

UNIVERSITAT DE VALÈNCIA



Técnicas para la correspondencia y reconocimiento
de estructuras de imágenes digitales. Aplicación al
diagnóstico asistido de enfermedades oculares
mediante análisis de angiografías.

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA

AL DPTO. DE INFORMÁTICA Y ELECTRÓNICA

FACULTAD DE FÍSICA

UNIVERSITAT DE VALÈNCIA

Esther de Ves Cuenca

Dirección

Juan Domingo Esteve, Guillermo Ayala Gallego y Amelia Simó Vidal

28 de junio de 1999

UMI Number: U607752

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U607752

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSITAT DE VALÈNCIA
BIBLIOTECA CIÈNCIES

Nº Registre13.544.....

DATA29.X.1999.....

SIGNATURA T.D.367 FÍSICAS

Nº LIBIS: j 201 27534

29 cm.



D. JUAN DE MATA DOMINGO ESTEVE, profesor titular de Ciencias de la Computación e Inteligencia Artificial de la Universitat de Valencia,

D. GUILLERMO AYALA GALLEGO, profesor titular de Estadística e Investigación Operativa de la Universitat de Valencia, y

DÑA. AMELIA SIMO VIDAL, profesora titular de Estadística e Investigación Operativa de la Universitat Jaume I de Castellón

CERTIFICAMOS que la presente memoria

“Técnicas para la correspondencia y reconocimiento de estructuras de imágenes digitales. Aplicación al diagnóstico asistido de enfermedades oculares mediante análisis de angiografías”

ha sido realizada bajo nuestra dirección, en el Departamento de Informática y Electrónica de la Universitat de Valencia por Dña. Esther de Ves Cuenca, y constituye su tesis para optar al grado de Doctora en Ingeniería Informática.

Y para que conste, en cumplimiento de la legislación vigente, presentamos ante la Facultad de Física de la Universitat de Valencia, a 30 de Junio de 1999.

LOS DIRECTORES

Juan de M. Domingo Esteve

Guillermo Ayala gallego

Amelia Simó Vidal



A mis padres por su apoyo y confianza incondicional.

A Oscar con todo mi cariño.



Agradecimientos

Me gustaría, en primer lugar, manifestar mi más sincero agradecimiento a mis directores de tesis, los doctores Juan Domingo, Guillermo Ayala y Amelia Simó por su apoyo continuo y sus valiosos consejos desde que tuve la suerte de conocerles.

Me gustaría también agradecer a las doctoras Lucía Martínez-Costa y Pilar Marco su colaboración y ayuda por aportar sus valiosos conocimientos y experiencia clínica en oftalmología, así como gran parte de las imágenes usadas. Sin ellas este trabajo no habría sido posible.

Al Dr. Edwin Hancock que ha sido mi tutor durante mis estancias en la Universidad de York, por su disponibilidad y todas sus sugerencias para el desarrollo de este trabajo.

También, a todos los miembros del departamento de Informática y Electrónica y del Instituto de Robótica y, especialmente, a su director, Gregorio Martín Quetglás gracias al cual los estudios de Ingeniería Informática son una realidad.

A mis compañeros de trabajo, Elena Díaz, Ricardo Ferrís, Juan Gutiérrez e Irene Epifanio que me han apoyado desinteresadamente en todo momento. A Vicente Cerverón, Jesús Albert y Francesc Ferri por su generosa ayuda.

A mis buenas amigas y compañeras Isabel Gosálvez y Xaro Benavent por su comprensión y su amistad.

A todas las personas que desinteresadamente producen programas de excelente calidad para aplicaciones de todo tipo. En particular, me gustaría mencionar Linux, VISTA, GNU, xfig y latex.

La autora quiere agradecer también su apoyo:

A la Conselleria de Cultura, Educació y Ciència de la Generalitat Valenciana por la concesión de una beca de formación de personal investigador asociada al proyecto GV-2221 en el departamento de Informática y Electrónica de la Facultad de Físicas de la Universitat de València, desde Enero de 1996 hasta Octubre de 1998.

A la Comisión Interministerial de Ciencia y Tecnología por el proyecto TIC98-1019 titulado "Fusion de imágenes en color, aneritras y angiográficas para la detección automática de microaneurismas en enfermos de retinopatía diabética".

Índice General

1	Introducción	1
1.1	Introducción	1
1.2	Análisis de imágenes médicas	2
1.3	Descripción del problema y objetivos	3
1.4	Estructura del sistema	8
2	Imágenes de fondo ocular	11
2.1	Introducción	11
2.2	El ojo humano	11
2.2.1	Anatomía del ojo humano	12
2.2.2	Fisiología del ojo humano	14
2.3	Imágenes de fondo de ojo	15
2.3.1	¿Qué es una cámara de fondo de ojo?	16
2.4	Sistema óptico de las cámaras de fondo ocular	18
2.5	Tipos de imágenes de fondo de ojo	20
2.5.1	Fotografía de fondo monocromática	21
2.5.2	Angiografías	23
2.6	Enfermedades que afectan a la retina	25
2.6.1	Retinopatía diabética	25
2.6.2	Trombosis venosa	27
2.6.3	Neovascularización subretiniana	28
3	Alineamiento de angiografías fluoresceínicas	29
3.1	Introducción	29
3.1.1	Descripción del problema, objetivos y revisión de la literatura	30
3.1.2	Descripción del método	32
3.2	Detector de puntos relevantes	34
3.3	Características locales	36

3.4	Modelo estocástico para el encaje de imágenes	37
3.4.1	Correspondencias iniciales	39
3.4.2	Un procedimiento iterativo de estimación	41
3.4.3	Un campo aleatorio de Markov en la matriz de correspondencias	42
3.4.4	Transformación geométrica	44
3.5	Resultados experimentales	45
3.6	Discusión	48
4	Alineamiento de angiografías mediante el algoritmo EM Generalizado	57
4.1	Introducción	57
4.2	Conceptos y notación	58
4.2.1	Transformación afín de un conjunto de puntos	58
4.2.2	Grafos relacionales	59
4.3	Extracción de los datos	60
4.4	El algoritmo EM generalizado	61
4.5	El proceso de encaje	62
4.5.1	Esperanza	64
4.5.2	Maximización utilizando agrupamiento difuso	66
4.6	Resultados experimentales	68
5	Modelo de transformación no lineal para alineamiento de angiografías	75
5.1	Introducción	75
5.2	Modelo de transformación	77
5.3	Estimación de parámetros	81
5.4	El proceso de alineamiento	82
5.5	Resultados experimentales	86
5.6	Discusión	93
6	Segmentación de angiografías fluoresceínicas	95
6.1	Introducción	95
6.2	Aproximaciones bayesianas	99
6.3	Un modelo estocástico	101
6.4	Modelo <i>a priori</i> para las etiquetas	104
6.5	Estimación de x	107

6.5.1	Estimadores iniciales	109
6.5.2	Convergencia	109
6.6	Parámetros e hiperparámetros	109
6.6.1	Parámetros en la distribución <i>a priori</i>	110
6.6.2	Distribución de los parámetros de linealidad	111
6.6.3	Parámetros de la distribución de intensidad	111
6.7	Resultados	111
7	Evaluación del edema macular	117
7.1	Introducción	117
7.2	Métodos	118
7.2.1	Preprocesamiento de la secuencia	118
7.2.2	Segmentación de la difusión	119
7.2.3	Concordancia con observadores externos	123
7.3	Material	123
7.4	Resultados	124
7.5	Discusión	127
8	Conclusiones, aportaciones y trabajo futuro	133
8.1	Resultados y aportaciones	133
8.2	Trabajo futuro	137
A	Campos aleatorios	139
A.1	La imagen como un campo aleatorio	139
A.2	Campos aleatorios	139
A.3	El teorema Hammersley-Clifford	141
B	El algoritmo EM	145
B.1	Descripción del algoritmo	145
B.1.1	Mixturas de densidades	147
	Bibliografía	149

Índice de Tablas

3.1	Resultados del experimento 1	55
3.2	Resultados del experimento 2	56
3.3	Resultados del experimento 3	56
3.4	Resultados del experimento 4	56
5.1	Angulos medidos y estimados para varios pares de imágenes	88
5.2	Esta tabla muestra el número de encajes obtenidos con el algoritmo de preprocesamiento y el número de encajes correctos. $R1$ y $R2$ son respectivamente $R1 = \frac{nc}{ne}$ y $R2 = \frac{nf}{nc}$ donde ne es el número de encajes iniciales, nc es el número de encajes correctos del total y nf es el número de encajes correctos obtenidos por el procedimiento propuesto	88
5.3	Comparación de la transformación afín y el desarrollo en serie de Taylor de la transformación propuesta para el angiograma de la figura 5.5 . . .	91
6.1	p-valores correspondientes a los tests de normalidad y no correlación .	104
6.2	Parámetros estimados utilizando datos de entrenamiento.	112
7.1	Comparación entre la detección realizada por los observadores y la automática Resultados experimentales correspondientes a seis angiogramas con un valor intermedio de k , 0.5.	131

Índice de Figuras

1.1	Estructura del sistema	9
2.1	Anatomía del ojo humano	13
2.2	Variación de parte de la escena fotografiada por una cámara según su amplitud de campo	17
2.3	Imágenes de fondo de ojo con distinta apertura de campo. Izquierda: 45 grados. Centro: 30 grados. Derecha: 20 grados	17
2.4	Sistema óptico utilizado en una cámara de fondo. Arriba: Sistema total de lentes. Izquierda: Primeras lentes en el camino de la luz. Centro: Filtros. Derecha: Fuente de luz	18
2.5	Haz de luz entrando y saliendo a través de la pupila	19
2.6	Izquierda:Lentes y espejos que dirigen la luz imagen hacia la película fotográfica. Derecha: Sistema óptico de una cámara reflex	20
2.7	Aplicación de diversos tipos de filtros en la captura de imágenes de fondo de ojo. Arriba-Izquierda: luz blanca. Arriba-Derecha: luz azul. Abajo-Izquierda: luz verde. Abajo-Derecha: luz roja	22
2.8	Primeras angiografías fluoresceínicas realizadas	25
3.1	Experimento 1: caso de retinopatía diabética; $ic = 7$ con ecualización	51
3.2	Experimento 2: caso de retinopatía diabética; $ic = 3$ con ecualización	52
3.3	Experimento 3: caso de oclusión venosa; $ic = 4$ sin ecualización	53
3.4	Experimento 4: caso de neovascularización; $ic = 7$ sin ecualización	54
3.5	Gráfica cuantil-cuantil de las coordenadas primera y segunda de los residuos obtenidos en el experimento 1.	55
4.1	Imágenes originales (a) primera image y (b) segunda imagen de la secuencia digital	69
4.2	Secuencia de angiografías sobre la que se aplicado el algoritmo propuesto	69

4.3	Vasos sanguíneos segmentados de (a) la primera image y (b) la segunda imagen de la secuencia digital	70
4.4	Resultado del encaje mediante un método basado en relajación probabilística	71
4.5	Correspondencia obtenidas para dos diferentes imágenes de la misma secuencia	72
4.6	Resultado del encaje con el algoritmo generalizado EM	72
4.7	Ejemplo en el que el algoritmo obtiene algunas correspondencias erróneas	73
4.8	Distribución de las transformaciones en el espacio de parámetros	73
5.1	Los dos sistemas de coordenadas de la imagen y su posición relativa. . .	78
5.2	Varias imágenes de la secuencia sintética	87
5.3	Representación de la función objetivo variando uno de los parámetros en cada caso. El resto de parámetros se han fijado a los valores donde se obtiene el mínimo.	89
5.4	Correspondencias entre diferentes pares de imágenes. Primera fila: Correspondencias obtenidas mediante un algoritmo de relajación probabilística. Última fila: correspondencias utilizando información de la transformación. Izquierda: imágenes 1 y 2 de la secuencia. Centro: imágenes 1 y 7. Derecha: imágenes 1 y 13	90
5.5	Correspondencias establecidas entre un par de angiografía	91
5.6	Comparación del ECM de una transformación afín y del modelo propuesto	92
6.1	Dos imágenes angiográficas de fondo de ojo.(a) Paciente con retinopatía diabética. (b) Paciente con oclusión venosa	96
6.2	<i>Cliques</i> que contribuyen a la función negpotencial	105
6.3	Estimadores iniciales para la segmentación de imágenes	114
6.4	Resultados utilizando 50 iteraciones del algoritmo ICM	114
6.5	Resultados utilizando 100 iteraciones del algoritmo SA	115
6.6	Imagen diferencia de los resultados obtenidos con el ICM y SA	115
6.7	Comparación del resultado obtenido por el ICM y una segmentación manual	116

6.8	Gráficas de convergencia para la imagen (b) de la figura (estabilización del número de <i>pixels</i> pertenecientes a la fovea) 6.8. (a) Algoritmo SA. (b) Algoritmo ICM.	116
7.1	Izquierda: Imagen inicial (arriba) y final del caso 1 del angiograma a: oclusión de la ramificación de la vena retiniana de un paciente de 24 años de edad. Derecha: algunas correspondencias establecidas entre las dos imágenes para su posterior alineamiento.	120
7.2	Imágenes diferencia correspondientes a los angiogramas a, b y c. Las imágenes están reescaladas al rango de intensidad [0, 255]. Los centros foveales y la región de interés aparecen resaltados.	122
7.3	Angiogramas a, comenzando por la fila superior y de izquierda a derecha: imagen diferencia, difusión segmentada por los expertos, difusión segmentada automáticamente y difusión detectada automáticamente después de una apertura.	126
7.4	Cuantificación correspondientes a los angiogramas a, b y c	130



Capítulo 1

Introducción

1.1 Introducción

Esta tesis tiene como objetivo fundamental proponer nuevos métodos de análisis de imágenes, y aplicar otros ya existentes a un campo concreto del diagnóstico médico: la interpretación de angiografías fluoresceínicas. Pretende ser, por tanto, un trabajo en el que, además del interés puramente científico de los algoritmos propuestos desde el punto de vista del análisis de imágenes, se busca la consecución de un objetivo práctico: proporcionar a los oftalmólogos herramientas que les ayuden a interpretar y analizar, en términos cualitativos y cuantitativos, un tipo particular de imágenes especialmente importante en su campo: las llamadas angiografías fluoresceínicas de fondo de ojo. Por ello, los métodos genéricos y los algoritmos concretos que se propondrán han sido desarrollados o escogidos con una finalidad en mente, y con unas restricciones impuestas por las imágenes que se manejan, que son, al fin, las impuestas por el propio problema.

Por ello, es necesario entender, primero de modo genérico, y a continuación restringiéndose a nuestro caso, cuál es el origen e importancia del análisis de imágenes en el diagnóstico médico. A ello se dedicará la primera sección de este capítulo. En la segunda se describirán con detalle los problemas concretos que pretendemos abordar y finalizaremos con una descripción de la estructura del sistema que proponemos.

1.2 Análisis de imágenes médicas

Con el descubrimiento de los rayos X en 1896 (Roetgen, 1845-1923) la medicina encontró un nuevo método para el diagnóstico de enfermedades. La novedad de esta técnica se basaba en la posibilidad de fotografiar estructuras internas del cuerpo humano, algo imposible hasta el momento, gracias a la diferente absorción de rayos X por estructuras de distinta densidad y espesor. La contribución de este descubrimiento es, obviamente inmensa. Pero inmediatamente se hace necesario encontrar métodos adecuados para la correcta interpretación de estas imágenes.

El uso de sistemas informáticos en medicina que permiten almacenar, manipular y analizar un gran número de imágenes médicas está siendo hoy la mejor ayuda en este campo. En la actualidad, la utilización de ordenadores va desde la simple visualización en un monitor de imágenes médicas (radiografías, ecografías, etc.) para que sean analizadas por un experto hasta los sistemas de diagnóstico asistido de enfermedades, que sólo pueden ser detectadas y correctamente evaluadas mediante el examen detallado de algún tipo de imagen médica.

Una de las aplicaciones informáticas en medicina que está teniendo mayor auge en los últimos años es la reconstrucción de estructuras anatómicas en tres dimensiones partiendo de la información contenida en imágenes planas que corresponden a secciones consecutivas del cuerpo humano. Esto permite crear reconstrucciones volumétricas en las que se pueden observar diversas partes internas del cuerpo humano, lo que puede servir tanto para la detección de lesiones que no serían observables de otra manera como para ampliar el conocimiento que tenemos sobre nuestro propio organismo. Otra aplicación relacionada directamente con la visualización en tres dimensiones son los sistemas de realidad virtual enfocados al aprendizaje, en los que estudiantes de medicina pueden aprender de una manera interactiva como desenvolverse en un quirófano gracias al gran realismo que ofrece la realidad virtual ([oM]).

Centrándonos en nuestro objetivo primordial, la Oftalmología es, después del Radiodiagnóstico, la especialidad médica que más se apoya en la imagen para emitir un diagnóstico. El diagnóstico asistido por ordenador consiste en el análisis automático de imágenes digitales mediante diversos métodos que nos ofrece la Visión por Computador

para detectar y cuantificar anomalías concretas que indiquen la presencia de una determinada enfermedad o la potencialidad de padecerla en el futuro. Las imágenes con las que se trabaja en oftalmología son imágenes de fondo de ojo, en las que se visualiza la retina humana. Estas imágenes se utilizan como una ayuda al diagnóstico para estudio de la evolución de una cierta enfermedad con o sin tratamiento y para la cuantificación, si es el caso, de la gravedad de la misma.

1.3 Descripción del problema y objetivos

El origen de este trabajo surge de un proyecto interdisciplinar para el estudio de la retinopatía diabética, enfermedad que padecen los enfermos diabéticos, con el objetivo de desarrollar algoritmos de Visión por Computador para el análisis de secuencias de imágenes de fondo de ojo que ayudarán al diagnóstico de esta patología, al estudio de la evolución de la enfermedad a lo largo del tiempo, y a la valoración de la gravedad de las lesiones sufridas. El objetivo concreto que nos planteamos es el desarrollo de nuevas técnicas para el análisis de imágenes de fondo de ojo, concretamente de angiografías fluoresceínicas, que puedan ser utilizadas para el diagnóstico de las patologías de la retina, en particular de la retinopatía diabética. Es importante hacer notar que aunque estas técnicas han sido desarrolladas en un marco muy concreto, son suficientemente generales de forma que su aplicación no está restringida a esta patología sino que se podrán utilizar en otro tipo de enfermedades oculares e incluso podrían, con las debidas modificaciones, extrapolarse al análisis de muchos otros tipos de imagen médica.

Con el fin de poder seguir más fácilmente el desarrollo de este trabajo es conveniente describir brevemente en que consiste la técnica de angiografía fluoresceínica y el tipo de imágenes que proporciona. La angiografía fluoresceínica es un método ampliamente utilizada en oftalmología para examinar la circulación sanguínea en la retina. Consiste en inyectar un líquido fluorescente en la arteria cubital del paciente y esperar hasta que este líquido alcance la región vascular del ojo. Cuando el tinte llega a estos vasos sanguíneos se procede a tomar una serie de fotografías en instantes sucesivos de tiempo tal como indique el oftalmólogo. Tendremos pues una secuencia de imágenes en la que se observa cómo se van llenando las arterias y posteriormente las venas del colorante

fluorescente inyectado, es decir, el fenómeno que se observa es la trayectoria de la sangre en la retina del paciente. Por lo tanto, estas imágenes poseen una información muy valiosa para el estudio de todas aquellas patologías relacionadas con una mala circulación sanguínea.

El análisis de estas imágenes se ha aplicado al diagnóstico de múltiples enfermedades de la retina. La obstrucción arterial y venosa es un ejemplo de lesión que puede ser detectada mediante la angiografía fluoresceínica. Los síntomas observados con esta técnica son un estrechamiento claro de la arteria o vena principal de la retina acompañado de hemorragias en la retina que se visualizan aquí como manchas fluorescentes con origen en los vasos sanguíneos. La retinopatía hipertensiva es otra enfermedad que puede ser diagnosticada mediante esta técnica. Los síntomas principales que se pueden observar de ella son un estrechamiento arterial mientras que la vena se hace más gruesa; es decir, por un lado la variación de la relación entre la anchura de las venas y de las arterias junto con el endurecimiento de estas últimas, y por otro lado, el aumento de la tortuosidad en las venas.

Por último la retinopatía diabética, enfermedad en la que se centra nuestro estudio, es una de las principales y más comunes causas de la ceguera en los países industrializados. La duración de la diabetes es el principal factor en la aparición de la retinopatía diabética; dos tercios de los diabéticos de 15 o más años de evolución han desarrollado alguna forma de retinopatía. Los síntomas asociados con esta enfermedad son: la aparición de microaneurismas que se manifiestan en las angiografías fluoresceínicas como puntos muy brillantes, numerosas hemorragias y exudados lipídicos duros.

Desde el punto de vista de la Visión por Computador el problema del diagnóstico asistido se reformula como el reconocimiento de ciertos patrones que indiquen la presencia o no de una determinada enfermedad y el seguimiento temporal de dichos patrones.

Para poder realizar un estudio temporal de los elementos en la secuencia de imágenes un paso necesario es el alineamiento de la secuencia, es decir, la superposición de la imagen tomada en un instante dado con una versión girada, trasladada o deformada de la tomada en un instante posterior, de tal modo que las mismas estructuras anatómicas, y en general, cualquier punto de la imagen, coincidan perfectamente. Por lo tanto para el diagnóstico asistido de enfermedades y su evolución temporal existen dos tareas

básicas: el alineamiento de las imágenes y la detección de elementos que indiquen una determinada patología. Además de estas dos tareas fundamentales, en ciertos casos es interesante la segmentación de la imagen, entendida como la clasificación de cada punto de ésta en una de las componentes anatómicas relevantes que constituyen la imagen que se está analizando. Esta segmentación es de interés ya que permite localizar las lesiones no de forma absoluta sino con relación a dichos elementos anatómicos, y ello es a su vez lo que posee significado médico puesto que los elementos anatómicos varían en tamaño y posición entre pacientes oscilando siempre entre unos rangos considerados como normales.

Por todo lo dicho, este trabajo deberá incluir el desarrollo de algoritmos específicos para tratar imágenes de fondo de ojo que cubran las tres fases antes citadas: alineamiento, segmentación y detección de elementos patológicos concretos. Veamos con más detalle cada uno de éstos.

