic Optics. Bellingham, WA: SPIE.

Serra, J., 1982. Image Analysis and Mathematical Morphology. Vol. 1. London: Academic Press.

Shen, H., Roysam, B., Stewart, C., Turner, J., Tanenbaum, H., 2001. Optimal scheduling of tracing computations for real-time vascular landmark extraction from retinal fundus images. IEEE Trans. Inf. Technol. Biomed. 5, 77–91.

Shi, Y., Sun, H., 1999. Image and Video Compression for Multimedia Engineering: Fundamental, Algorithm, and Standards. Boca Raton, FL: CRC.

Sinthanayothin, C., Boyce, J., Cook, H., Williamson, T., 1999. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. Br. J. Ophthalmol. 83, 902–910.

Soares, J., Leandro, J., Cesar Jr., R., Jelinek, H., Cree, M., 2006. Retinal vessel segmentation using the 2d gabor wavelet and supervised classification. IEEE Trans. Med. Imag. 25, 1214–1222.

Soille, P., 1999. Morphological Image Analysis: Principles and Applications. Berlin, Germany: Springer-Verlag.

Spencer, T., Olson, J., McHardy, K., Sharp, P., Forrester, J., 1996. An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. Comput. Biomed. Res. 29 (4), 284–302.

Staal, J., Abràmoff, M., Niemeijer, M., Viergever, M. A., van Ginneken, B., 2004. Ridge based vessel segmentation in color images of the retina. IEEE Trans. Med. Imag. 23, 501–509.

Streeter, L., Cree, M., 2003. Microaneurysm detection in colour fundus images. In: Proceedings of Image and Vision Computing New Zealand 2003. pp. 280–285.

Taylor, H., Keeffe, J., 2001. World blindness: a 21st century perspective. Br. J. Ophthalmol. 85, 261–266.

Tobin, K., Chaum, E., Govindasamy, V., Karnowski, T., 2007. Detection of anatomic structures in human retinal imagery. IEEE Trans. Med. Imag. 26 (12), 1729–1739.

Tolias, Y., Panas, S., 1998. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. IEEE Trans. Med. Imag. 17, 263–273.

Usher, D., Dumskyj, M., Himaga, M., Williamson, T., Nussey, S., Boyce, J., 2003. Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. Diabet. Med. 21, 84–90.

Walter, T., Klein, J., 2001. Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques. In: ISMDA '01:

Proceedings of the Second International Symposium on Medical Data Analysis. Springer-Verlag, London, UK, pp. 282–287.

Walter, T., Klein, J., 2005. Handbook of Biomedical Image Analysis. Vol. 2. Kluwer Academic/Plenum, Ch. Automatic analysis of color fundus photographs and its application to the diagnosis of diabetic retinopathy, pp. 315–368.

Walter, T., Massin, P., Erginay, A., Ordonez, R., Jeulin, C., Klein, J., 2007. Automatic detection of microaneurysms in color fundus images. Med. Image Anal. 11, 555–566.

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, 1047–1053.

Wong, D. K., Liu, J., Lim, J., Jia, X., Yin, F., Li, H., Wong, T., 2008. Level-set based automatic cup-to-disc ratio determination using retinal fundus images in argali. In: 30th Annual International IEEE EMBS Conference. pp. 2266–2269.

Xu, J., Chutatape, O., Sung, E., Zheng, C., Kuan, P. T., 2007. Optic disk feature extraction via modified deformable model technique for glaucoma analysis. Pattern Recognition 40, 2063–2076.

Youssif, A., Ghalwash, A. Z., Ghoneim, A. R., 2008. Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter. IEEE Trans. Med. Imag. 27, 11–18.

Zana, F., Klein, J., 1999. A multimodal registration algorithm of eye fundus images using vessels detection and hough transform. IEEE Trans. Med. Imag. 18, 419–428.

Zana, F., Klein, J., 2001. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. IEEE Trans. Image Processing 10, 1010–1019.

Zhang, B., Wu, X., You, J., Li, Q., Karray, F., 2010. Detection of microaneurysms using multi-scale correlation coefficients. Pattern Recognition 43, 2237–2248.

Zhou, L., Rzeszotarski, M., Singerman, L., Chokreff, J., 1994. The detection and quantification of retinopathy using digital angiograms. IEEE Trans. Med. Imag. 13, 619–626.

Index