El **alineamiento** de imágenes de fondo de ojo ([Cid95, JB⁺90, NSR85]), es un paso crucial en el análisis de fondo ocular ya que todos los resultados posteriores dependerán de la comparación entre imágenes de la secuencia una vez alineadas. El alineamiento consiste en encontrar una transformación geométrica que aplique los puntos de una imagen en sus correspondientes en la siguiente imagen de la secuencia de forma que puntos correspondientes sean el mismo punto físico en la escena. La estimación del movimiento modelado como una transformación geométrica ha sido ampliamente estudiado en la literatura de Visión por Computador ya que es un problema que aparece en diversas aplicaciones reales como la inspección y automatización industrial, ensamblaje por robots, navegación de vehículos autónomos, ingeniería biomédica, etc. Existen dos diferentes enfoques para la estimación del movimiento: el primero es mediante el cálculo del flujo óptico y el segundo es a través de correspondencias de estructuras más o menos complejas extraídas de la imagen.

Los métodos que hemos estudiado para la resolución del problema del alineamiento en angiografías se basan en el segundo de los enfoques ya en que las imágenes angiográficas no existen objetos que se muevan en un fondo fijo sino que el ojo se mueve globalmente y además no existe la posibilidad de predecir el movimiento por dos motivos básicos: por un lado los movimientos del ojo son muy rápidos y aleatorios y por

otro lado cada imagen dentro de la secuencia se toma a intervalos de tiempo irregulares y bastante separados (no existe continuidad en el movimiento).

Estos métodos basados en la correspondencia intentan estimar la transformación geométrica entre imágenes consecutivas mediante el encaje de ciertas estructuras de bajo nivel extraídas previamente de las imágenes. Se han desarrollado 3 diferentes métodos para el alineamiento. El primero de ellos realiza un encaje de puntos relevantes (aspectos circulares y esquinas) mediante un procedimiento que busca correspondencias entre puntos que poseen un vecindario similar. Las características asociadas a cada punto utilizadas son la autocorrelación a diferentes distancias. El segundo método se basa en encajes de grafos relacionales mediante un procedimiento inspirado en el algoritmo EM (*Expectation-Maximization*). Se pretende aprovechar las restricciones impuestas por los grafos extraídos de las imágenes para la estimación de la transformación geométrica. En estos dos métodos se asume que la transformación geométrica entre sucesivas imágenes se puede modelizar por una transformación afín. En el tercero y último de los métodos desarrollados se pretende encontrar una transformación más compleja que sea capaz de modelizar el movimiento real del ojo humano teniendo en cuenta que las imágenes de fondo ocular son proyecciones planas de superficies que no lo son.

La segmentación ([MNSM89, CC89, LS93, TP98, IS99]) de las imágenes es, como dijimos, el segundo de los problemas que se plantean cuando se analizan imágenes de fondo de ojo. La segmentación es básicamente un proceso de clasificación de los *pixels* en la imagen; los *pixels* se asignan a diferentes clases lo que causa la división de la imagen en subconjuntos cuyos puntos tienen alguna característica en común. La segmentación de imágenes en los elementos anatómicos normales, en nuestro caso, venas, arterias, disco óptico y fovea, es uno de los pasos más importantes en el análisis automático de imágenes ya que en él los objetos de interés se extraen de la imagen para posteriormente ser descritos o medidos cuantitativamente. En el caso de las imágenes médicas, la segmentación tiene un doble valor; por un lado, como hemos mencionado, permite extraer los elementos más relevantes dentro de la imagen y así caracterizarlos y describirlos; por otro, es importante dar la posición relativa de una lesión detectada con respecto a alguno de los elementos anatómicos existentes. Los algoritmos de segmentación generalmente están basados en dos propiedades básicas de los valores de los niveles de gris: la discontinuidad y la similitud. En los primeros la imagen se divide

según los cambios bruscos de intensidad. Los segundos se basan en la umbralización, el crecimiento y la división de regiones.

Como, por supuesto, en las imágenes médicas se conoce normalmente alguna información *a priori* sobre el tipo y la posición espacial de las estructuras que aparecen en la imagen, es obvio que se puede y se debe aprovechar dicha información para realizar la segmentación de una imagen. La mayoría de los métodos existentes para imágenes médicas, y concretamente para angiografías fluoresceínicas, usan esta información como guía en el proceso de segmentación. Otro aspecto importante a destacar de las técnicas utilizadas en imágenes médicas para la segmentación es el uso de procedimientos parciales, en el sentido de que no se divide la imagen de una sola vez en todos los elementos relevantes que aparecen en ella sino que se utilizan algoritmos diferentes para cada una de las componentes de interés existentes en la imagen. Por ejemplo, en el caso de angiografías existen algoritmos de segmentación específicos para la fovea, para los vasos sanguíneos y para el disco óptico. El método que hemos desarrollado para la segmentación de angiografías intenta superar el inconveniente que supone la segmentación parcial. Este método intenta por un lado aprovechar la información *a priori* disponible, y por otro que la segmentación sea global. Su algoritmo se basa en modelos estocásticos y en Estadística Bayesiana. En particular, los modelos de Campos Aleatorios de Markov han demostrado ser herramientas apropiadas para el etiquetado de imágenes en un proceso de competición y cooperación local. La principal ventaja de estos métodos es que son capaces de utilizar toda la información disponible, tanto los datos de la imagen que se analiza como el tipo de elementos esperados en dicha imagen.

Una vez que la secuencia de imágenes ha sido alineada y segmentada podemos pasar al tercer y último paso en el análisis de imágenes de fondo ocular. Este tarea final consiste en la **detección y cuantificación** de los síntomas relacionados con la retinopatía diabética ([WTT89, Gro91, JAB⁺92, PSR⁺91, PRT⁺91, SPSF92, SOMH⁺96]). Entre todos los distintos síntomas indicativos hemos seleccionado la difusión macular por ser uno de los más significativos de la patología. Además, es la angiografía fluoresceínica la técnica que más resalta esta lesión de entre todas las imágenes de fondo ocular. La difusión macular, síntoma clave en el diagnóstico de la retinopatía diabética y la oclusión de los vasos sanguíneos, ha sido analizada por comparación de angiografías de una misma secuencia previamente alineadas. El método desarrollado trabaja sobre una imagen

diferencia obtenida a partir de un par de imágenes consecutivas. En esta imagen se observa claramente la difusión, ya que el resto de elementos fijos en la imagen se han eliminado de la imagen diferencia. El uso de morfología matemática sobre la imagen previamente binarizada cuantifica el área dañada, dando más importancia a las áreas más cercanas a la fóvea, ya que la difusión es tanto más peligrosa cuanto más próxima de la zona foveal esté, al ser ésta la de mayor agudeza visual.

1.4 Estructura del sistema

Las imágenes se procesan de forma secuencial en tres diferentes fases. La estructura propuesta se observa en la figura 1.1 . Estas tres fases serán analizadas en detalle en capítulos posteriores. A continuación se describe brevemente este sistema:

1. **Alineamiento de las imágenes:** Este es el primer paso a seguir antes de realizar el análisis de la secuencia. Las imágenes deben estar alineadas ya que todos los estudios relacionados con una evolución temporal necesitan de un proceso de comparación.
2. **Segmentación :** Los procedimientos de segmentación trabajan sobre las imágenes originales, sin alinear, por lo que esta tarea podría realizarse antes de que las imágenes estén alineadas. Por esta razón en la figura 1.1 que representa la estructura global de este trabajo el bloque correspondiente a la segmentación aparece al mismo nivel que la extracción de estructuras necesarias para el encaje. Además, la imagen segmentada puede utilizarse para la fase del diagnóstico sin necesidad del alineamiento de la secuencia en aquellas enfermedades en que no exista evolución temporal de ningún fenómeno.
3. **Detección de lesiones:** Esta tarea viene dirigida directamente por las necesidades reales de los expertos de área, en nuestro caso, los oftalmólogos. Son ellos los que deciden que patologías deberían ser estudiadas y los síntomas relacionados con esa enfermedad.

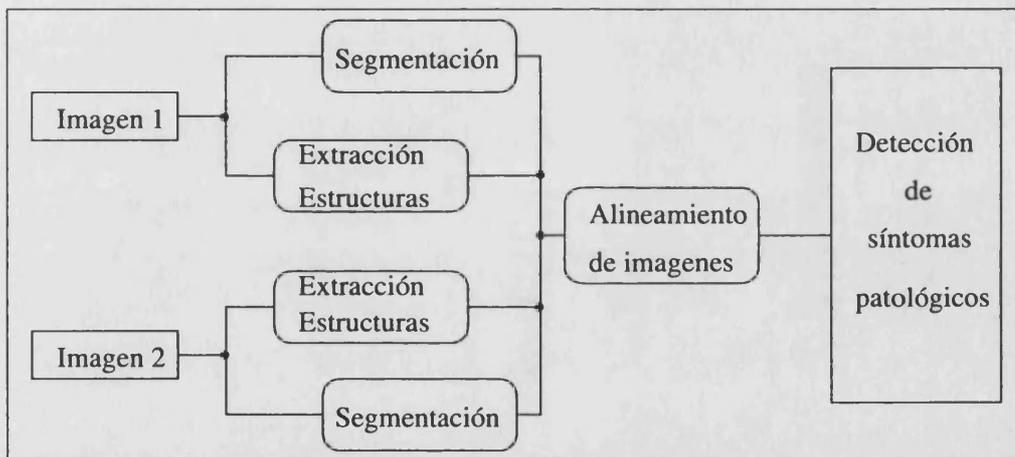


Figura 1.1: Estructura del sistema

Capítulo 2

Imágenes de fondo ocular

2.1 Introducción

En este capítulo se describe en detalle el aspecto y características de las imágenes con las que trabajaremos. Como dicho aspecto es el resultado tanto de los objetos (en este caso, estructuras anatómicas) que las produjeron, como de los instrumentos y procedimientos que se usan para su adquisición. Describiremos, pues, ambas cosas. Desde la experiencia adquirida en este trabajo podemos decir que un conocimiento lo más exhaustivo posible de la naturaleza, restricciones y características del problema ayuda enormemente al diseño de buenos algoritmos, orientados a la búsqueda y extracción de los aspectos verdaderamente relevantes del problema que se aborda. Para ello, veamos la anatomía del ojo, y los métodos para fotografiarlo.

2.2 El ojo humano

Las imágenes con los que trabajamos son imágenes que fotografían una de las partes más importantes del ojo humano, la retina, por lo que merece la pena hacer una descripción anatómica del globo ocular, con el objetivo de localizar el fin de nuestro estudio.

El ojo es el órgano de la visión en los seres humanos y en los animales. Los ojos de las diferentes especies varían desde las estructuras más simples, capaces de diferenciar

sólo entre la luz y la oscuridad, hasta los órganos complejos que presentan los seres humanos y otros mamíferos, que pueden distinguir variaciones muy pequeñas de forma, color, luminosidad y distancia. En realidad, el órgano que efectúa el proceso de la visión es el cerebro; la función del ojo es traducir las vibraciones electromagnéticas de la luz en un determinado tipo de impulsos nerviosos que se transmiten al cerebro.

2.2.1 Anatomía del ojo humano

El ojo en su conjunto, llamado globo ocular, es una estructura esférica de aproximadamente 2,5 cm de diámetro con un marcado abombamiento sobre su superficie delantera. La figura 2.1 muestra un esquema simplificado de las diversas partes que forman el globo ocular. Pasemos a ver una breve descripción de éstas partes y su papel dentro del sistema visual humano.

La parte exterior del ojo forma una especie de cubierta que se compone de tres capas de diversos tejidos:

- La capa más externa o esclerótica tiene una función protectora, cubre unos cinco sextos de la superficie ocular y se prolonga en la parte anterior con la córnea transparente;
- La capa media conocida como úvea tiene a su vez tres partes diferenciadas: la coroides, zona muy vascularizada, que continúa con el cuerpo ciliar, formado por los procesos ciliares, y a continuación el iris, que se extiende por la parte frontal del ojo.
- La capa más interna es la retina, que es la parte del ojo sensible a la luz.

La **córnea** es una membrana resistente, compuesta por cinco capas, a través de la cual la luz penetra en el interior del ojo. Por detrás, hay una cámara llena de un fluido claro y húmedo (el humor acuoso) que separa la córnea de la lente del **crystalino**. En sí misma, la lente es un doble casquete esférico aplanado constituida por un gran número de fibras transparentes dispuestas en capas. Está conectada con el músculo ciliar, que tiene forma de anillo y la rodea mediante unos ligamentos. El músculo ciliar y los tejidos circundantes forman el cuerpo ciliar y esta estructura aplanada o redondea la lente, cambiando su distancia focal. El **iris** es una estructura pigmentada suspendida

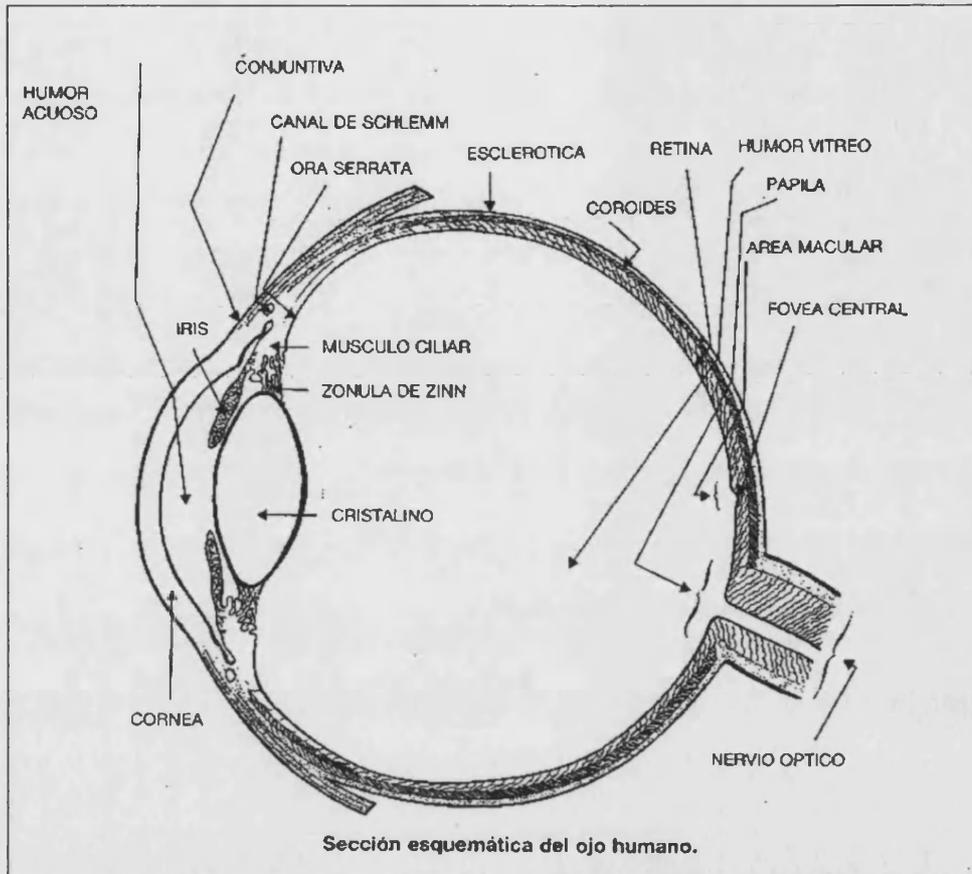


Figura 2.1: Anatomía del ojo humano

entre la córnea y el cristalino y tiene una apertura circular en el centro, la **pupila**. El tamaño de la pupila depende de un músculo que rodea sus bordes, aumentando o disminuyendo cuando se contrae o se relaja, controlando la cantidad de luz que entra en el ojo. Se puede decir que el iris constituye el diafragma del sistema óptico humano. Por detrás de la lente, el cuerpo principal del ojo está lleno de una sustancia transparente y gelatinosa conocida como **humor vítreo** que está encerrado en un saco delgado que recibe el nombre de membrana hialoidea. La presión del humor vítreo mantiene tenso el globo ocular.

La **retina** es una capa compleja compuesta sobre todo por células nerviosas. Las células receptoras sensibles a la luz se encuentran en su superficie exterior detrás de una capa de tejido pigmentado. Estas células son conocidas con el nombre de bastones y conos debido a su forma. Situada detrás de la pupila, la retina tiene una pequeña

mancha de color amarillo, llamada mácula lútea; en su centro se encuentra la fovea central que es la zona del ojo con mayor agudeza visual. La capa sensorial de la fovea se compone sólo de células con forma de conos, mientras que en torno a ella también se encuentran células con forma de bastones. Según nos alejamos del área sensible, las células con forma de cono se vuelven más escasas y en los bordes exteriores de la retina sólo existen las células con forma de bastones. El nervio óptico entra en el globo ocular por debajo y algo inclinado hacia el lado interno de la fovea central, originando en la retina una pequeña mancha redondeada llamada **disco óptico**. Esta estructura forma el punto ciego del ojo, ya que carece de células sensibles a la luz.

2.2.2 Fisiología del ojo humano

En general, los ojos de los animales funcionan como unas cámaras fotográficas sencillas. La lente del cristalino forma en la retina una imagen invertida de los objetos que enfoca y la retina se corresponde con la película sensible a la luz. Como ya se ha dicho, el enfoque del ojo se lleva a cabo debido a que la lente del cristalino se aplana o redondea; este proceso se llama acomodación. En un ojo normal no es necesaria la acomodación para ver los objetos distantes, pues se enfocan en la retina cuando la lente está aplanada gracias al ligamento suspensorio. Para ver los objetos más cercanos, el músculo ciliar se contrae y por relajación del ligamento suspensorio, la lente se redondea de forma progresiva.

Debido a la estructura nerviosa de la retina, los ojos ven con la mayor nitidez sólo en la región de la fovea. Las células con forma de conos están conectadas de forma individual con otras fibras nerviosas, de modo que los estímulos que llegan a cada una de ellas se reproducen y permiten distinguir los pequeños detalles. Por otro lado, las células con forma de bastones se conectan en grupo y responden a los estímulos que alcanzan un área general (es decir, los estímulos luminosos), pero no tienen capacidad para separar los pequeños detalles de la imagen visual. La diferente localización y estructura de estas células conducen a la división del campo visual del ojo en una pequeña región central de gran agudeza y en las zonas que la rodean, de menor agudeza y con una gran sensibilidad a la luz. Así, durante la noche, los objetos confusos se pueden ver por la parte periférica de la retina, aun cuando son invisibles para la fovea

central. El mecanismo de la visión nocturna implica la sensibilización de las células en forma de bastones gracias a un pigmento, la púrpura visual o rodopsina, sintetizado en su interior. La rodopsina se blanquea por la acción de la luz y los bastones deben reconstituirla en la oscuridad, de ahí que una persona que entra en una habitación oscura procedente del exterior con luz del sol, no puede ver hasta que el pigmento no empieza a formarse; cuando los ojos son sensibles a unos niveles bajos de iluminación, quiere decir que se han adaptado a la oscuridad.

Los ojos están en constante movimiento y la retina se excita en una u otra parte, según la atención se desvía de un objeto a otro. Los movimientos del globo ocular hacia la derecha, izquierda, arriba, abajo y a los lados se llevan a cabo por los seis músculos oculares y son muy precisos. Se ha estimado que los ojos pueden moverse para enfocar en, al menos, cien mil puntos distintos del campo visual. Los músculos de los dos ojos funcionan de forma simultánea, por lo que también desempeñan la importante función de hacer converger sus ejes ópticos en un punto para que las imágenes de ambos coincidan; cuando esta convergencia no existe o es defectuosa se produce la doble visión. El movimiento ocular y la fusión de las imágenes también contribuyen en la estimación visual del tamaño y la distancia.

2.3 Imágenes de fondo de ojo

Las imágenes de fondo ocular son un tipo de imágenes médicas ampliamente utilizadas para el diagnóstico de enfermedades oculares. Estas imágenes muestran la retina, el tejido sensitivo del ojo que traduce las imágenes ópticas que vemos en impulsos electroquímicos que nuestro cerebro puede procesar. La retina puede ser considerada como la película fotográfica del sistema óptico humano.

En la sección anterior se ha visto la función desempeñada por las células fotorreceptoras junto con su localización en la capa retiniana. Puesto que la distribución de conos y bastones en la retina no es uniforme, el daño causado en la retina por una determinada patología será diferente dependiendo de la posición relativa dentro de ella. De ahí el interés en su localización; en general, una lesión será tanto más peligrosa cuanto más cerca se encuentre de la fovea.

La retina puede fotografiarse directamente utilizándose la pupila como entrada y salida para el haz de rayos. La utilización de la información contenida en estas imágenes para el seguimiento, diagnóstico y tratamiento de ciertas enfermedades oculares ha demostrado ser una herramienta práctica y fiable.

Existen diversos métodos para la obtención de imágenes de fondo de ojo. Estos métodos se pueden clasificar en dos tipos: los que se basan en la elección del tipo de luz usado como fuente y los que se basan en la inyección de colorantes en la sangre para resaltar ciertas estructuras anatómicas. Todas estas técnicas dan como resultado imágenes de retina en las que se visualizan de formas distintas los elementos de ésta; de ahí que la elección de una de estas técnicas depende fundamentalmente del tipo de enfermedad que se pretende detectar.

Antes de continuar con la explicación detallada de estas técnicas, nos detendremos en la descripción del sistema óptico utilizado en la captura de las imágenes de fondo ocular.

2.3.1 ¿Qué es una cámara de fondo de ojo?

Una cámara de fondo de ojo es un microscopio de baja potencia junto con una cámara fotográfica (o cualquier otro instrumento que permita capturar imágenes como cámara de video). Su diseño óptico se basa en un oftalmoscopio indirecto. Estas cámaras se caracterizan por el ángulo de visión que es el ángulo de apertura de campo de la lente. Este parámetro de la lente está relacionado directamente con la ampliación de la imagen obtenida con respecto a la escena real. Cuanto más amplio es el ángulo de visión menor nivel de detalle se obtiene en la imagen. Por contra, cuanto más pequeño es el campo de visión más pequeña es el área de la escena fotografiada y el nivel de detalle se incrementa.

La figura 2.2 muestra las diferentes áreas de la retina que se fotografían en función de la amplitud de campo de la cámara utilizada. La figura 2.3 muestra tres imágenes de fondo de ojo tomadas con cámaras con distinta apertura de campo. En la primera de ellas se observa un campo de 45 grados, en la segunda el campo de visión se ve reducido a 30 grados, lo que se considera un campo normal de visión, y en la tercera y última se

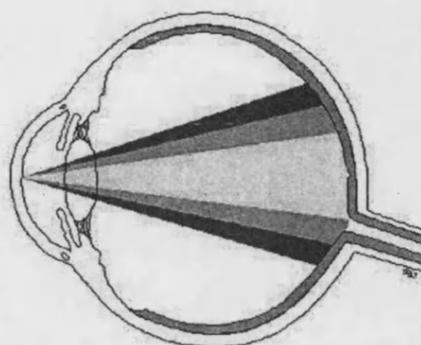


Figura 2.2: Variación de parte de la escena fotografiada por una cámara según su amplitud de campo



Figura 2.3: Imágenes de fondo de ojo con distinta apertura de campo. Izquierda: 45 grados. Centro: 30 grados. Derecha: 20 grados

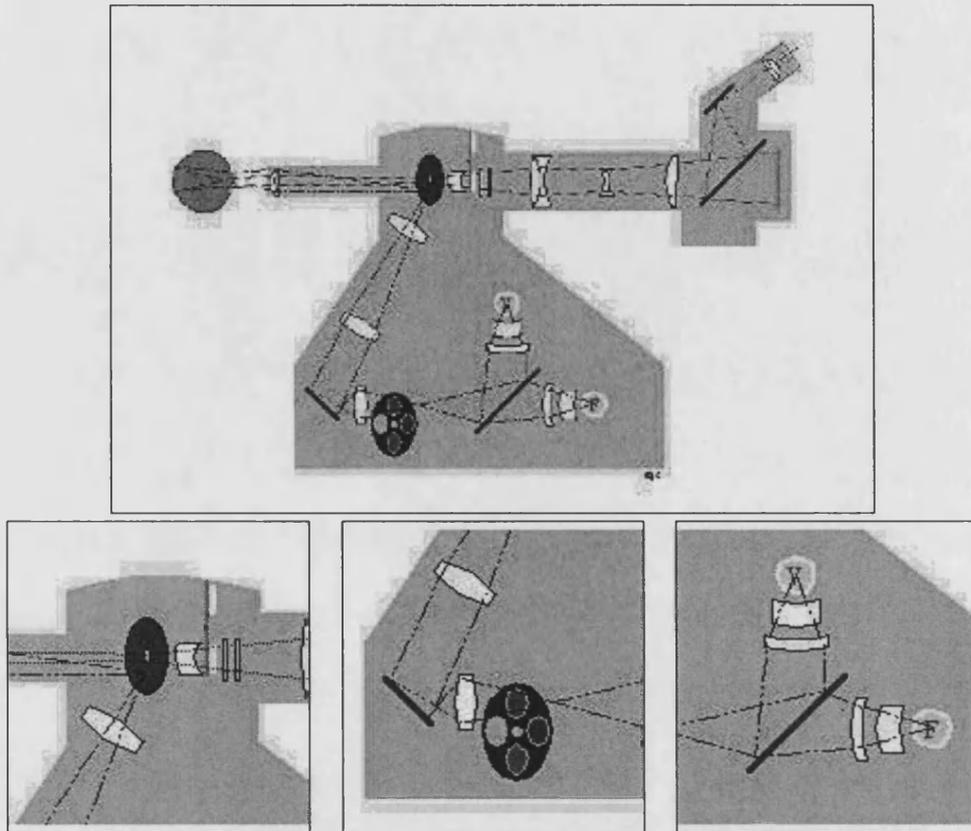


Figura 2.4: Sistema óptico utilizado en una cámara de fondo. Arriba: Sistema total de lentes. Izquierda: Primeras lentes en el camino de la luz. Centro: Filtros. Derecha: Fuente de luz

muestra una imagen en la que el campo es mínimo (20 grados).

2.4 Sistema óptico de las cámaras de fondo ocular

El sistema óptico total de una cámara de fondo de ojo ([Sai]) está compuesto por un conjunto de lentes y espejos junto con el sistema de lentes de la cámara utilizada. Para la descripción del sistema se ha supuesto que la cámara es una cámara fotográfica reflex típica. Sin embargo, cualquier otro tipo de cámara puede servir para montar un sistema óptico equivalente.

En la figura 2.4 aparece un esquema con todos los elementos que componen el sistema. Veamos paso a paso el camino seguido por el haz de luz generado por la fuente

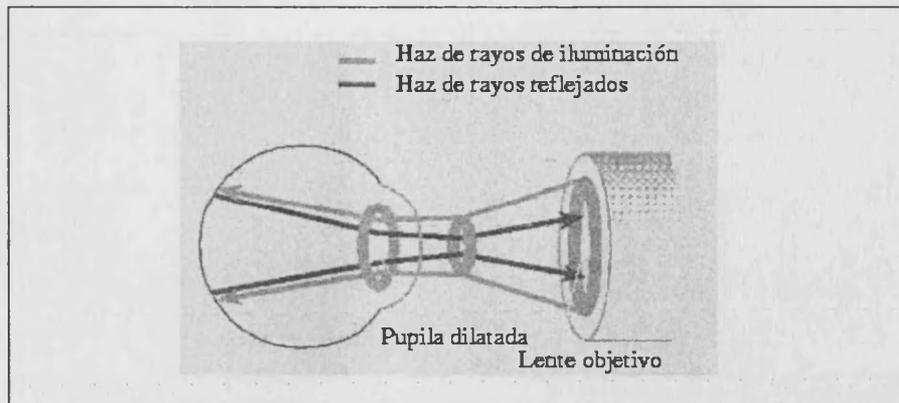


Figura 2.5: Haz de luz entrando y saliendo a través de la pupila

hasta que se forma la imagen en la película fotográfica. La luz generada por el flash electrónico o por una lámpara se proyecta a través de un conjunto de filtros sobre un espejo circular. Una ampliación de los elementos del sistema utilizados como fuente de luz se muestra en esa misma figura. Los filtros que se colocan en el camino de la luz tienen como objetivo eliminar ciertas longitudes de onda del haz de luz proyectado. La utilización de unos filtros u otros viene dado por el elemento que quiere ser visualizado con un mayor contraste en la imagen resultado. La influencia de cada uno de estos filtros en la imagen resultado de fondo de ojo se explica en la sección 2.5.

En el camino de la luz se sitúan un conjunto de lentes y espejos mediante los cuales la luz es dirigida hacia la entrada del globo ocular. La última de las lentes situada en el camino del haz de luz antes de que entre en el ojo, cumple dos papeles: por un lado enfoca la luz hacia un espejo, que, a su vez la reflejará hacia el fondo del ojo, y por otro se comporta como una máscara que modela los rayos en forma de corona circular. Estos rayos en forma de corona circular se proyectan hacia un espejo que reflejará la luz hacia la córnea del ojo humano.

Suponiendo que el sistema óptico y el globo ocular están correctamente alineados y enfocados, la imagen resultado de la retina sale de la córnea a través de la parte central, no iluminada de la corona circular. La figura 2.5 muestra la entrada y salida del haz de luz a través de la parte más externa de la corona circular y su parte central, respectivamente.

Una vez que se tiene el haz de luz imagen procedente de la retina es necesario

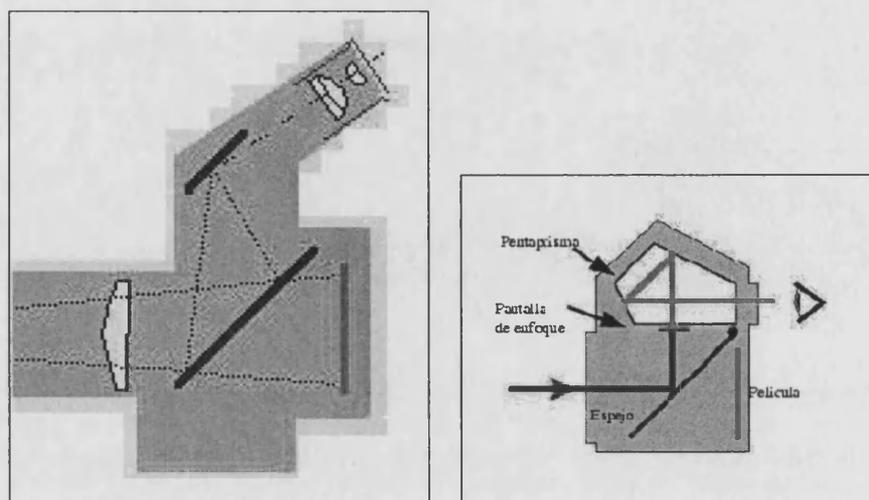


Figura 2.6: Izquierda: Lentes y espejos que dirigen la luz imagen hacia la película fotográfica. Derecha: Sistema óptico de una cámara reflex

dirigirlo hacia la película fotográfica donde se capturará la imagen. La luz continúa a través de la apertura central del espejo que hemos descrito, a través del instrumento de corrección de astigmatismo y las lentes de compensación de dioptrías, y finalmente llega al sistema de lentes de la cámara reflex donde el haz de luz de la imagen se recoge en la película fotográfica.

2.5 Tipos de imágenes de fondo de ojo

Existen numerosas técnicas para la captura de imágenes de fondo ocular. Como ya se ha mencionado, estas técnicas se pueden clasificar en dos grupos: las que utilizan algún tipo de líquido de contraste inyectado en las arterias del paciente y las que se basan en la aplicación de filtros para la eliminación de ciertas longitudes de onda de la fuente de rayos luminosos. Nos referiremos a las primeras como angiografías y a las segundas como fotografía monocromática. Veamos en detalle los principios en los que se basan cada una de ellas.

2.5.1 Fotografía de fondo monocromática

La fotografía de fondo monocromática es un término empleado para designar las técnicas oftalmoscópicas que se ayudan del uso de iluminación de color o monocromática para la captura de imágenes de fondo ocular. Es en 1925 cuando se describe por primera vez el uso de luz libre de rojo en oftalmoscopia para resaltar el contraste visual de detalles anatómicos del fondo. Limitando el rango espectral de la fuente de iluminación, la visibilidad de varias estructuras del fondo puede ser resaltada. Esta técnica es mucho más efectiva cuando se combina con películas fotográficas de blanco y negro de alta resolución. El contraste se puede manipular para la visualización óptima de detalles específicos del fondo de ojo.

La fotografía monocromática de fondo de ojo se basa en dos principios de la luz y la fotografía: el uso de los filtros de contraste para alterar los tonos en fotografías de blanco y negro y el incremento de la dispersión de la luz de longitud de onda corta.

Los filtros de contraste ajustan la tonalidad monocromática de diferentes colores introduciendo diferencias de intensidad entre colores que normalmente se reproducirían como tonos similares de gris. Fotógrafos profesionales han utilizado filtros de contraste durante muchos años para ajustar los tonos de las imágenes haciéndolas habitualmente más bruscas. Un determinado color aparecerá más brillante cuando se fotografía a través de un filtro de su mismo color y más oscuro cuando se utiliza un filtro de su color complementario. Por ejemplo, un objeto rojo aparecería más brillante si se expone a través de un filtro rojo; este mismo color aparecerá más oscuro si se fotografía a través de un filtro de color azul celeste (ciano).

El espectro visible de la luz puede dividirse en tres partes: la luz de alta frecuencia (longitud de onda corta), la intermedia, y la de baja frecuencia (azul, verde y rojo, respectivamente). Cada uno de los filtros rojo, verde y azul transmiten aproximadamente una tercera parte de la totalidad de la luz blanca y bloquean las otras dos partes.

Hemos mencionado que muchas veces puede ser ventajosa la dispersión de la luz visible de alta frecuencia (color azul). Este efecto incrementa la visibilidad de las capas anteriores de la retina. Si la claridad del medio ocular se ve comprometida, la dispersión puede tener un efecto adverso en las imágenes de fondo de ojo. Cuando la luz se dispersa

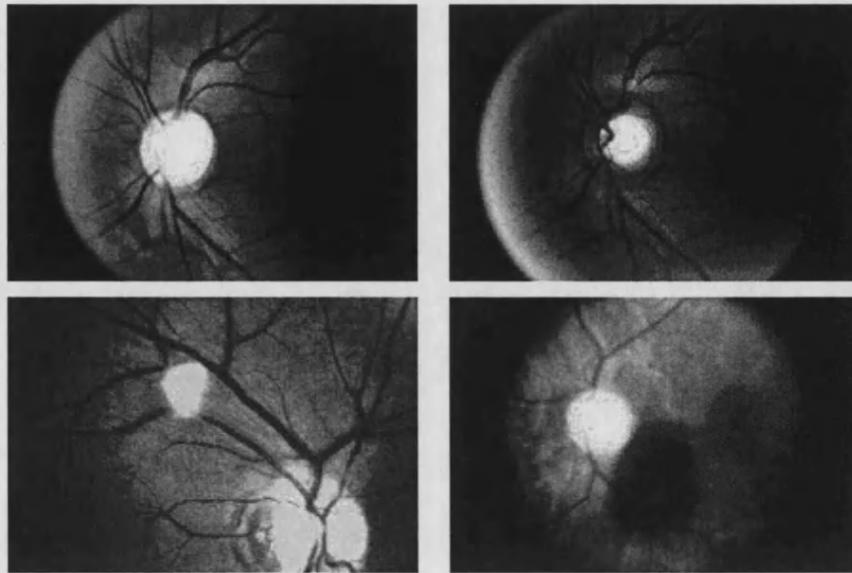


Figura 2.7: Aplicación de diversos tipos de filtros en la captura de imágenes de fondo de ojo. Arriba-Izquierda: luz blanca. Arriba-Derecha: luz azul. Abajo-Izquierda: luz verde. Abajo-Derecha: luz roja

antes de que alcance la retina, la imagen puede aparecer bastante borrosa con luz de alta frecuencia.

Las imágenes que aparecen en la figura 2.7 demuestran los principios antes mencionados y sus efectos sobre imágenes de fondo de ojo. Se han aplicado diversos filtros de diferente color. Se observa que dependiendo del color del filtro utilizado algunas estructuras del fondo ocular se hacen más o menos visibles.

La luz de color azul del espectro visible es absorbida por los vasos sanguíneos de la retina, por los vasos coroidales y el nervio óptico. Por ello, cuando se utiliza un filtro azul que elimina el resto de frecuencias del espectro, la imagen proporcionada tendrá un fondo oscuro sobre el cual las reflexiones especulares y la dispersión en las capas anteriores del ojo se ven resaltadas. Sin embargo, la dispersión en el medio ocular puede limitar la efectividad de esta longitud de onda. Ejemplos de estructuras que se hacen más evidentes en las imágenes de fondo de ojo tomadas con un filtro azul son: la capa del nervio óptico, la membrana epiretinal, etc. Como se ha visto la luz de muy alta frecuencia puede tener un efecto negativo por su excesiva dispersión por lo que normalmente se emplean filtros azul celeste de una longitud de onda un poco mayor

(490 nm frente a las 450 nm de la luz azul). La luz verde también es absorbida por la sangre, pero se refleja más por el epitelio pigmentoso de la retina que la luz azul. Existe una menor dispersión que con longitudes de onda más corta, por lo que las opacidades del medio tienen un efecto menos perjudicial.

Este tipo de luz proporciona un excelente contraste y una mejor vista global del fondo de ojo. Entre las estructuras resaltadas destacan las zonas vasculares de la retina y las hemorragias. Por esta razón estas imágenes junto con las angiografías fluoresceínicas son las técnicas rutinarias más empleadas cuando se necesita realizar una exploración general de la capa retiniana.

El uso de filtros rojos convierte el epitelio pigmentoso retiniano en más transparente revelando una mejor vista de la zona coroidal. El contraste global se ve claramente reducido con este tipo de filtro; los vasos sanguíneos de la retina aparecen más brillantes y se hacen menos obvios conforme aumenta la longitud de onda de la luz de la fuente utilizada. Además existe una mayor separación de intensidad entre venas y arterias (las arterias se ven más luminosas que las venas). Por otro lado el nervio óptico se hace también más brillante y sin rasgos distintivos. En resumen, se puede decir que este tipo de luz es muy útil para fotografiar ciertas clases de trastornos de los pigmentos de la retina, rupturas coroidales y melanomas malignos.

2.5.2 Angiografías

Las angiografías son imágenes de fondo de ojo caracterizadas por la inyección de líquidos colorantes en los vasos sanguíneos del paciente ([Bay]). Estos tests son muy útiles para detectar la difusión de la sangre fuera de su circuito normal y los vasos sanguíneos dañados. El colorante circula a través del sistema circulatorio hasta alcanzar tanto los vasos en la retina como aquellos de la capa de tejido situada más profundamente llamada coroides. Ninguno de estos test involucran el uso de rayos X u otro tipo de radiación dañina.

La fluoresceína es un tinte amarillo que brilla en la luz visible. La indocianina es un tinte verde cuya fluorescencia se manifiesta en el rango de luz infraroja, no visible para el ojo humano; por ello, se necesita una cámara especial sensible a este tipo de

radiación. La angiografía de indocianina verde recientemente se está convirtiendo en una técnica habitual gracias a que las cámaras necesarias para esta práctica son cada vez más accesibles.

Estos dos tipos de imágenes de la retina pueden ayudar a los especialistas a diagnosticar y estudiar la evolución de una determinada lesión ocular. El colorante fluoresceínico es mejor para estudiar la circulación en la retina mientras que la indocianina es más útil para estudiar los vasos sanguíneos de la coroides, que se encuentran en una capa más profunda. Para algunas enfermedades que afectan principalmente a la circulación sanguínea de la retina, como la retinopatía diabética y la retinopatía vascular, se utiliza la angiografía fluoresceínica. Para otro tipo de lesiones, como la degeneración macular, donde la difusión se produce de los vasos de la capa coroidal, los dos exámenes pueden ser de gran utilidad. Además este último se utiliza cuando existe difusión de sangre que puede hacer difícil la interpretación de la angiografía fluoresceínica.

Las angiografías son, por lo tanto, imágenes de fondo de ojo tomadas después de la inyección de un líquido de contraste en la vena cubital del paciente. Como ya hemos dicho, las angiografías permiten estudiar la distribución del colorante en la retina, y si, además, se toma una secuencia de imágenes, podremos estudiar la evolución temporal. Cada una de las imágenes que componen la secuencia ha sido tomada en un instante clave a criterio del especialista.

Historia de la angiografía fluoresceínica

A finales de los años 50, el estudiante de medicina Harold Novotny y el doctor David Alvis trabajaban en el departamento de medicina de la Universidad de Indiana en determinar la saturación de oxígeno en las arteriolas de la retina mediante un método fotográfico.

Novotny comenzó a pensar en la posibilidad de introducir algún tipo de colorante a través del sistema circulatorio para determinar los niveles de saturación. Se pensó que la observación de fluorescencia podría hacer más fácil esta tarea, y el sodio fluoresceínico fue uno de los tintes utilizados.

Mientras tanto, Alvis mezcló una gota de su propia sangre con sodio fluoresceínico

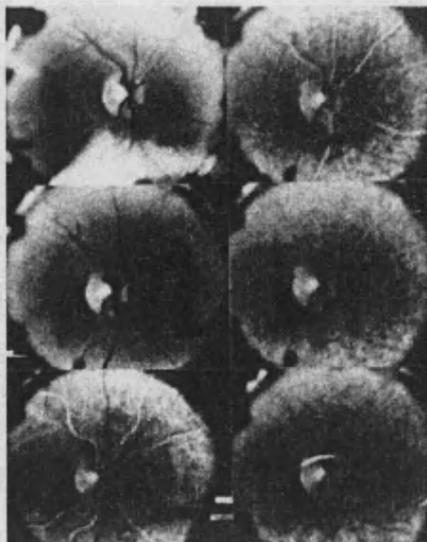


Figura 2.8: Primeras angiografías fluoresceínicas realizadas

y envió la mezcla para que le realizaran un análisis en un espectrofluorómetro que determinó la longitud de onda para la que la emisión era máxima. Así fue como estos médicos comenzaron a realizar estudios fluoresceínicos en pacientes diabéticos e individuos con hipertensión. En 1961 apareció la primera descripción de esta técnica en *Circulation* ([NA61]).

2.6 Enfermedades que afectan a la retina

2.6.1 Retinopatía diabética

La retina es una delicada membrana situada en la pared posterior del ojo. En ella se encuentran, como ya hemos indicado, gran cantidad de células fotorreceptoras las cuales captan las imágenes que entran a través de la pupila. Estas imágenes se envían en forma de señales electroquímicas a través del nervio óptico al cerebro que las interpretará posteriormente.

La diabetes provoca una hiperglucemia mantenida a través de los años que conduce a graves alteraciones vasculares. En los pequeños vasos de la retina (y otros órganos como el riñón) se produce la llamada microangiopatía diabética.

Las dilataciones de los vasos capilares, clásicamente llamados microaneurismas (MA) son probablemente las primeras lesiones observables en la retinopatía diabética. Únicamente los MA de tamaño mayor que 30 micras de diámetro son identificables en imágenes en color del fondo ocular. Puesto que las angiografías fluoresceínicas proporcionan una mejor definición los MA pueden ser detectados como diminutos puntos fluoresceínicos que se llenan en las primeras imágenes de la secuencia de angiogramas.

La retinopatía diabética (RD) sigue siendo la causa más frecuente de ceguera en los países desarrollados. Entre los mecanismos que pueden producir pérdida de visión en estos enfermos, hay dos que derivan de la propia enfermedad, como son la retinopatía proliferante con hemorragia vítrea y el edema macular diabético

La primera de estas formas se considera muy seria y conduce a la ceguera del paciente si no se trata a tiempo. Los enfermos diabéticos deberían conocer que esta manifestación de la enfermedad es prácticamente indolora y sus síntomas inexistentes. Por lo tanto, es esencial que se sometan a exámenes periódicos por los especialistas.

En la retinopatía diabética proliferante, algunos de los vasos sanguíneos de la retina se cierran, dejando grandes áreas de la retina sin los nutrientes necesarios (isquemia retiniana). El cierre de estos vasos sanguíneos es también bastante serio, y paradójicamente, la gravedad se intensifica cuando comienzan a desarrollarse nuevos vasos sanguíneos anormales para compensar la pérdida del cierre de vasos. Éstos son muy peligrosos para la salud del ojo ya que pueden crecer anárquicamente hacia el vítreo y no nutren la retina de forma adecuada; además pueden conducir a otros problemas:

- Sangrar hacia la cavidad donde se encuentra el humor vítreo y bloquear el paso de luz (Hemorragia vítrea que produce la pérdida súbita de la visión).
- Formar tejido cicatrizante que puede estirar la retina hacia el vítreo (desprendimiento de retina tradicional).
- Producir proliferación de vasos en el iris bloqueando la salida de humor acuoso y dando lugar el denominado glaucoma neovascular (fase terminal de la retinopatía diabética).

Todos estos problemas pueden conducir a una pérdida de visión e incluso pueden causar ceguera, si no son controlados.

En la segunda forma de manifestación de la enfermedad, el edema macular, los vasos de la retina cambian y se debilitan. Algunos vasos decrecen en tamaño, mientras que otros se hacen más grandes y terminan en estructuras globulares. Estos vasos débiles tienden a tener hemorragias y difunden sangre y suero. Esta difusión en la retina provoca una disminución de la agudeza visual si se sitúa en la parte central de la macula denominada fovea.

El edema macular diabético se define actualmente como un área de engrosamiento retiniano situado a menos de un diámetro papilar del centro de la mácula. Se produce esencialmente por una alteración en el endotelio capilar (capa más interna del vaso), dando lugar a una salida de plasma y componentes hemáticos hacia la retina circundante. Estas zonas de salida o escape de sustancias se ponen de manifiesto mediante Angiografía Fluoresceínica, ya que producen extravasación del colorante fuera del vaso. Se producen a partir de lesiones propias de la enfermedad microvascular, como son los microaneurismas.

El edema macular es la causa más frecuente de pérdida de visión en pacientes con retinopatía diabética de fondo. Admitiendo la definición dada antes, su tasa de prevalencia se sitúa aproximadamente en un 14% de la población diabética total; de éstos, un 40% tienen afectación del centro de la mácula, lo cual conlleva un pronóstico visual mucho peor.

Tanto la retinopatía diabética proliferante como el edema macular se tratan mediante la retino-fotocoagulación con laser.

- En la RD proliferante se trata toda la retina isquémica que es el estímulo para la formación de neovasos.
- En el edema macular los impactos de laser se dirigen selectivamente hacia las lesiones que provocan la salida de sustancias fuera del vaso, como son los microaneurismas.

2.6.2 Trombosis venosa

La trombosis venosa es una enfermedad común de los vasos retinianos con posibles complicaciones de ceguera. Siguiendo a la retinopatía diabética, la oclusión venosa retiniana es la enfermedad vascular más común. El edema macular es la principal

causa de pérdida de visión en la oclusión de venas de la retina. La trombosis venosa retiniana se produce debido a la oclusión de una vena por un coágulo sanguíneo. Es más frecuente a partir de los 60 años y en enfermos con factores de riesgo cardiovascular como son la hipertensión arterial, la arteriosclerosis, el tabaquismo o la tasa elevada de colesterol en sangre.

2.6.3 Neovascularización subretiniana

El término neovascularización subretiniana (NSR) designa al crecimiento de nuevos vasos sanguíneos bajo la retina. Los vasos pueden caer entre la coroides y el epitelio pigmentoso de la retina (EPR) y la retina. NRS es una anomalía encontrada en muchas enfermedades de la retina. Las dos enfermedades que más frecuentemente provocan neovascularización subretiniana son la degeneración macular ligada a la edad y la degeneración macular en la miopía elevada.

Capítulo 3

Alineamiento de angiografías fluoresceínicas

3.1 Introducción

La angiografía fluoresceínica es un procedimiento muy común en oftalmología, principalmente para evaluar la retinopatía vascular (retinopatía diabética, oclusión de venas). La técnica consiste en la inyección de fluoresceína sódica en la vena cúbica del brazo. Después de 8-10 segundos, el colorante fluoresceínico alcanza la zona vascular del ojo y su distribución se puede observar y fotografiar a través de un filtro adecuado que mostrará el fenómeno fluoresceínico. Sólo los vasos sanguíneos anormales difunden fluoresceína sódica, causando un progresivo incremento de fluorescencia en ciertas áreas. Esta exploración permite al experto planificar y evaluar el tratamiento láser. Cuando se toman decisiones y se hace un diagnóstico a partir de estas imágenes, una inspección visual directa podría no ser suficientemente fiable y un análisis cuantitativo más preciso podría ser necesario.

Sin embargo, una fijación de la mirada del paciente no constante y el movimiento natural de su ojo a menudo provoca que éste y la cámara no estén alineados, causando variaciones de iluminación y de posición de la imagen de fondo de ojo. Estos cambios en iluminación y posición pueden alterar significativamente los resultados del análisis cuantitativo de la imagen. Como Rosenfeld and Kak (1988) dicen: "Dadas dos imágenes

de una misma escena tomadas en diferentes instantes, si se registran podemos determinar los puntos en los que difieren, y así analizar los cambios que han tenido lugar.” La AF es un ejemplo médico importante de esta situación. Por lo tanto, el encaje de imágenes es un paso crucial en el análisis de una AF digitalizada para conseguir una buena precisión y una velocidad aceptable en el proceso del diagnóstico.

3.1.1 Descripción del problema, objetivos y revisión de la literatura

Nótese que estamos tratando con la rotación de una superficie curva (el fondo ocular). La transformación entre imágenes consecutivas de este tipo de secuencia no es exactamente el movimiento de un sólido rígido. La estimación del movimiento, modelado como una transformación lineal ha sido un tema común de investigación, y varios enfoques se han propuesto para ello. Es un problema que surge en el procesamiento de imágenes médicas, particularmente en AF (Goldbaum 1990 [GK⁺90]; Todd-Pokropek and Viergever, 1992 [TPV92]), cuando se intenta comparar varias imágenes de una misma secuencia.

Cada imagen de la secuencia que se va a analizar, que es una proyección plana de la superficie curva del fondo ocular, ha sido tomada a intervalos de tiempo irregulares, ya que es el doctor quien decide cuándo es interesante tomar una imagen. El paciente puede realizar un movimiento errático del ojo entre una imagen y la siguiente. El nivel de fluorescencia también cambia considerablemente a lo largo de la secuencia. Primeramente, los vasos coroidales se llenan creando una fluorescencia del fondo de la imagen. Uno o dos segundos más tarde, la fluoresceína llena las arterias de la retina y después llega a las venas de ésta. Esto crea cambios bruscos de iluminación afectando a regiones de tamaño considerable. A diferencia de otros problemas en la estimación del movimiento en el que uno o más objetos se desplazan sobre un fondo estático, en nuestro caso toda la imagen puede contener información importante relacionada con el movimiento. La oclusión de estructuras que aparecen en la imagen por parte de otras no ocurrirá pero algunas de ellas pueden dejar de verse y otras pueden aparecer, debido a que el movimiento hace que algunas partes entren y salgan del campo de visión de la cámara. Es importante indicar que el área más interesante en el análisis del fondo de ojo es la zona de la fovea y el área que la rodea. Esta región se encuentra localizada

alrededor del centro de la imagen, y es allí donde mayor precisión hace falta.

Ya que el objetivo final de la AF es el análisis de los cambios que han tenido lugar en cada punto de la imagen, es obvio que el primer paso es deshacer el movimiento del globo ocular. Sólomente de esta forma la comparación de los niveles de gris en un determinado pixel de diferentes imágenes tiene sentido.

El problema que estamos considerando ha sido tratado ampliamente en la literatura de Visión por Computador. Sin embargo las soluciones a los problemas derivados de su aplicación a técnicas de exploración son escasas. Los principales métodos de encaje de angiografías son:

- Alineamiento manual ([PRT⁺91], [PL86], [TNSW87]),
- Correlación cruzada ([NSR85]),
- Métodos basados en el encaje de ciertas estructuras extraídas de la imagen ([RJ⁺93], [RMP⁺96]).

El alineamiento manual es una tarea muy laboriosa, lo que representa un claro inconveniente para un protocolo clínico. La mayoría de las referencias en el encaje de angiografías de fondo de ojo utilizan este tipo de método.

En cuanto al uso de la correlación cruzada tal como apunta Rangarajan y colaboradores ([RMP⁺96], "El método no es robusto a cambios en los patrones de intensidad...", lo cual es nuestro caso, ya que la difusión del contraste produce cambios severos de este tipo.

El método que hemos desarrollado para el encaje de imágenes de la secuencia se clasifica dentro de los basados en la detección de puntos de interés. Una aportación previa de este tipo corresponde a los métodos dados por Jagoe y Blauth ([RJ⁺93]. Estos utilizan como puntos relevantes (marcas) los cruces de venas, para cuya detección se diseña y prueba un procedimiento específico. Las correspondencias de puntos se establecen entre triplete de puntos usando una medida de similitud entre los triángulos que forman. Además se utiliza información adicional asociada con cada vértice. La transformación final generada es descrita mediante polinomios de primer orden. Para desechar correspondencias erróneas se utilizan ciertas suposiciones bastante razonables sobre las imágenes iniciales: la distancia entre puntos correspondientes debe ser menor

que un cierto límite y la diferencia entre los ángulos que forman los triángulos también debe ser menor que un valor umbral. Otra restricción usada se aplica sobre la longitud de los lados que forman los triángulos: los lados deben ser mayores que un cierto valor para evitar agrupamientos locales.

Una contribución más reciente en un contexto médico diferente pero con un enfoque similar es la propuesta por Rangarajan. Su grupo trabaja en el alineamiento automático de radiografías planteado como problema de optimización. La transformación geométrica y las correspondencias se estiman de la siguiente forma: la transformación, minimizando una función objetivo usando un método de continuación y las correspondencias, mediante una doble normalización de la matriz de correspondencias. Como estableceremos más tarde, también nosotros actualizamos una matriz de correspondencias y estimamos la transformación geométrica en turnos, pero utilizamos un procedimiento de predicción basado en un modelo estocástico en lugar de solucionar un problema de optimización. Una descripción detallada del método de optimización de utilizado por Rangarajan se encuentra en [RMP⁺96].

3.1.2 Descripción del método

El procedimiento que proponemos consiste en 4 pasos sucesivos. En el primero, un conjunto de puntos relevantes se detecta automáticamente en cada una de las imágenes de la secuencia. En el segundo, un vector de características locales se asocia a cada uno de estos puntos (en nuestro caso, los valores de la función de autocorrelación). En el tercero, unas pocas correspondencias entre puntos relevantes de diferentes imágenes se establecen usando las distancias euclídeas entre puntos, y medidas de similitud entre vectores de características de las distintas imágenes. Y por último, en el cuarto, un procedimiento iterativo basado en un modelo estocástico mejora el conjunto inicial de correspondencias y calcula un estimador de la transformación lineal asumida. Cada uno de estos pasos se describirá a continuación.

Necesitamos justificar la elección de puntos relevantes como primitivas para el encaje en lugar de bordes u otros tipos de estructuras de más alto nivel (venas u otras estructuras anatómicas). Cambios locales y globales de intensidad aparecen en nues-

tras imágenes; los cambios globales se deben a que el colorante fluoresceínico alcanza la retina. Pero además aparecen cambios locales, debidos a situaciones patológicas como la difusión. Esto difumina algunos contornos que no ya no serán detectados y hace que otros aparezcan a lo largo de la secuencia. Pero los cruces de venas, encontrados mediante un detector de puntos relevantes, son normalmente suficientemente robustos. Además, la asociación de características locales invariantes a estos cambios es más fácil con puntos de interés que con bordes. Y finalmente, los movimientos del fondo ocular dejan importantes regiones fuera del campo de visión, lo que tiene significativas consecuencias para los bordes, ya que cualquier medida entre ellos se verá afectada severamente (por corte de los límites). El correspondiente efecto en puntos (aparición y desaparición de ellos) será elegantemente tratado en nuestro algoritmo.

Como se mostrará, alrededor de 30 puntos relevantes pueden ser fácilmente detectados en esta clase de imágenes por el método que hemos empleado. Aproximadamente una tercera parte de ellos tienen correspondencia en la imagen siguiente. Probaremos que esto es suficiente para la estimación del movimiento.

Para cumplir nuestros objetivos, algunos aspectos y sus correspondientes métodos deben ser claramente establecidos:

- Un operador fiable para la detección de puntos relevantes, de forma que sea tan insensible como sea posible frente al ruido y a cambios de intensidad y contraste.
- Un conjunto de características asociado a cada punto que tendrá que ser invariante tanto ante transformaciones de la intensidad como ante rotaciones, y un procedimiento para la extracción de las características elegidas utilizando sólo vecindades locales del punto relevante.
- Un algoritmo fiable para establecer correspondencias iniciales.
- Una formulación estadística para establecer los mejores encajes, y un procedimiento para estimar la transformación geométrica.

En la sección siguiente se describe brevemente el detector de puntos relevantes utilizado. La sección 3 describe las características locales empleadas. El modelo estocástico para la estimación de la transformación geométrica se define en la sección 4. Finalmente, la sección 5 mostrará la aplicación de esta metodología a varias secuencias de

AFs.

3.2 Detector de puntos relevantes

El concepto de relevancia aplicado a un punto en una imagen puede ser definido de muchas formas diferentes: como distinguibilidad de los puntos de su vecindad, como unicidad (que sea diferente del resto de puntos relevantes), significación (de acuerdo con la tarea), etc. Una breve discusión sobre el tema se encuentra en [HS92a]. Sin embargo, la mayoría de los autores están de acuerdo en que esquinas y centros de regiones con simetría circular están entre los candidatos más apropiados para ser puntos relevantes.

Según Zhang y colaboradores ([ZHRB96]), los detectores de esquinas pueden ser clasificados de modo genérico en dos categorías diferentes: aquellos que actúan sobre imágenes umbralizadas o sobre una imagen de bordes extraída de la imagen original, y aquellos que operan directamente sobre la imagen en niveles de gris. La primera de las categorías no es apropiada para nuestro caso, ya que la elección de umbral (que debido a la difusión tendría que ser local) o el proceso de detección de bordes serían tareas demasiado complicadas. Más aún, por su propia naturaleza estos detectores no están diseñados para tratar con centros de regiones radialmente simétricas, lo cual aparece con frecuencia en nuestras imágenes.

El segundo grupo incluye detectores que usan un criterio basado en alguna medida de la curvatura derivada de la geometría diferencial para establecer si un punto dado es una esquina. Un buen ejemplo es [WB92].

Otros detectores que también operan sobre imágenes en niveles de gris son aquellos basados en modelos. Uno de los más conocidos es debido a Förstner ([HS92a]). Förstner supone que un punto es una esquina si un número de segmentos rectos acaban en un mismo punto desconocido, la posición estimada de la esquina. Una definición análoga puede establecerse para centros de aspectos con simetría circular, para los que los segmentos de borde no deben intersectarse, sino al contrario, ser perpendiculares a las hipotéticas líneas que convergen en el punto.

Una descripción detallada del detector de Förstner está fuera del propósito de este

trabajo. La elección de este detector ha sido motivada por su habilidad para detectar aspectos circulares, lo que permite contar los impactos de láser que poseen los pacientes tratados previamente (ver figura 3.1). Otra razón importante es que el número de puntos detectados puede ser controlado por el valor de los parámetros utilizados. Sin embargo, pensamos que un estudio comparativo de varios detectores aplicados a este tipo particular de imágenes tendría indudablemente un alto interés práctico. A priori, no tenemos ninguna razón para creer que alguno de los detectores que aparecen en la literatura es el más apropiado para este tipo de imágenes.

Otro comentario interesante a nuestro parecer sobre el detector de Förstner es que utiliza un umbral derivado directamente de la magnitud del gradiente. Esto induce un comportamiento pobre en áreas de bajo contraste. Para minimizar este problema hemos experimentado con imágenes en las que hemos incrementado localmente el contraste mediante una ecualización del histograma ([GW87]) aplicada no a toda la imagen de una vez sino a diferentes partes separadamente.

La influencia de este procedimiento en el número y organización espacial de los puntos detectados debe ser medida por un índice de regularidad. Una medida ampliamente aceptada en el contexto de procesos puntuales es el índice de dispersión discutido por Diggle en 1983 ([Dig83]). Se calcula dividiendo la región de interés en varias áreas (cuadrantes, en la literatura) y contando el número de puntos presentes en cada uno de los cuadrantes. Sea N_i ($i = 1, \dots, M$) tal número. Un estadístico natural es el índice de dispersión,

$$I_{mv} = \frac{\sum_{i=1}^M (N_i - \bar{N})^2}{(M-1)\bar{N}} \quad (3.1)$$

Bajo la hipótesis de aleatoriedad espacial, es decir, puntos distribuidos uniforme e independientemente, la distribución que sigue el estadístico $(M-1)I_{mv}$ es una χ_{M-1}^2 con $M-1$ grados de libertad.

3.3 Características locales

Una vez que un conjunto apropiado de puntos relevantes se ha detectado, el paso siguiente es la elección de parámetros locales que constituyan un vector de características asociado a cada punto que se utilizará para establecer la correspondencia entre ellos. Nuestro enfoque difiere de algunos de los enfoques anteriores [HS92b] en que en éstos los puntos se encajan mediante la correlación cruzada de ventanas centradas en cada punto o desplazadas unos pocos *pixels*; pero esto no es conveniente para nuestra aplicación debido a los cambios de contraste y la posibilidad de rotación. Al igual que los otros, nuestro enfoque asume que los puntos correspondientes en diferentes imágenes tienen vecindades similares; no obstante, en lugar de usar la correlación, intenta dar una descripción numérica robusta de la vecindad de cada punto, invariante a rotaciones, traslaciones y cambios de contraste. Dicha descripción, junto con un estimador del movimiento esperado, se necesita para comparar la vecindad de puntos relevantes en imágenes diferentes. Para construir el vector de características invariantes hemos elegido la función de autocorrelación de intensidad evaluada a diferentes distancias del punto de interés en una ventana circular centrada en éste.

Una imagen digital en niveles de gris puede ser considerada como una realización de un proceso aleatorio definido en un espacio bidimensional muestreado en la cuadrícula de puntos [Bil89]. Si asumimos que el proceso aleatorio es estacionario (distribución invariante con respecto a traslaciones) tiene sentido definir la correlación cruzada entre variables cuyas posiciones están unidas por cualquier vector \mathbf{h} .

Sea \mathbf{h} la función de correlación cruzada. Además, para procesos estacionarios e isotrópicos (distribuciones invariantes con respecto a rotaciones alrededor del origen), la correlación cruzada depende solamente del módulo t de \mathbf{h} . Sea $C(t)$ una función que denota la correlación cruzada entre valores de gris observados en puntos separados una distancia t . $C(t)$ puede ser estimado en cualquier punto de la imagen para cualquier valor de t utilizando una ventana circular W centrada en el punto de la siguiente forma: la imagen original I es en primer lugar normalizada en W sustrayendo el valor de la media en W y dividiendo por la desviación estándar en W . Esto generará la imagen I_s . $C(\mathbf{h})$ se puede evaluar directamente si las componentes de \mathbf{h} son múltiplos de la distancia interpixel como $C(\mathbf{h}) = \frac{\sum_{x: x, x+h \in W} I_s(x+h)I_s(x)}{\#\{x: x, x+h \in W\}}$ donde $\#$ es la cardinalidad. Para

cualquier otro valor positivo de t , $C(t)$ se puede estimar a partir de $C(\mathbf{h})$ utilizando el *kernel* de Epanechnikov [Sil86]:

$$\hat{C}(t) = \sum_{\mathbf{h}} L(\mathbf{h}) K(t - \|\mathbf{h}\|), \quad (3.2)$$

con

$$K(t) = \begin{cases} \frac{3}{4\sqrt{5}}(1 - \frac{1}{5}t^2) & \text{if } -\sqrt{5} \leq t \leq \sqrt{5} \\ 0 & \text{en otro caso} \end{cases}$$

Ya que $\hat{C}(t)$ se calcula como el promedio sobre una circunferencia centrada en el punto de interés, $\hat{C}(t)$ es una característica invariante ante el cambio de orientación de la imagen bajo rotación de este punto. Así, podemos asociar a cada punto relevante un vector de características mediante la evaluación de $\hat{C}(t)$ en 5 valores equidistantes de t con una ventana circular de radio 10. Esta ventana debe estar completamente incluida en I , en otro caso el punto se descartará. Sea $\{s_i\}$ y $\{t_j\}$ los conjuntos de puntos relevantes detectados en la primera y segunda imagen respectivamente y sean m_i y l_i los vectores de características asociados a cada punto de los conjuntos definidos. Dos puntos s_i y t_j se asume que son correspondientes si sus vectores de características m_i y l_j son similares de acuerdo con alguna medida de similitud $D(m_i, l_j)$. Nosotros hemos utilizado la distancia euclídea entre m_i y l_j como medida de similitud D .

3.4 Modelo estocástico para el encaje de imágenes

La angiografía fluoresceínica es un conjunto dinámico de imágenes que cambian a lo largo del tiempo. El nivel de gris de una angiografía en cualquier punto puede cambiar dramáticamente en unos pocos segundos. Un punto relevante puede desaparecer o puede cambiar su localización, o algunos nuevos puntos relevantes pueden aparecer en la segunda imagen. En términos generales, dos efectos tienen que ser considerados:

- Efecto de nacimiento y muerte: algunos puntos son detectados solamente en la segunda imagen. Otros puntos detectados en la primera imagen desaparecen. Este es quizá el principal problema a la hora de establecer la correspondencia entre puntos relevantes.
- Efecto de desplazamiento: como cualquier detector de puntos relevantes que utilice

solamente información local, cambios de intensidad local pueden introducir pequeños desplazamientos de los puntos detectados sin un efecto de nacimiento o muerte.

Estos dos efectos o limitaciones no deseados no pueden ser controlados pero pueden tenerse en cuenta. Se pueden modelar estocásticamente, y se pueden considerar estimadores estadísticos para una estimación robusta de la transformación geométrica.

Siguiendo el procedimiento aquí considerado, han sido detectados dos conjuntos de puntos: $\{s_i : i = 1, \dots, n_1\}$ en la primera imagen y $\{t_j : j = 1, \dots, n_2\}$ en la segunda imagen.

Denotemos por $\theta = [\theta(i, j)]_{i=1, \dots, n_1, j=1, \dots, n_2}$ a la matriz de correspondencias de puntos, es decir, $\theta(i, j) = 1$ si el i -ésimo punto de la primera imagen corresponde a la j -ésimo de la segunda, y $\theta(i, j) = 0$ en cualquier otro caso. El problema del encaje de imágenes puede entonces ser modelado por la ecuación

$$\theta(i, j)t_j = \theta(i, j)(T(s_i) + \epsilon_j) \quad (3.3)$$

donde s_i , con $i = 1, \dots, n_1$ y t_j , con $j = 1, \dots, n_2$ son los elementos del primer y segundo conjunto de puntos relevantes; ϵ_j son vectores normales 2×1 , independientes e idénticamente distribuidos, con varianza Σ ($\epsilon_j \sim N_2(0, \Sigma)$), que dan cuenta del desplazamiento errático local debido a cambios de intensidad locales; T denota la transformación geométrica global.

Si asumimos que T es lineal, entonces la ecuación (3.3) puede ser reescrita como

$$\theta(i, j)t_j = \theta(i, j)(As_i + B + \epsilon_j), \quad i = 1, \dots, n_1 \quad j = 1, \dots, n_2, \quad (3.4)$$

donde A es una matriz 2×2 y B es una matriz 2×1 , representando la transformación geométrica global de la primera imagen en la segunda. A partir de ahora, este será el modelo asumido, que es intrínsecamente un modelo no lineal debido al producto por la matriz θ [DS81]. No obstante, nótese que dado θ , se convertiría en un modelo lineal (pero no necesariamente una rotación, ya que A no está restringida a ser ortogonal). Los vectores aleatorios ϵ_j se asume que son normalmente distribuidos con una matriz de covarianzas dada por Σ ($\epsilon_j \sim N_2(0, \Sigma)$). La conveniencia de este modelo y la asunción de que ϵ_j está normalmente distribuido se verifica en la sección 5. La verosimilitud la

podemos expresar como

$$L(t_1, \dots, t_{n_2} \mid A, B, \Sigma, \theta, s_1, \dots, s_{n_1}) = \quad (3.5)$$

$$|2\pi\Sigma|^{-\frac{n_2}{2}} \exp\left\{-\frac{1}{2} \sum_{j=1}^{n_2} \sum_{i=1}^{n_1} \theta(i, j) (t_j - B - As_i)^t \Sigma^{-1} (t_j - B - As_i)\right\},$$

donde $\theta(i, j) = \sum_i \theta(i, j)$. Estamos interesados en la estimación de A , B , Σ y θ . El número de parámetros es mayor que la información de muestra. Si las correspondencias fueran conocidas y la matriz θ también, los estimadores comunes para A , B y Σ serían los estimadores de mínimo error cuadrático, que coinciden con los estimadores de máxima verosimilitud. El principal problema, sin embargo, es la estimación de la matriz de correspondencias. Algunos comentarios previos deben tenerse en cuenta:

1. En primer lugar, la vecindad de los puntos correspondientes tiene que ser similar. En la sección 3.3 se describen los detalles referentes a la similitud de la vecindad.
2. Cada punto relevante puede corresponder con un punto o ninguno, es decir, $\sum_i \theta(i, j) \leq 1$ y $\sum_j \theta(i, j) \leq 1$.

Estos dos comentarios nos han sugerido dos diferentes procedimientos de estimación: un procedimiento iterativo basado en la predicción y un procedimiento Bayesiano usando una distribución *a priori* para θ . El estimador para θ en el segundo de los enfoques es el máximo *a posteriori*, es decir, el máximo de la distribución *a posteriori*.

Estos dos procedimientos requieren un paso previo en el que se determina un estimador inicial de la matriz de correspondencias. Estimaciones mejoradas para θ y el resto de parámetros serán entonces obtenidas de modo iterativo.

3.4.1 Correspondencias iniciales

La matriz de correspondencias iniciales $\hat{\theta}_0$ ha sido determinada como sigue:

Resumen del algoritmo

SEAN S Y T los conjuntos de puntos relevantes en la primera y segunda imagen respectivamente

SEA $M \subset S \times T$ el conjunto de correspondencias. Inicialmente, $M \leftarrow \emptyset$

SEA $Nivel \leftarrow 0$

HACER

Entre todos los pares de puntos $P = (s_i, t_j) \in S \times T$ tal que $P \notin M$

Encontrar el par $P_{min} = (s_u, t_v)$ tal que

($D(m_u, l_v)$ es el mínimo entre las correspondencias no prohibidas en este nivel)

SI (P_{min} existe) ENTONCES

SI P_{min} es CONSISTENTE con M

ENTONCES

$M \leftarrow M \cup \{P_{min}\}$

$Nivel \leftarrow Nivel + 1$

SINO

Marcar P_{min} como encaje prohibido en este nivel

SINO

$Nivel \leftarrow Nivel - 1$

MIENTRAS ($Nivel > 0$)

La incorporación de un nuevo par $P = (s_k, t_l)$ al conjunto $M = \{(s'_i, t'_i) : i = 1 \dots Nivel\}$ se considera

CONSISTENTE si $|\|s'_i - s_k\| - \|t'_i - t_l\|| \leq \delta \forall i = 1, \dots, Nivel$, donde δ es una constante positiva que se ha fijado en cuatro *pixels*. Esta restricción en las distancias ayuda en la imposición de similitud entre los dos árboles formados por los puntos actualmente encajados en las dos imágenes.

En resumen, este procedimiento para establecer correspondencias iniciales es una búsqueda heurística en el árbol de correspondencias, donde la función de evaluación que guía la búsqueda a través de la rama más prometedora se basa en la distancia entre puntos definida en la sección 3.3, junto con la restricción adicional de la consistencia geométrica.

3.4.2 Un procedimiento iterativo de estimación

Veamos el procedimiento iterativo propuesto para la estimación de la matriz de correspondencias. La matriz de correspondencias iniciales $\hat{\theta}_0$ se determina utilizando el procedimiento descrito en la sección 3.4.1, dando como resultado un conjunto inicial $\{(t_j, s_{i(j)}) : \hat{\theta}_0(i(j), j) = 1, \sum_i \hat{\theta}_0(i, j) = 1\}$ de puntos correspondientes. Los estimadores de mínimo error cuadrático \hat{A}_0 , \hat{B}_0 and $\hat{\Sigma}_0$ ([DS81]) se calculan a partir de esas correspondencias. Una nueva matriz de correspondencias $\hat{\theta}_1$ es entonces construida, admitiendo correspondencias entre cada punto s_i y su punto más cercano t_j si t_j está contenido en la región de confianza asociada con el transformado de s_i bajo (\hat{A}, \hat{B}) y si las características asociadas con s_i y t_j son suficientemente similares. Si t_j no satisface estas restricciones, entonces s_i no tiene punto correspondiente. Estas condiciones pueden expresarse matemáticamente como: $\hat{\theta}_i(i, j) = 1$ si y solo si

$$(t_j - (\hat{A}_0 s_i + \hat{B}_0))^t \hat{\Sigma}_0^{-1} (t_j - (\hat{A}_0 s_i + \hat{B}_0)) = \min_k \{(t_k - (\hat{A}_0 s_i + \hat{B}_0))^t \hat{\Sigma}_0^{-1} (t_k - (\hat{A}_0 s_i + \hat{B}_0))\} \leq \epsilon_1 \quad (3.6)$$

y

$$D(m_i, l_j) \leq \epsilon_2, \quad (3.7)$$

La región de confianza ϵ_1 se define como sigue:

$$\epsilon_1 = F(1 - \alpha, 2, n - r(X) - 1) \frac{2}{n - r(X) - 1} (n - r(X)) (1 + x_j^t (X^t X)^{-1} x_j) \quad (3.8)$$

con

$$n = \sum_{ij} \hat{\theta}_0(i, j),$$

$F(1 - \alpha, 2, n - r(X) - 1)$ es el cuantil $1 - \alpha$ de una distribución F con 2 y $n - r(X) - 1$ grados de libertad

$$\{s_{i_1}, \dots, s_{i_n}\} = \{s_i : \sum_j \hat{\theta}_0(i, j) = 1\}$$

X una matriz cuya j -ésima fila es $[s_j^t, 1]$

$r(X)$ es el rango de X

$$x_j^t = [s_j^t, 1].$$

ϵ_1 resulta ser altamente dependiente del número y localización de los puntos, pero valores típicos para él están en el rango de 30-40. Hemos utilizado para α los valores

0.05 o 0.01. ϵ_2 se ha fijado a dos veces la mayor componente de la matriz de covarianzas de (m_i, l_j) .

Nuevas estimaciones para A , B y Σ son calculadas utilizando la matriz de correspondencias actualizada $\hat{\theta}_1$. Este procedimiento se repite hasta que $\hat{\theta}_i = \hat{\theta}_{i+1}$.

$D(m_i, l_j)$ es una medida de similitud entre los vectores de atributos m_i y l_j . La definición de ϵ_2 depende de D . Las características locales tienen que ser similares para evitar encajes absurdos producidos por reflexiones especulares u otro tipo de razones que pueden causar que el algoritmo caiga en un mínimo local.

Veamos la idea fundamental del método anterior: para cada punto s_i , elegimos como punto correspondiente el punto más cercano t_j contenido en la región de confianza asociada con el transformado de s_i . Obviamente, la región de confianza se determina basándose en el modelo (3.4). Si no existe ningún t_j en la región entonces s_i no tiene punto correspondiente.

3.4.3 Un campo aleatorio de Markov en la matriz de correspondencias

En este segundo enfoque, la matriz de correspondencias θ se considera generada por un campo aleatorio de Markov [Bes74, GG84] con respecto al sistema de vecindad $\mathcal{V} = \{V_{ij}\}$, donde V_{ij} es la vecindad de $\theta(i, j)$ definida como:

$$V_{ij} = \{\theta(i, k)\}_{k \neq j} \cup \{\theta(l, j)\}_{l \neq i}$$

es decir, la vecindad de un elemento θ_{ij} la componen aquellos elementos situados bien en la misma fila o bien en la misma columna. La distribución *a priori* propuesta se expresa de la siguiente forma:

$$\mathcal{P}(\theta(i, j) / \{\theta(k, l) \in V_{ij}\}) \propto \exp\{-\alpha n_{ij} + \theta(i, j) \ln(d_{ij}) + (1 - \theta(i, j)) \ln(1 - d_{ij}) - \beta(1 - \theta(i, j))\}, \quad (3.9)$$

donde: (i) $0 \leq d_{ij} \leq 1$ es una medida de similitud entre las características locales de s_i y t_j

(ii) β es una constante que mide nuestro conocimiento *a priori* del número de corres-

pondencias

- (iii) $n_{ij} = 1$ si existe un vecino de $\theta(i, j)$ teniendo valor 1 y $n_{ij} = 0$ en otro caso
- (iv) α es una constante introducida para evitar que exista más de un vecino de $\theta(i, j)$ con valor uno; intuitivamente : para que cada punto tenga a lo más un correspondiente, pues no admitimos desdoblamiento.

Puede forzarse el cumplimiento de esta condición haciendo que $\alpha \rightarrow \infty$ pero esto contradice la condición de positividad en el teorema de Hammersley-Clifford [Cre93, Bes74] pues es un requisito para la demostración de este teorema (Ver apendice A). Este problema puede resolverse incrementando α iterativamente.

Teniendo en cuenta las ecuaciones (6.4) y (3.9), la distribución de probabilidad a posteriori en la matriz de correspondencias se puede deducir:

$$\begin{aligned} \mathcal{P}(\theta(i, j)/\{s_i\}, \{t_j\}, V_{ij}, A, B, \Sigma) \propto & \quad (3.10) \\ \exp\{-\theta(i, j)(t_j - (As_i + B))^t \Sigma^{-1} (t_j - (As_i + B)) & \\ -\alpha n_{ij} + \theta(i, j) \ln(d_{ij}) + (1 - \theta(i, j)) \ln(1 - d_{ij}) - \beta(1 - \theta(i, j))\}, & \end{aligned}$$

Como se ha establecido anteriormente, la estimación de la matriz de correspondencias implica la maximización de la distribución *a posteriori*. Con este objetivo, utilizamos el algoritmo de relajación estocástica propuesto por Geman y Geman basado en el algoritmo conocido como *Simulated annealing* y el *Gibbs Sampler* (SA) [GG84]. Otras opciones podrían ser el algoritmo de relajación determinista conocido como *iterated conditional modes* (ICM) [Bes86]. Aunque este segundo algoritmo tiene una convergencia muy rápida, puede fácilmente caer en un máximo local. La naturaleza estocástica del primer procedimiento permite que el método escape de cualquier máximo local. Hemos escrito una versión de este algoritmo que incorpora la estimación de A , B , y Σ . Proponemos el siguiente procedimiento iterativo:

Resumen del algoritmo

Paso 0, Inicialización : Se determina una matriz de correspondencias iniciales $\hat{\theta}_0$. α y el parámetro de temperatura τ son inicializados y las distancias d_{ij} , $i = 1, \dots, n_1, j = 1, \dots, n_2$ son calculadas.

Paso 1, Estimación de A , B y Σ : A , B , y Σ se estiman mediante el máximo de la

distribución de pseudoverosimilitud de la $\hat{\theta}$ actual para obtener los nuevos \hat{A} , \hat{B} y $\hat{\Sigma}$.

Paso 2, Estimación de θ : Se realiza un ciclo del SA basado en las $\hat{\theta}$, \hat{A} , \hat{B} y $\hat{\Sigma}$ actuales para obtener una nueva $\hat{\theta}$. Una iteración del SA es un barrido completo sobre todas las componentes de θ de forma que el nuevo valor de $\hat{\theta}(i, j)$ es aleatoriamente elegido con una probabilidad proporcional a $\mathcal{P}^{1/\tau}(\theta(i, j)/\{s_i\}, \{t_j\}, V_{ij}, A, B, \Sigma)$.

Paso 3, Actualización de los parámetros de control: α se incrementa y la temperatura τ se decrementa.

Paso 4: Se vuelve al paso 1 durante un número fijo de ciclos o hasta una convergencia aproximada de $\hat{\theta}$.

Para asegurar la convergencia del método τ y α deben ser decrementados e incrementados (respectivamente) muy lentamente [GG87], pero esto provoca una convergencia muy lenta. En la práctica, hemos obtenido buenos resultados utilizando $\frac{1}{\tau} = \alpha = 1.01^k$, donde k es el número de barridos completos realizados. El valor de β utilizado ha sido de 0.3.

3.4.4 Transformación geométrica

Utilizando alguno de los procedimientos anteriormente descritos hemos obtenido unas matrices A y B . La transformación geométrica se ha aplicado mediante un método de interpolación. Si \hat{T} denota la transformación estimada y f es la imagen original, la imagen f transformada $f_{\hat{T}}$ se define en el pixel h como:

$$f_{\hat{T}}(h) = \sum_{\hat{T}(s) \in N(h)} K(\|s - h\|) f(s), \quad (3.11)$$

donde K denota el kernel de Epanechnikov y $N(h)$ es la función soporte de K (es decir, aquellos puntos s tal que $K(\|s - h\|) > 0$).

3.5 Resultados experimentales

La metodología propuesta puede aplicarse en principio a muchos tipos diferentes de imágenes. Sin embargo, la motivación inicial y el contexto en el que ha sido aplicada es el de ayuda a los oftalmólogos en el diagnóstico de ciertas patologías oculares donde la técnica exploratoria de la angiografía fluoresceínica juega un papel crucial. Las imágenes que han sido elegidas son de pacientes con tres tipos diferentes de enfermedades: retinopatía diabética, oclusión de venas y membranas neovasculares coroidales.

Los negativos de angiografías fluoresceínicas tomados con una cámara de fondo de ojo *Canon CF-60U* y *Nikon NFC 50* se transformaron en señal de vídeo con un conversor SONY y fueron digitalizadas con una tarjeta *Matrox MVP-AT* de 512×380 píxeles. Las librerías de dominio público VISTA ¹ se han utilizado para la implementación de todos los algoritmos.

Una valoración razonable del método de alineamiento tiene que identificar los factores que se ha considerado que tienen una influencia sobre el resultado final. Según el sentido común y la verificación experimental en un conjunto de 100 pares de imágenes, se han identificado como: (i) el número de puntos relevantes detectados en cada imagen; (ii) la distribución espacial de estos puntos y (iii) el número de correspondencias iniciales.

Se han diseñado algunos experimentos teniendo en cuenta estos factores. Los parámetros del detector de puntos afectan al número de puntos encontrados: dichos parámetros han sido elegidos *ad hoc* de forma que se obtuviera un número similar de puntos relevantes. Un amplio conjunto de puntos relevantes debería producir una mejor estimación de la transformación geométrica pero el proceso sería más lento. Además, una disposición más regular de estos puntos debería mejorar tanto la velocidad a la que las correspondencias se encuentran como la precisión en la estimación de la matriz de transformación global. La precisión del alineamiento se ha evaluado como el error cuadrático medio entre las posiciones de los puntos relevantes reales y transformados para aquellos puntos no utilizados en la estimación de la transformación. Estos puntos adicionales (marcados

¹Desarrollado principalmente por Art Pope en el departamento de Ciencias de la Computación, Universidad de British Columbia

manualmente) han sido escogidos normalmente lejos de la zona foveal para proporcionar una estimación de un límite alto de error; casi todas las correspondencias no encontradas por el programa se han utilizado, con un rango de entre 4 y 7 correspondencias.

El detector de Förstner se ha aplicado a 122 imágenes. El número medio de puntos detectados ha sido 31.3 y los cuantiles más bajo y más alto han sido 14.7 y 44, respectivamente. Para cada imagen, el índice I_{mv} y el p -valor correspondiente se ha calculado. El valor medio de I_{mv} ha sido 4.1 y sus cuantiles más bajo y más alto, 2.2 y 5.3 respectivamente. Los p -valores observados (todos ellos más pequeños que 0.001) permiten rechazar la aleatoriedad espacial completamente en favor de un patrón agregado.

El programa se ejecutó en un ordenador *HP-735* con un tiempo de CPU entre 113.9 y 151.9 segundos para cada par de imágenes, con un promedio de 134.9 s.

La figura 3.1 muestra 2 imágenes de una angiografía fluoresceínica tomadas con unos pocos segundos de diferencia a un paciente con retinopatía diabética después de un tratamiento de laser en un edema macular (de ahí los puntos de hiperfluorescencia). Estos puntos relevantes y las correspondencias establecidas están indicados en las imágenes. Los resultados se resumen en la tabla 3.1. La primera columna indica si se aplica ecualización antes de la detección de los puntos relevantes. Las columnas etiquetadas con $n_1, I_{mv}(1)$, $n_2, I_{mv}(2)$ representan el número de puntos detectados y sus índices de dispersión en la primera y segunda imagen respectivamente. En todas las tablas cada fila refleja un número diferente de correspondencias iniciales, obtenidas dejando que el detector encontrara inicialmente un número diferente de puntos. Esto se ha conseguido eligiendo manualmente ciertos parámetros usados por el detector de Förstner.

El primer experimento es un caso donde la diferencia entre utilizar una imagen ecualizada (con puntos distribuidos más regularmente) o el detector de Förstner sin ecualización no es muy relevante. Si el número de puntos detectados es suficiente, incluso cuando los puntos no están bien distribuidos (véase el índice I_{mv}), se consigue un buen encaje y los valores de MSE no son más grandes de dos o tres *pixels*. Es importante destacar que intentar empezar con un número demasiado pequeño de correspondencias iniciales puede llevar al algoritmo a caer en un mínimo local. Si este es el caso, no es posible encontrar correspondencias y, por lo tanto, no se puede estimar

la transformación geométrica.

El segundo caso (ver figura 3.2), que pertenece a otro paciente con retinopatía diabética, es similar al primero pero en imágenes muy mal contrastadas. Esta situación ocurre normalmente en paciente ancianos con cataratas y obliga a un paso previo de ecualización, lo que en este caso produce los resultados que se muestran. El caso no ecualizado no aparece en la tabla, ya que no se encontraron en el primer paso un número suficiente de puntos para dejar al algoritmo iterativo refinar las correspondencias. En este caso, los errores con las correspondencias iniciales son ligeramente más grandes, sobre 7 puntos.

El tercer caso (figura 3.3) es de un paciente con oclusión de la vena superotemporal. La difusión (áreas hiperfluoresceínicas) aparece en la segunda imagen. Este caso es uno de los más difíciles. Debido a una correspondencia inicial errónea para un par de puntos, el algoritmo cae en un mínimo local del cual el algoritmo iterativo de refinamiento no es capaz de salir, dando como resultado un encaje incorrecto. El alineamiento es aparentemente bueno, porque el área de alineamiento aceptable coincide con las venas mientras que el área de alineamiento erróneo es en un área uniforme de gris. Sin embargo, el error cuadrático medio es 15.4 puntos. Obsérvese la diferencia con la primera fila, para la que el error cuadrático medio es 4.2 puntos.

La cuarta y última secuencia (figura 3.4) es un caso de neovascularización coroidal (aparición de nuevos vasos en la coroides) adyacente al disco óptico en un paciente con miopía. Aparece una zona de hiperfluorescencia que rodea la membrana neovascular.

Con respecto a los modelos propuestos, digamos que el modelo representado por la ecuación (3.4) asume que los residuos ϵ_j están distribuidos normalmente. Hemos verificado la asunción inspeccionando la distribución del residuo $t_j - (\hat{A}s_i + \hat{B})$ de todos los puntos correspondientes (s_i, t_j) para todos los experimentos. Hemos encontrado que cada componente se distribuye normalmente con p-valores mayores que 0.5 en todos los casos. Las figuras 5(izquierda y derecha) muestran las gráficas de cuantil-cuantil de la primera y segunda componente respectivamente de los residuos obtenidos en el primer experimento (figura 1), lo cual indica claramente que la asunción de normalidad es válida.

3.6 Discusión

A continuación haremos algunos comentarios sobre la comparación de este trabajo con los enfoques más parecidos en la solución del problema del alineamiento [JB⁺90, RMP⁺96]. A diferencia de estos dos trabajos, en nuestra aproximación se ha propuesto un modelo para el ruido. Basado en este modelo estocástico dos diferentes estimadores de la transformación geométrica han sido igualmente propuestos. Muchos otros métodos de estimación pueden utilizarse y compararse con nuestro método estadístico. Jagoe y colaboradores ([RJ⁺93]) estiman la matriz de correspondencias y la transformación geométrica separadamente. Sin embargo, una vez que los puntos marcados han sido encajados se realiza una eliminación para borrar los que dan mayores errores residuales. Por otra parte, hemos estimado las correspondencias y la transformación alternativamente de una forma iterativa, como lo hace Rangarajan, pero utilizando diferentes estimadores.

Algunas críticas que podrían indicarse acerca del método presentado tienen que ver con la velocidad y la precisión. El número de puntos relevantes que se pasa al algoritmo que calcula las correspondencias iniciales no debe ser ni muy pequeño (para permitir establecer suficiente número de correspondencias) ni muy grande (ya que el árbol de búsqueda tiene una complejidad exponencial, incluso si es podado mediante el conocimiento heurístico). Esto requiere que se haga una elección cuidadosa de los puntos, y una forma de clasificarlos por relevancia si alguno de ellos debe ser descartado.

La precisión observada es significativamente peor cuando se está lejos de las áreas de alta densidad de puntos relevantes. Puesto que estas son áreas periféricas en nuestras imágenes, esto no altera significativamente la aplicabilidad médica para este uso particular; pero otras aplicaciones del método puede requerir mejoras. Una posible solución consiste en ordenar los puntos relevantes no solamente por su unicidad sino también por su localización, intentando detectar más de los necesarios y manteniendo un número suficiente de puntos bien distribuidos. Algunas veces, esto puede perder algunas correspondencias interesantes. Otra posible solución sería el uso de una transformación no lineal o no local, en lugar de las matrices globales que hemos asumido.

El algoritmo utilizado para establecer la matriz de correspondencias iniciales de-

pende fuertemente del número de puntos relevantes detectados. Los resultados más satisfactorios se han obtenido con 20 a 50 puntos por imagen. Sin embargo, la distribución espacial parece no ser tan importante. Resultados similares se han obtenido con o sin una ecualización previa.

Una deficiencia observada es que el conjunto de puntos relevantes son consistentes solamente si las distancias euclídeas entre ellos son similares. Esto no es un grave problema si se asume un movimiento cuasi-rígido, pero sólo hasta cierto punto puede asumirse para puntos en la periferia de la imagen, debido a la deformación producida por la rotación del ojo (esto se tratará en el capítulo 5). Esto indica que debería usarse un procedimiento menos restrictivo, por ejemplo el encaje del mínimo *spanning tree* de los dos conjuntos de puntos [Sha95].

En cuanto la elección de las características relacionadas con cada punto relevante, la función de autocorrelación se ha utilizado como una medida estandar de la dependencia local en el contexto de los procesos aleatorios. Otros descriptores locales alternativos merecen un estudio futuro.

Sobre el número de componentes del vector de características, p , el cual fija el tamaño de la ventana que tiene que ser de al menos $2p$, nuestra elección de $p = 5$ viene sugerida por el algoritmo de Förstner utilizado en la detección de puntos relevantes donde se han considerado las intensidades con una ventana de 9×9 unidades. Haciendo esto, las dependencias entre puntos localizados más lejos de 9 unidades no se tienen en cuenta. Los buenos resultados experimentales obtenidos apoyan esta elección aunque nuevos experimentos podrían realizarse.

Algunas otras medidas más sofisticadas y quizás, más difíciles de calcular, pueden proponerse como alternativas adecuadas para la distancia Euclídea entre los vectores de características (p.ej., la distancia de Mahalanobis, ...).

Finalmente, hemos estimado la transformación geométrica asumiendo un modelo lineal. Como se mostró en la sección 7.4, este modelo es suficientemente bueno para una angiografía fluoresceínica. Recordemos que nuestra área de interés primaria es la zona avascular de la fovea, que al situarse en el centro de la imagen hace que la transformación real desconocida se puede aproximar mediante una transformación lineal. Claramente,

esta transformación lineal corresponde a la aproximación de Taylor de primer orden de la transformación geométrica real.

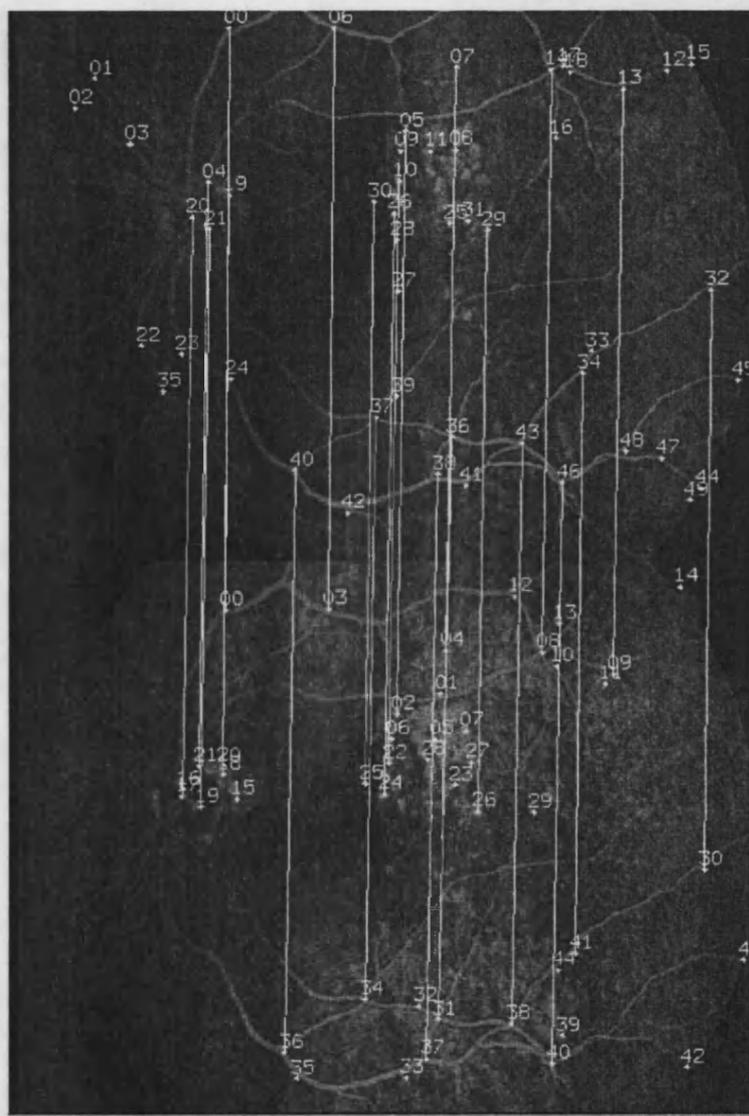


Figura 3.1: Experimento 1: caso de retinopatía diabética; $ic = 7$ con ecualización

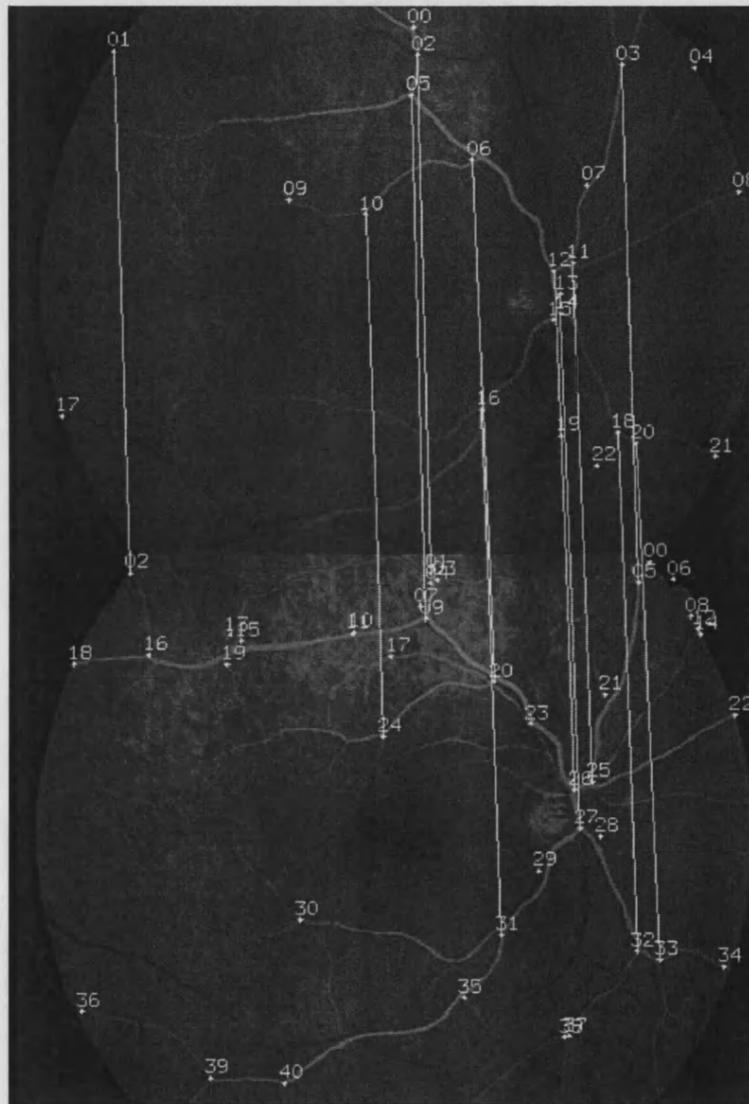


Figura 3.2: Experimento 2: caso de retinopatía diabética; $ic = 3$ con ecualización

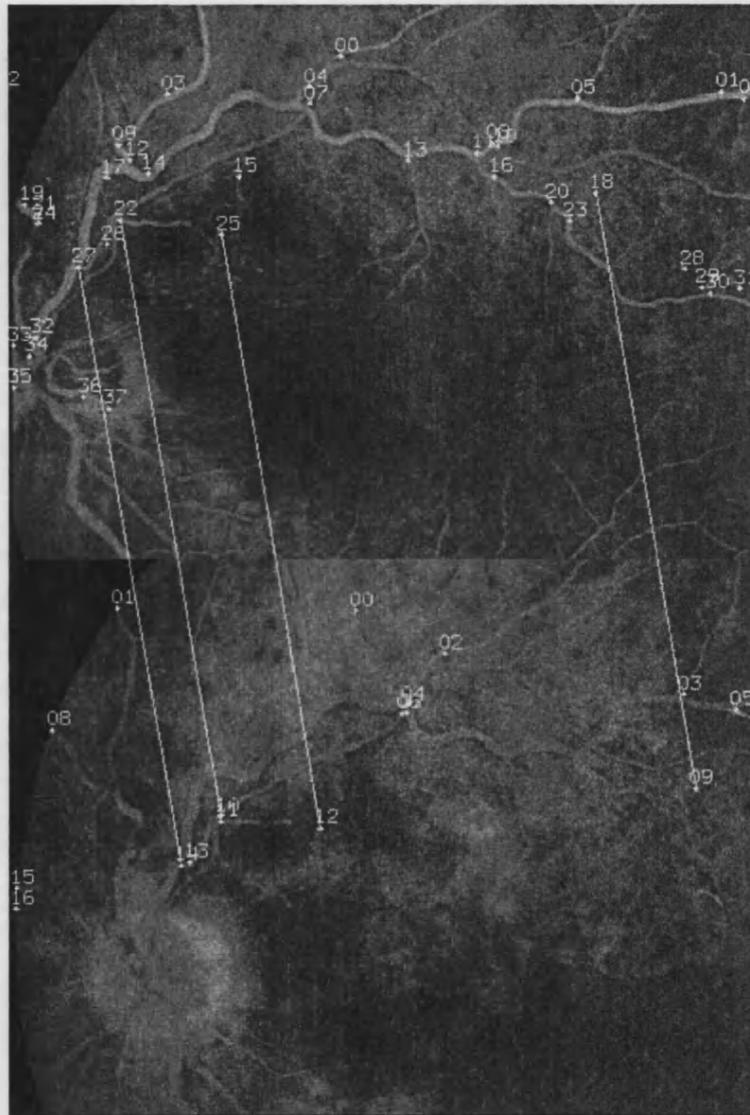


Figura 3.3: Experimento 3: caso de oclusión venosa; $ic = 4$ sin ecualización

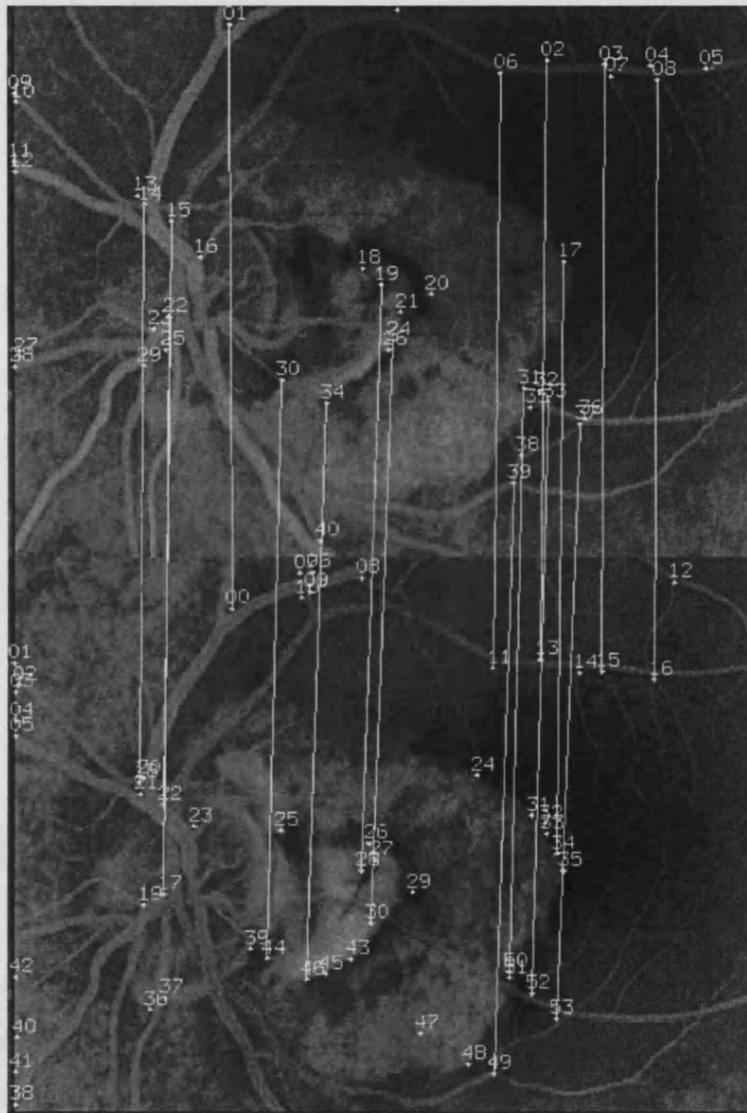


Figura 3.4: Experimento 4: caso de neovascularización; $ic = 7$ sin ecualización

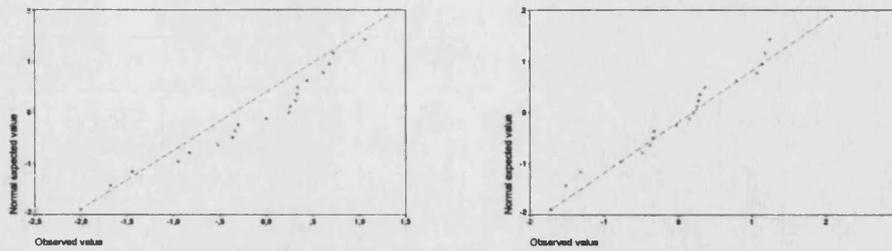


Figura 3.5: Gráfica cuantil-cuantil de las coordenadas primera y segunda de los residuos obtenidos en el experimento 1.

						Método 1			Método 2		
Ec.	n_1	$I_{mv}(1)$	n_2	$I_{mv}(2)$	ic	fc_1	MSE1	MAX1	fc_2	MSE2	MAX2
SI	17	0.91	15	1.36	3	4	3.01	4.48	4	1.28	2.53
					3	4	3.01	4.19	4	1.28	2.19
					3	4	3.01	4.10	4	1.41	2.77
SI	39	1.80	33	2.32	3	-	N.A.	N.A.	-	N.A.	N.A.
					5	10	2.33	4.04	17	1.24	1.86
					7	10	2.40	3.25	17	1.24	2.01
SI	50	2.37	45	3.86	3	-	N.A.	N.A.	-	N.A.	N.A.
					5	10	3.33	6.16	19	1.55	2.39
					7	10	3.33	4.77	22	1.24	1.56
NO	33	4.22	27	10.14	3	13	2.86	3.68	19	1.27	1.89
					5	13	2.86	5.21	15	0.89	1.64
					7	13	3.21	4.95	15	0.89	1.67

Tabla 3.1: Resultados del experimento 1

						Método 1			Método 2		
Ec.	n_1	$I_{mv}(1)$	n_2	$I_{mv}(2)$	ic	fc_1	MSE1	MAX1	fc_2	MSE2	MAX2
SI	23	1.58	21	1.02	3	6	5.04	6.59	4	1.45	2.51
					5	6	4.73	5.48	3	1.41	2.60
					5	6	4.73	4.84	3	1.41	2.57
SI	23	1.58	26	1.35	3	8	5.37	5.60	–	N.A.	N.A.
					5	8	5.26	6.21	7	0.86	1.37
					7	8	5.26	5.97	9	1.33	2.17
SI	23	1.58	41	2.51	3	7	5.53	7.77	13	1.90	3.41
					5	7	5.41	6.49	7	0.96	1.62
					7	7	5.10	6.28	10	1.39	1.79

Tabla 3.2: Resultados del experimento 2

						Método 1			Método 2		
Ec.	n_1	$I_{mv}(1)$	n_2	$I_{mv}(2)$	ic	fc_1	MSE1	MAX1	fc_2	MSE2	MAX2
NO	41	6.33	15	1.78	3	4	4.23	8.19	4	0.84	1.28
					4	4	15.40	18.43	5	0.99	1.81
					4	4	15.40	20.73	4	1.27	1.92

Tabla 3.3: Resultados del experimento 3

						Método 1			Método 2		
Ec.	n_1	$I_{mv}(1)$	n_2	$I_{mv}(2)$	ic	fc_1	MSE1	MAX1	fc_2	MSE2	MAX2
NO	41	2.51	54	2.52	3	5	7.86	10.47	13	1.67	2.43
					5	5	16.60	21.84	19	2.81	5.35
					7	8	12.40	14.11	19	2.33	4.48

Tabla 3.4: Resultados del experimento 4

Capítulo 4

Alineamiento de angiografías mediante el algoritmo EM Generalizado

4.1 Introducción

El problema del alineamiento de imágenes ha sido ampliamente estudiado en la literatura de visión por computador [Ume88, SP95, LTC93, AK96]. Entre las técnicas utilizadas en el análisis de angiografías fluoresceínicas destacan el alineamiento manual [PRT⁺91], la correlación cruzada [SR92] y alineamientos basados en encaje de puntos [DAS⁺97]. En este capítulo se describe un nuevo método de encaje de imágenes que, a diferencia de los métodos propuestos en el capítulo anterior, utiliza la red de vasos sanguíneos como estructuras para realizar el encaje.

El método que proponemos utiliza grafos relacionales que representan la estructura de vasos sanguíneos en la retina. Los nodos representan cruces de vasos mientras que las aristas indican que hay un vaso sanguíneo de conexión. La estructura de aristas de la red de vasos se utiliza para restringir el encaje de los angiogramas. Estamos interesados, no solamente en estimar las correspondencias entre los cruces de la red, sino también en la estimación de los parámetros de la transformación de coordenadas entre diferentes imágenes. El método que emplearemos es una variante del conocido algoritmo EM

([DRR77, JJ94, MH97b, MH97a, CH98]). Pretendemos extender esta metodología para demostrar como la estimación de los parámetros de la transformación puede restringirse mediante el uso de estructuras relacionales en la función del algoritmo EM. Partiremos del marco convencional incorporando restricciones contextuales proporcionadas por las aristas de los grafos relacionales. El modelo usado para las restricciones se apoya en un modelo de compatibilidad presentado por Wilson and Hancock [WH96].

La estructura que seguiremos en este capítulo es la siguiente: en la sección primera se describirá la notación, seguidamente se explicará el proceso de extracción de los grafos, se continuará con una breve descripción del algoritmo EM generalizado, método en el que se inspira el proceso de encaje, después se explicará el procedimiento de alineamiento y se finalizará con los resultados experimentales.

4.2 Conceptos y notación

En esta sección se presentan los conceptos teóricos que vamos a utilizar posteriormente. El conjunto de vasos de la retina ha sido descrito mediante un grafo relacional. En la sección 4.2.1 se describe la transformación afín que se va a utilizar mientras que en la 4.2.2 se indicará como se construye el grafo relacional atribuido y como se aplica a nuestro problema.

4.2.1 Transformación afín de un conjunto de puntos

Se pretende estimar iterativamente los parámetros de la transformación geométrica $\Phi^{(n)}$ (donde n corresponde al índice de la iteración) que mejor transforma un conjunto de puntos característicos de una imagen \mathbf{w} en sus homólogos en el modelo \mathbf{z} . Para ello, representamos cada punto del conjunto de datos de la imagen por el vector de posiciones aumentado $\underline{\mathbf{w}}_i = (x_i, y_i, 1)^T$ dónde i es el índice del punto. Este vector representa las posiciones en dos dimensiones del punto en un sistema homogéneo de coordenadas. Asumiremos que todos estos puntos caen en un mismo plano en la imagen. Denotaremos el conjunto entero de puntos de la imagen por $\mathbf{w} = \{\underline{\mathbf{w}}_i, i \in \mathcal{D}\}$ donde \mathcal{D} es el conjunto de puntos. Los puntos correspondientes que constituyen el modelo son

igualmente representados por $\mathbf{z} = \{z_j, j \in \mathcal{M}\}$ donde \mathcal{M} denota el conjunto de índices para el modelo de puntos característicos z_j .

En este trabajo estamos interesados en las transformaciones afines. La transformación afín tiene seis parámetros libres. Estos modelizan las dos componentes de la traslación del origen en el plano de la imagen, la rotación global del sistema de coordenadas y el cambio de escala. Estos parámetros se pueden combinar de forma concisa en una matriz aumentada que toma la forma:

$$\Phi^{(n)} = \begin{pmatrix} \phi_{1,1}^{(n)} & \phi_{1,2}^{(n)} & \phi_{1,3}^{(n)} \\ \phi_{2,1}^{(n)} & \phi_{2,2}^{(n)} & \phi_{2,3}^{(n)} \\ 0 & 0 & 1 \end{pmatrix} \quad (4.1)$$

Con esta representación, la transformación afín de coordenadas se calcula utilizando la multiplicación de matrices

$$\mathbf{w}_i^{(n)} = \Phi^{(n)} \mathbf{z}_j \quad (4.2)$$

Claramente, el resultado de esta multiplicación nos da un vector de la forma $\mathbf{w}_i^{(n)} = [x, y, 1]^T$. El superíndice n indica que los parámetros son tomados de la n -ésima iteración del algoritmo. Nuestro objetivo es estimar los elementos $\phi_{i,j}^{(n)}$ de la matriz de parámetros $\Phi^{(n)}$, la cual describe una transformación de sistema de coordenadas que mejor encaja el conjunto de puntos de la imagen \mathbf{w} en los del modelo \mathbf{z} .

4.2.2 Grafos relacionales

Nos gustaría aprovechar las restricciones proporcionadas por la red de vasos sanguíneos para mejorar la estimación de los parámetros afines que encajan los pares de angiografías. Para formalizar este proceso, representaremos la estructura de vasos sanguíneos segmentados de la imagen utilizando una representación de grafos relacionales. La extracción de estas estructuras de las imágenes se describirá en la sección 4.3. Aquí los nodos serán cruces en la red de vasos sanguíneos. Las aristas, por otro lado, indicarán que existe una estructura vascular de conexión entre pares de nodos. Más formalmente, el conjunto de cruces serán los nodos del grafo de datos $G_D = \{\mathcal{D}, E_D\}$ y del grafo del

modelo $G_M = \{\mathcal{M}, E_M\}$. Aquí $E_D \subseteq \mathcal{D} \times \mathcal{D}$ y $E_M \subseteq \mathcal{M} \times \mathcal{M}$ son los conjuntos de aristas del grafo de datos y del grafo del modelo respectivamente. Para etiquetar los nodos no encajados en el grafo de datos, hemos añadido al conjunto de nodos del grafo del modelo el nodo nulo denotado por \emptyset . La clave de nuestro proceso de encaje es la idea de utilizar la estructura de los cruces de la red para restringir los encajes entre los dos conjunto de puntos. Este encaje se denota por la función $f : \mathcal{D} \rightarrow \mathcal{M} \cup \emptyset$ de los nodos de un grafo de datos a aquellos de un grafo modelo. De acuerdo con esta notación $f^{(n)}(i) = j$ indica que hay un encaje entre el nodo $i \in \mathcal{D}$ del grafo de datos y el nodo $j \in \mathcal{M}$ del modelo en la iteración n-ésima del algoritmo.

4.3 Extracción de los datos

En esta sección pretendemos describir la extracción de estos grafos a partir de nuestras imágenes originales. Si observamos nuestras imágenes encontramos un claro patrón de vasos sanguíneos repetido en la secuencia. Los vasos sanguíneos aparecen en nuestras imágenes con un nivel de intensidad mayor que el fondo. Además esta intensidad se va incrementando a medida que el colorante fluoresceínico, inyectado en la vena del paciente, llega a la retina. Por esta razón, el patrón de vasos sanguíneos irá cambiando de una imagen a la siguiente, ya que algunos vasos van perdiendo contraste mientras que otros aparecen más contrastados.

La segmentación de los vasos sanguíneos se realiza mediante un detector de líneas. Este detector ([KH89]) aplica filtros de líneas orientados en cuatro direcciones diferentes en la imagen para realzar los picos de intensidad. Las conexiones de las líneas encontradas de este modo se consiguen mediante un procedimiento de relajación probabilística que utiliza un diccionario de estructuras de líneas permitidas. Detalles sobre este trabajo se dan en [WH93]. El resultado de aplicar este detector a nuestras imágenes se muestra en la figura 4.3.

Los contornos de líneas se han segmentado limpiamente del fondo de las imágenes. Se puede observar que los patrones de venas segmentadas tienen una parte común pero existen partes que no aparecen en las dos imágenes debido, por un lado, al proceso de segmentación y, por otro, al movimiento del ojo del paciente. Puesto que el diccionario

del detector de líneas utilizado codifica los puntos de unión, estos puntos también se detectan.

Una vez realizada una segmentación de la red de vasos sanguíneos, se necesita una representación de esta red en forma de grafo. Con este propósito, se toman las uniones tipo T y los finales de línea que delimitan los segmentos de línea. Estos elementos son fácilmente identificables utilizando la representación presente en el detector de líneas para puntos unión. Las aristas del grafo corresponden a los vasos sanguíneos. Los puntos unión y final de línea se caracterizan mediante información extraída de las imágenes originales; cada punto se describe mediante el número de líneas que convergen en él y las líneas se describen mediante su longitud y dirección angular.

4.4 El algoritmo EM generalizado

El algoritmo EM es una técnica general para obtener el estimador de máxima verosimilitud de los parámetros de una distribución θ dado un conjunto de datos observados X ([JJ94, DRR77]). Consiste en un algoritmo iterativo compuesto por dos pasos fundamentales: la estimación (E) y la maximización (M). La M implica la búsqueda del máximo de la función de verosimilitud que es redefinida durante el paso E. Si el algoritmo en lugar de maximizar la función en el paso M simplemente la incrementa, el algoritmo toma el nombre de algoritmo generalizado EM. Esta será la versión del algoritmo utilizada en este capítulo.

El EM se aplica generalmente a aquellos problemas en los que la optimización de la función de verosimilitud $\mathcal{L}(\theta; X)$ se simplificaría si se conocieran los valores para un conjunto adicional de variables, llamadas variables ocultas. En este contexto, nos referiremos a los datos observados X como el conjunto de datos incompletos y denominaremos conjunto completo Y al que incluye las X y las variables ocultas Z . Se necesita especificar una distribución de probabilidad que una las variables ocultas con los datos reales: $P(Y, Z|X, \theta)$. El logaritmo de la densidad P define la verosimilitud del conjunto completo de datos. $\mathcal{L}_c(\theta; Y)$. La verosimilitud original, $\mathcal{L}(\theta; X)$, se le llama en este contexto verosimilitud de datos incompletos. La relación entre estas dos funciones de verosimilitud es la que motiva el algoritmo EM. Es importante notar que la verosimi-

litud del conjunto de datos completo es una variable aleatoria, porque las variables Z son realmente desconocidas. El algoritmo EM en primer lugar calcula el valor esperado de la verosimilitud del conjunto completo de datos, dados los datos observados y los parámetros del modelo actual. Este es el llamado paso E:

$$Q(\theta, \theta^{(n)}) = E[\mathcal{L}_c(\theta; Y)|X] \quad (4.3)$$

donde $\theta^{(n)}$ es el valor de los parámetros en la n -ésima iteración y la esperanza se calcula con respecto a estos parámetros. El paso M maximiza la función con respecto a θ para encontrar nuevos estimadores para los parámetros $\theta^{(n+1)}$:

$$\theta^{(n+1)} = \arg \max_{\theta} Q(\theta, \theta^{(n)}) \quad (4.4)$$

El algoritmo EM generalizado elige valores para los parámetros que incrementan el valor de Q , la esperanza de la verosimilitud completa. El efecto de este incremento en la verosimilitud de los datos incompletos es también un incremento:

$$\mathcal{L}(\theta^{(n+1)}; X) \geq \mathcal{L}(\theta^{(n)}; X) \quad (4.5)$$

Así la verosimilitud \mathcal{L} se incrementa monótonamente a lo largo de la secuencia de estimaciones de parámetros generadas por el algoritmo EM. En la práctica esto implica la convergencia a un máximo local.

4.5 El proceso de encaje

Una vez establecido el tipo de estructuras con las que se trabajará en este capítulo, se procederá a describir el proceso de alineamiento utilizado. Este procedimiento está inspirado en el conocido algoritmo EM generalizado en el que la transformación geométrica y el conjunto de correspondencias se estiman mediante un algoritmo iterativo.

El encaje de dos conjuntos de puntos restringido a una transformación geométrica es un problema que se ajusta muy bien a una resolución mediante el algoritmo EM. Estamos interesados en estimar la transformación afín Φ entre los dos conjuntos de puntos. Esto sería muy sencillo si previamente conociéramos las correspondencias entre los puntos. Sin embargo, este no es el caso. Deseamos calcular los parámetros de la



transformación a la vez que se obtienen los encajes entre el conjunto $\mathbf{w} = \{\mathbf{w}_i, i \in \mathcal{D}\}$ de datos y el conjunto $\mathbf{z} = \{\mathbf{z}_j, j \in \mathcal{M}\}$ del modelo. En este contexto, el conjunto de datos faltantes lo representa la función $f : \mathcal{D} \rightarrow \mathcal{M} \cup \emptyset$ y los parámetros del modelo corresponden a los de la transformación afín asumida entre las dos imágenes.

Asumimos que en principio cada nodo del grafo de datos \mathbf{w}_i puede asociarse con cada uno de los nodos \mathbf{z}_j con una cierta probabilidad. Bajo la hipótesis de independencia de los diferentes \mathbf{w}_i la verosimilitud adopta la siguiente forma:

$$P(\mathbf{w}|\Phi) = \prod_{i \in \mathcal{D}} P(\mathbf{w}_i|\Phi) \quad (4.6)$$

Cada \mathbf{w}_i puede corresponder a cualquier \mathbf{z}_j pero sólo a uno de ellos de donde:

$$P(\mathbf{w}_i|\Phi) = \sum_{j \in \mathcal{M}} P(\mathbf{w}_i, f(i) = j|\Phi) \quad (4.7)$$

y,

$$P(\mathbf{w}_i|\Phi) = \sum_{j \in \mathcal{M}} P(\mathbf{w}_i|f(i) = j, \Phi)P(f(i) = j|\Phi) \quad (4.8)$$

Podemos interpretar las probabilidades anteriores como una mixtura de densidades, donde el término $P(f(i) = j|\Phi)$ corresponde a la proporción de la densidad de probabilidad $P(\mathbf{w}_i|f(i) = j, \Phi)$. Utilizando las ecuaciones 4.6 y 4.8 la función de logverosimilitud se expresa de la siguiente forma:

$$\mathcal{L}(\Phi) = \sum_{i \in \mathcal{D}} \ln \sum_{j \in \mathcal{M}} P(\mathbf{w}_i|f(i) = j, \Phi)P(f(i) = j|\Phi) \quad (4.9)$$

Como ya hemos mencionado, el algoritmo EM propone un método mediante el cual se pueden encontrar iterativamente los parámetros que maximizan la verosimilitud. Vamos a utilizar el algoritmo EM en su versión para mixturas de densidades. En este caso, los datos faltantes corresponden a las variables que indican la pertenencia a una determinada clase o componente de la mixtura ([DRR77]). Interpretando la ecuación 4.8 como una mixtura de densidades, la esperanza de la función de logverosimilitud se expresa de la siguiente forma (ver apéndice B),

$$Q(\Phi^{(n+1)}, \Phi^{(n)}) = \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} P(f(i) = j|\mathbf{w}_i, \Phi^{(n)}) \{ \ln P(\mathbf{w}_i|f(i) = j, \Phi^{(n+1)}) + P(f(i) = j|\Phi^{(n)}) \} \quad (4.10)$$

La maximización de esta expresión implica maximizar dos términos. Sin embargo, el segundo término de la expresión corresponde a las proporciones del modelo de mixtura y no tiene una relevancia directa en la verosimilitud de los datos. De este modo, nos centraremos en la maximización de la siguiente función:

$$Q(\Phi^{(n+1)}, \Phi^{(n)}) = \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} P(f(i) = j | \mathbf{w}_i, \Phi^{(n)}) \ln P(\mathbf{w}_i | f(i) = j, \Phi^{(n+1)}) \quad (4.11)$$

En la fórmula anterior $P(f(i) = j | \mathbf{w}_i, \Phi^{(n)})$ es la probabilidad de las correspondencias entre los nodos del grafo modelo j y los nodos del grafo de datos i , dado el actual conjunto de parámetros estimados $\Phi^{(n)}$. Las proporciones de las componentes de la mixtura corresponden con las probabilidades de correspondencia entre los nodos i , j asumiendo la transformación $\Phi^{(n)}$. Estas probabilidades se actualizan en el paso de esperanza del algoritmo EM utilizando las restricciones relacionales proporcionadas por la estructura de cruces de las venas y arterias.

La cantidad $P(\mathbf{w}_i | f(i) = j, \Phi^{(n+1)})$ es la densidad de la posición aleatoria \mathbf{w}_i , a la que es transformado un nodo i del grafo de datos si ha sido originado por el nodo j del grafo modelo, situado en la posición \mathbf{z}_j bajo los parámetros actualizados $\Phi^{(n+1)}$. Estos parámetros se actualizan en el paso de maximización del algoritmo EM utilizando un procedimiento de agrupamiento difuso.

4.5.1 Esperanza

Pretendemos actualizar las probabilidades de encaje utilizando información contextual. Normalmente, el algoritmo EM realiza solamente una actualización nodo a nodo que ignora el contexto del objeto. Nuestra idea es introducir en el esquema de actualización de estas probabilidades información local en cada nodo del grafo. Wilson y Hancock [WH96] proponen un esquema de actualización de estas probabilidades que hace uso de las restricciones impuestas por los aristas del grafo. Este será el algoritmo de actualización utilizado introduciendo una pequeña variación. Puesto que, dada la naturaleza de nuestros grafos, las correspondencias deberían ser compatibles con una determinada transformación geométrica, parece razonable introducir en el proceso de actualización

de las probabilidades de encaje un término que de más peso a aquellas correspondencias que sean consistentes con la transformación y que penalice al resto. Así hemos introducido en el esquema original de Wilson y Hancock el término $P(\underline{w}_i | \underline{z}_j, \Phi_c^{(n)})$. Con este nuevo ingrediente las probabilidades $P(f(i) = j | \underline{w}_i, \Phi^{(n+1)})$ se actualizan de la siguiente forma:

$$P(f(i) = j | \underline{w}_i, \Phi^{(n+1)}) = P(\underline{w}_i | f(i) = j, \Phi^{(n)}) \times \frac{\prod_{I \in N_i} \sum_{J \in \mathcal{M} \cup \emptyset} P(\underline{w}_I | f(I) = J, \Phi^{(n)}) R_{ij}(I, J)}{\sum_{j \in \mathcal{M} \cup \emptyset} P(\underline{w}_i | f(i) = j, \Phi^{(n)}) \prod_{I \in N_i} \sum_{J \in \mathcal{M} \cup \emptyset} P(\underline{w}_I | f(I) = J, \Phi^{(n)}) R_{ij}(I, J)} \quad (4.12)$$

En la ecuación anterior $N_i = \{j : (i, j) \in E_D\}$ es el conjunto de nodos del grafo de datos conectados al nodo i por una arista (es decir, N_i es la vecindad del nodo i). La cantidad $R_{ij}(I, J)$ es un coeficiente de compatibilidad que mide la similitud entre las vecindades de los nodos i y j . Existen cuatro casos diferentes. En el primero de ellos, la arista del grafo de datos $(i, I) \in E_D$ se encaja con una arista consistente del grafo modelo $(j, J) \in E_M$. El segundo caso se da cuando uno de los nodos del borde del grafo de datos es encajado con el nodo nulo. El tercer caso ocurre cuando ambos nodos del borde son encajados con el nodo nulo. Finalmente, existe el caso de un encaje en el que (i, I) es una arista en el primer grafo y (j, J) no lo es. Abreviadamente:

$$R_{ij}(I, J) = \begin{cases} \frac{|\mathcal{M}|^2}{|E_M|} & \text{si } (i, j) \in E_m, \\ 1 & \text{si } (i, j) \in (\mathcal{M} \times \emptyset), \cup (\emptyset \times \mathcal{M}) \\ 1 & \text{si } (i, j) \in \emptyset \times \emptyset, \\ 0 & \text{si } (i, j) \in \mathcal{M} \times (\mathcal{M} - E_m), \end{cases} \quad (4.13)$$

La fórmula 4.12 necesita la especificación de un modelo para la distribución de las posiciones aleatorias de los datos. Aquí, hemos adoptado un modelo sencillo en el que los errores de alineamiento se asumen gaussianos, es decir,

$$P(\underline{w}_i | f(i) = j, \Phi^{(n)}) = \frac{1}{2\pi \sqrt{|\Sigma|}} \exp \left[-\frac{1}{2} (\epsilon_{ij}(\Phi^{(n)})^T \Sigma^{-1} (\epsilon_{ij} \Phi^{(n)})) \right] \quad (4.14)$$

donde $\epsilon_{ij}(\Phi^{(n)}) = \underline{w}_i - \Phi^{(n)} \underline{z}_j$ y Σ es la matriz de covarianzas para la transformación de los encajes según Φ^n .

Las probabilidades de encaje calculadas en el paso de la esperanza proporcionan un medio por el cual podemos asignar fuertes encajes entre los nodos del grafo del modelo y de los datos. En cada iteración asumimos como nodo correspondiente al i -ésimo del grafo de datos aquel del nodo del modelo con una probabilidad mayor. En otras palabras

$$g^{(n+1)}(i) = \arg \max_{j \in \mathcal{M} \cup \emptyset} P(f(i) = j | \mathbf{w}_i, \mathbf{z}_j, \Phi_c^{(n+1)}) \quad (4.15)$$

Esta función $g^{(n)}(i)$ nos proporciona un modo de tener en cada iteración n la correspondencia más probable.

4.5.2 Maximización utilizando agrupamiento difuso

Con el conjunto de correspondencias fuertes fijadas, podemos concentrarnos en la estimación de la transformación afín requerida para representar el alineamiento de las dos imágenes. De acuerdo con el algoritmo EM, el proceso de estimación de parámetros implica encontrar la transformación afín que minimiza la cantidad

$$Q(\Phi^{(n+1)}, \Phi^n) = \sum_{(i,j) \in g^{(n+1)}} (\mathbf{w}_i - \Phi^{(n)} \mathbf{z}_j)^T \Sigma^{-1} (\mathbf{w}_i - \Phi^{(n)} \mathbf{z}_j) \quad (4.16)$$

La estimación de los parámetros mediante la minimización de la función anterior depende en gran medida de las probabilidades iniciales utilizadas puesto que si se calcula usando unas correspondencias iniciales completamente incorrectas dará errores grandes

Para evitar esto, la estimación de la transformación se realiza mediante un método de agrupamiento no supervisado en el espacio de parámetros de transformaciones afines ([KRD80, SKB80]). El grupo con menor dispersión representa la transformación que mejor alinearía los dos grafos.

El conjunto de datos se obtiene calculando para cada terna de correspondencias posibles en la iteración n dadas por $\{(\mathbf{w}_i, \mathbf{z}_j) : g^n(i) = j\}$ una transformación afín. Aquellas correspondencias $g^{(n)}(i) = j$ que son correctas deberían definir una transformación afín común entre las posiciones de los nodos de los grafos, mientras que aquellas que son

erróneas generarán transformaciones distintas, que se verán como puntos dispersos en el espacio de parámetros de las transformaciones. Por lo tanto, parece razonable la estimación de la transformación geométrica buscando un grupo denso en el espacio de parámetros afines.

Debido al gran número de transformaciones obtenidas de este modo es conveniente reducirlo. Con este propósito, hemos aplicado restricciones físicas para eliminar aquellas que son imposibles por la naturaleza del problema. En concreto se han eliminado del conjunto de datos aquellas que implican un desplazamiento fuera de los límites de la imagen.

Una vez obtenido el conjunto de datos, es decir, un conjunto de transformaciones se procede a su agrupamiento en el espacio de parámetros. Se ha elegido el procedimiento de agrupamiento difuso llamado Gustafson-Kessel [Bez81] que es una extensión del algoritmo K-medias. La principal idea del agrupamiento difuso consiste en la especificación de los pesos de pertenencia de los datos a dos o más clases. La pertenencia de un dato i a una clase j es proporcional a una medida de distancia d_{ij} a los centros de las clases. Sea $T = \{t_j : j \in N\}$ nuestro conjunto de datos. En nuestro problema concreto T corresponde al conjunto de transformaciones obtenidas de la forma descrita anteriormente. Los datos t_j serán vectores en un espacio de seis dimensiones puesto que las transformaciones tienen seis parámetros libres. Sea $V = \{v_i : i = 1 \dots C\}$ el conjunto de los centros de las clases en el espacio de parámetros. Las pertenencias difusas de cada una de estas clases se determinan mediante el conjunto de distancias $d(t_j, v_i)$ entre los datos y los centros de las clases actuales. Estas pertenencias difusas se almacenan en una matriz $C \times N$, $U = [u_{ij}]$. Para cada dato, la suma de las pertenencias extendida a todas las clases es la unidad,

$$\sum_{i=1}^C u_{ij} = 1, \forall j = 1 \dots N \quad (4.17)$$

El agrupamiento difuso es un procedimiento iterativo cuyo objetivo es la minimización de la siguiente función:

$$J(U, V : T) = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m d^2(t_j, v_i) \quad (4.18)$$

donde m es una constante.

La métrica para el cálculo de las distancias en el algoritmo de Gustafson y Kessel es diferente para cada clase. Estas métricas se van calculando en cada iteración del algoritmo como aquellas que minimizan la función objetivo anterior.

Este procedimiento nos permite agrupar el conjunto de transformaciones en varias clases. Una vez realizado el agrupamiento en el espacio de parámetros, es necesario decidir qué clase es la buscada. Con este propósito se selecciona una posición espacial sobre la que se aplicarán todas las transformaciones de cada clase. De este modo, obtenemos una nueva posición espacial para cada una de las transformaciones. Se calcula la matriz de covarianzas de este nuevo conjunto de puntos. La clase que ha generado puntos con varianzas menores se toma como correcta y los parámetros estimados corresponden al centro (en el espacio de transformaciones) de dicha clase (media ponderada con los grados de pertenencia).

4.6 Resultados experimentales

Un ejemplo del tipo de imágenes con las que vamos a trabajar se muestra en la figura 4.1. Lo más característico de estas imágenes es la red de vasos sanguíneos. Los vasos en la imagen de la derecha están extremadamente borrosos. Nuestro objetivo es alinear estas imágenes utilizando los cruces de los vasos como puntos característicos.

Como ya dijimos, el primer paso era extraer el sistema de vasos sanguíneos de los angiogramas utilizando un procedimiento de detección de líneas por relajación. Este algoritmo utiliza un diccionario de estructuras de contorno para resaltar las estructuras de líneas conectadas. La imagen segmentada donde aparecen las venas y arterias se muestra en 4.3. La reconstrucción de los vasos que aparecen borrosos en la imagen de la derecha es satisfactoria.

Construidos los grafos, el método propuesto requiere unas probabilidades de encaje iniciales. Estas probabilidades se han calculado mediante medidas tomadas para cada par de nodos i y j de los dos grafos. Concretamente se han utilizado el número de aristas que convergen en un nodo, la longitud de cada una de estas aristas y los ángulos que forman.

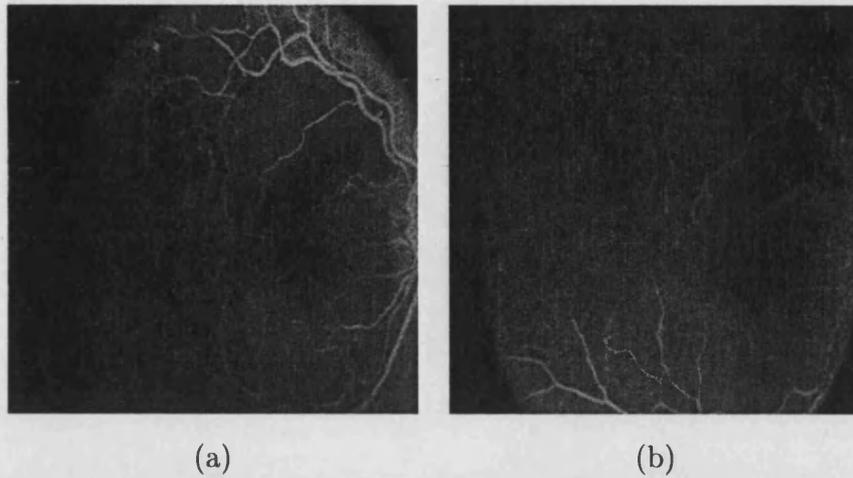


Figura 4.1: Imágenes originales (a) primera image y (b) segunda imagen de la secuencia digital

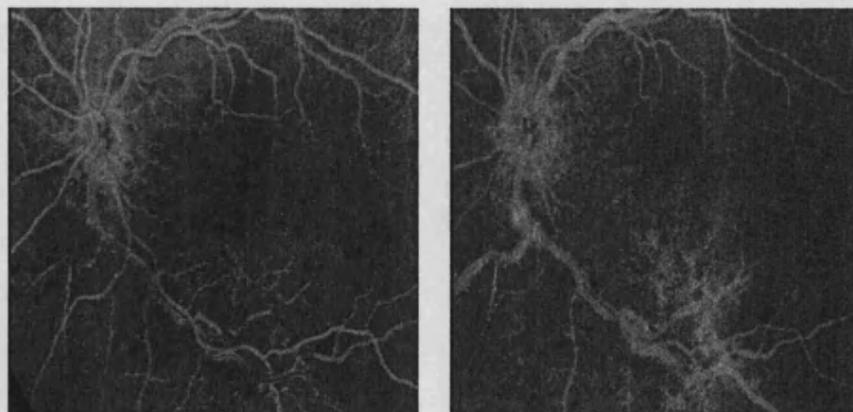


Figura 4.2: Secuencia de angiografías sobre la que se aplicado el algoritmo propuesto

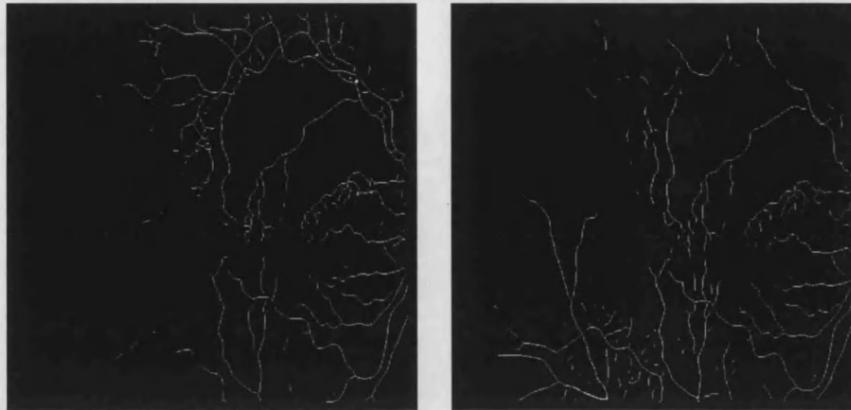


Figura 4.3: Vasos sanguíneos segmentados de (a) la primera image y (b) la segunda imagen de la secuencia digital

Hemos aplicado el proceso de encaje de grafos basado en relajación probabilística ([WH96]) a dos imágenes de una secuencia. Como se ve en la figura 4.4 el algoritmo obtiene además de un gran número de encajes correctos, otros muchos erróneos que pueden ser eliminados si se les aplican restricciones asociadas con una transformación. En la figura 4.6, se muestran los resultados del encaje obtenidos mediante el proceso de alineamiento propuesto. Las líneas representan puntos correspondientes. Solamente existe un encaje erróneo después de la convergencia del algoritmo EM.

La figura 4.5 representa otro ejemplo en el que se han obtenido los resultados esperados tras la aplicación del algoritmo EM generalizado. Por el contrario, la figura 4.7 muestra un caso en el que aparecen correspondencias erróneas tras la convergencia del algoritmo. Es importante destacar que la convergencia correcta del algoritmo EM va a depender en gran medida de las correspondencias iniciales. Si las probabilidades iniciales son erróneas, el procedimiento de agrupamiento no será capaz de obtener una buena estimación de la transformación y en consecuencia, la actualización de los encajes no mejorará. Otro punto importante a tratar es el número de clases considerado en el espacio de las transformaciones. En este ejemplo se han considerado cuatro clases. La utilización de un número diferente de clases, en principio no afectaría a la estimación de la transformación. La distribución de las transformaciones no es conocida por lo que podrían haber varios grupos. Lo único relevante para el algoritmo es que exista un grupo bastante numeroso y denso. La figura 4.8 representa la distribución de las

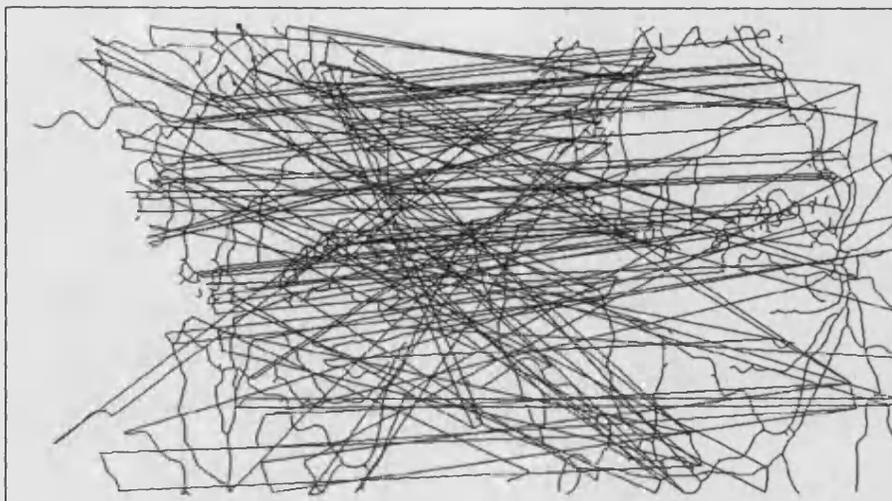


Figura 4.4: Resultado del encaje mediante un método basado en relajación probabilística

transformaciones en el espacio paramétrico para la figura 4.6. Como se ve existe una clara acumulación alrededor del origen.

La transformación estimada con este procedimiento se ha comparado con la obtenida fijando tres correspondencias manuales en cada imagen y el resultado ha sido correcto en 6 de las 10 imágenes estudiadas. La convergencia es muy rápida, en los casos estudiados una media de 8 iteraciones han sido suficientes (aproximadamente 100 s). Sin embargo, en 4 de los casos estudiados la transformación obtenida era incorrecta y, por lo tanto, aparecían correspondencias erróneas. Este funcionamiento es debido, como ya hemos dicho, al efecto de la inicialización de las probabilidades de encaje. El procedimiento muestra, pues, una clara dependencia de su inicialización.

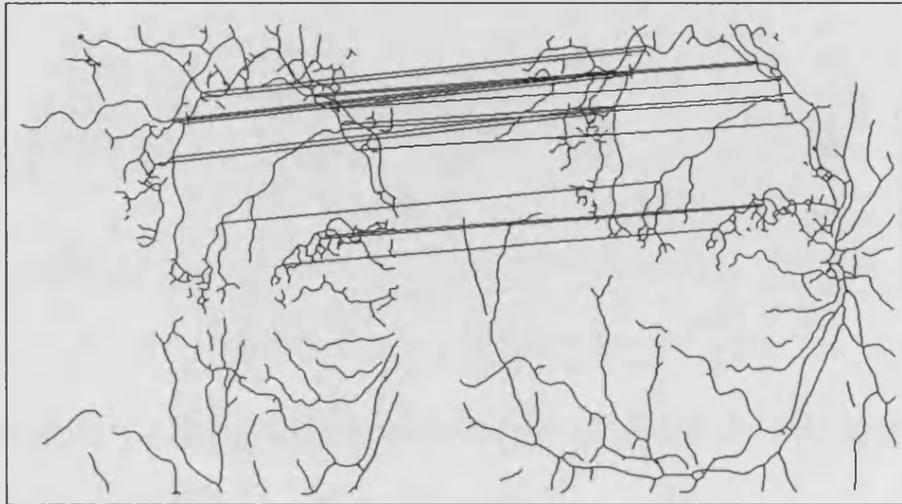


Figura 4.5: Correspondencia obtenidas para dos diferentes imágenes de la misma secuencia

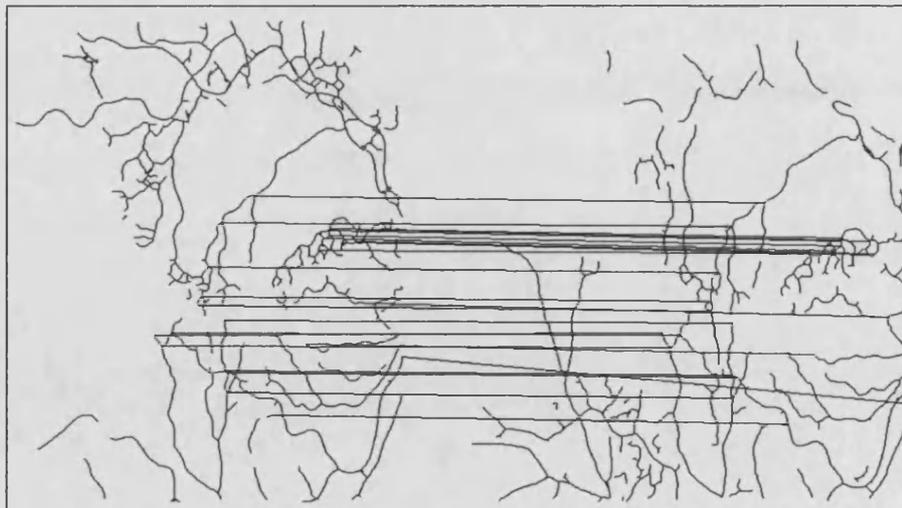


Figura 4.6: Resultado del encaje con el algoritmo generalizado EM

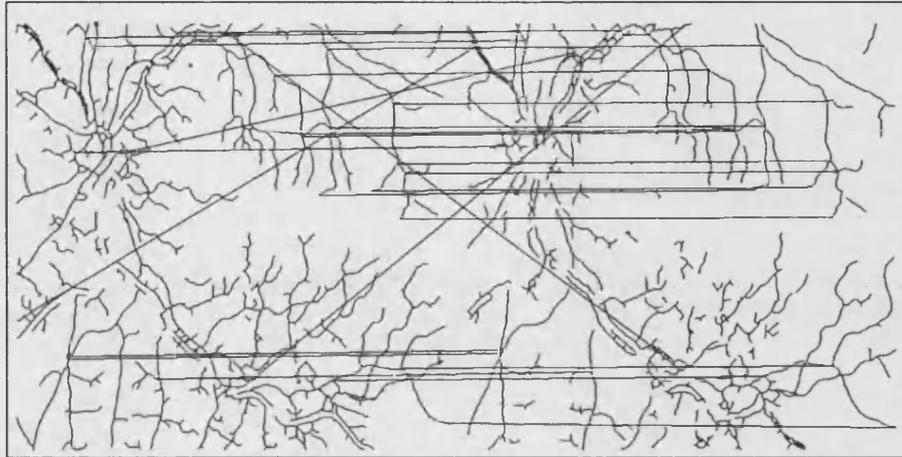


Figura 4.7: Ejemplo en el que el algoritmo obtiene algunas correspondencias erróneas

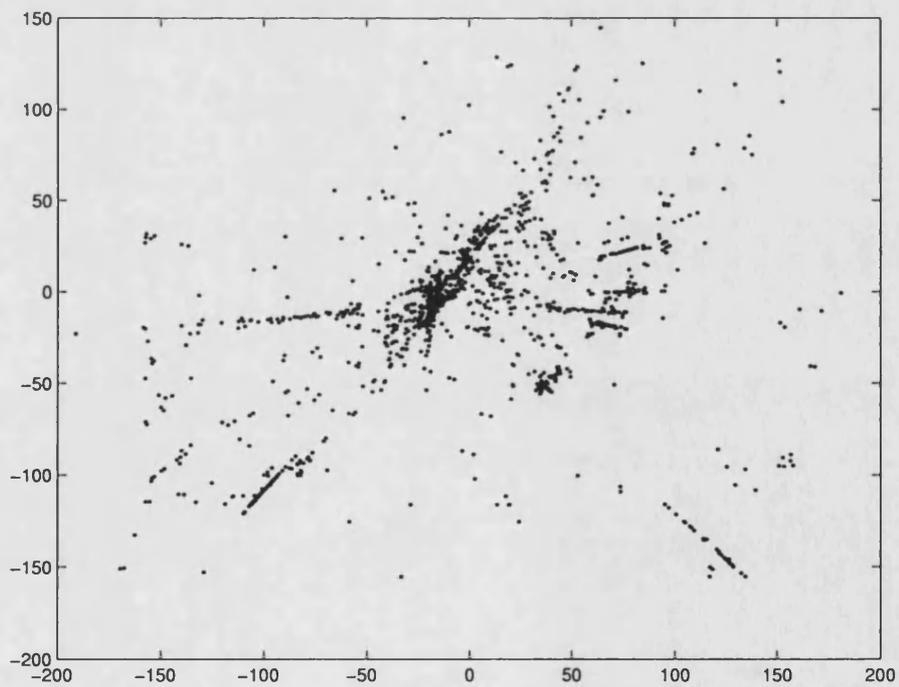


Figura 4.8: Distribución de las transformaciones en el espacio de parámetros

Capítulo 5

Modelo de transformación no lineal para alineamiento de angiografías

5.1 Introducción

En los dos capítulos anteriores se han propuesto dos métodos alternativos para el encaje de angiografías fluoresceínicas. Estos dos métodos asumen que el movimiento existente en una secuencia se puede considerar descrito por una transformación afín. Sin embargo, puesto que cada imagen de la secuencia corresponde a una proyección plana de la superficie curva del fondo ocular, la transformación entre imágenes consecutivas no corresponderá exactamente a la de un movimiento rígido. Por esta razón se pretende dar un modelo de transformación geométrica mediante una transformación que tenga en cuenta el movimiento real del globo ocular y la proyección plana de la superficie de la retina que supondremos aproximadamente esférica. Este nuevo modelo se fundamenta en el proceso físico de la captura de imágenes así como en el movimiento del globo ocular.

En la literatura se pueden encontrar varias aproximaciones que tratan el problema del alineamiento de imágenes cuando las superficies no se comportan como sólidos rígidos. Una aproximación general consiste en utilizar una expresión polinómica de un grado superior al lineal para la transformación que de cuenta de pequeñas distorsiones. La ventaja de esta aproximación radica en que la construcción de esta función

polinómica dado un conjunto de puntos de control se consigue mediante la resolución de un sistema lineal de ecuaciones ([JBBL87]).

Sin embargo, en muchos casos una expresión polinómica no es suficiente para suplir la transformación compleja requerida. Es en estos casos cuando se utilizan funciones de interpolación más o menos complejas para llevar a dos conjuntos de puntos de control hasta un solapamiento completo. En esta línea de trabajo se encuentra un método para el encaje de superficies con claras deformaciones mediante un conjunto de funciones base para la interpolación conocidas como *Thin-Plate Splines* propuesto por Bookstein ([Boo89]).

Otra aproximación a este mismo problema consiste en dar un modelo de transformación afín para la totalidad de la imagen y sólo en aquellos puntos donde existe una clara distorsión aplicar una transformación local diferente. Pérez [PPJ⁺99] sigue esta línea y propone un método para el encaje de imágenes que no pueden ser modelizadas según una transformación afín global. La esencia del algoritmo consiste en encontrar primeramente una transformación lineal global entre dos conjuntos de puntos de interés obtenidos de dos imágenes, en buscar a continuación los puntos con errores mayores que un cierto umbral, y finalmente encontrar una transformación afín diferente para esos puntos utilizando información del entorno.

Si nos centramos en el alineamiento de angiografías, la transformación más comúnmente utilizada es la transformación afín. Cideciyan [Cid95] presenta un algoritmo para el encaje de angiografías que utiliza la correlación cruzada entre las transformadas de Fourier en coordenadas log-polares para encontrar la rotación y el cambio de escala y otra correlación cruzada entre las imágenes originales para obtener la translación. Jagoe [JB⁺90] utiliza una función polinómica de primer orden como modelo de transformación. Las correspondencias entre puntos se obtienen mediante la comparación de los triángulos sostenidos por cada conjunto de tres puntos en cada una de las imágenes. Para reducir el número de combinaciones se aplican restricciones de movimiento. La transformación afín se calcula mediante un ajuste por mínimos cuadrados dado el conjunto de correspondencias. Otro método propuesto para el encaje de imágenes de fondo ocular se debe a Hart [HG94]. Hart presenta un procedimiento que asume también una transformación afín. Para estimar la transformación afín se necesita un conjunto de correspondencias

obtenidas probando todas las combinaciones y se descartan aquellos encajes con un desplazamiento mayor que un valor dado. De todas las posibles correspondencias se utilizarán aquellas con una correlación cruzada más alta para la estimación de la transformación.

Vemos que todos los métodos propuestos para el encaje de angiografías asumen una transformación afín como modelo. Parece, por lo tanto, necesario justificar el uso de una nueva transformación. Pretendemos, por lo tanto, construir una transformación que reduzca el error de posición cometido en el encaje de imágenes con respecto a la transformación afín o, al menos, justificar el uso generalizado de la transformación afín para el encaje de angiografías. Puesto que las distorsiones locales en nuestras imágenes proceden de que el ojo es aproximadamente esférico y las imágenes que analizamos corresponden a su proyección plana aprovecharemos esta información para la obtención de un modelo geométrico.

Existe un conjunto de ventajas derivadas de la utilización de esta transformación geométrica en el alineamiento de imágenes: (i) Los parámetros de este modelo tienen un significado intuitivo, relacionado con el proceso de captura de imágenes, y (ii) el movimiento del ojo se describe únicamente con dos parámetros mientras que eran seis los necesarios para la especificación completa del movimiento mediante una transformación afín.

En las secciones siguientes veremos en detalle el método propuesto para el alineamiento de imágenes. Comenzaremos con la derivación del modelo geométrico asumido en la sección 5.2. Continuaremos en 5.3 con una breve descripción de los parámetros de esta transformación. En la sección 5.4 se explicará el algoritmo de alineamiento utilizado en este capítulo y finalizaremos con las secciones 5.5 y 5.6 en las que se mostrarán y discutirán los resultados experimentales obtenidos.

5.2 Modelo de transformación

Una transformación geométrica entre dos imágenes puede ser de varios tipos: rígida, afín, de proyección o deformable. Estas categorías indican el grado de elasticidad de la

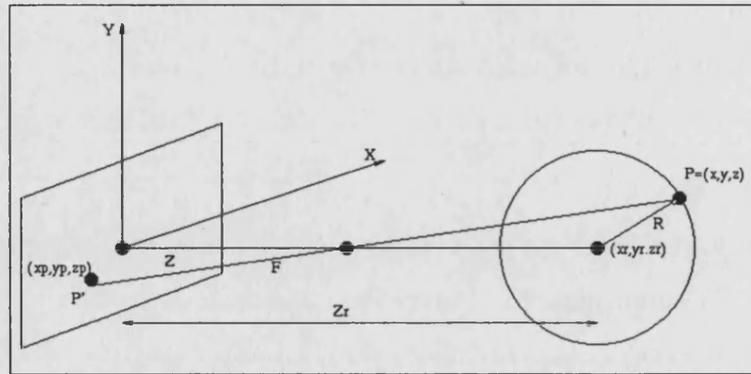


Figura 5.1: Los dos sistemas de coordenadas de la imagen y su posición relativa.

transformación. Una transformación rígida corresponde a un movimiento de rotación y translación entre un par de imágenes y está definida por cuatro parámetros. La transformación afín, más general que la anterior, corresponde a una transformación donde una línea recta se convierte en otra línea recta y el paralelismo entre rectas se conserva. Tiene seis parámetros libres, lo que la hace menos restrictiva que la rígida y además posee la propiedad de ser lineal si se expresa en coordenadas homogéneas. Una transformación proyectiva (llamada también transformación de perspectiva) convierte una línea recta en otra pero no conserva el paralelismo. No es lineal, por lo que la estimación de parámetros de la transformación es una tarea más difícil. Por último, existen las que hemos llamado deformables que incluyen al resto. Este tipo de transformación geométrica convierte una línea en una curva ([Wol90]).

La transformación elegida para este trabajo podemos clasificarla en la categoría de transformaciones deformables, puesto que, como veremos, es una transformación no lineal que no transforma una línea recta en otra. Esta transformación captura el movimiento de la escena (en este caso, el fondo ocular) proyectado sobre una imagen plana en diferentes instantes de tiempo. Se sabe que el movimiento del ojo humano se restringe a rotaciones, y que estas rotaciones pueden reproducirse aproximadamente por medio de giros alrededor de los tres ejes de un sistema de coordenadas cartesiano con el origen en el centro del ojo.

Las imágenes que queremos encajar son de fondo de ojo el cual podemos suponer idealmente esférico, y de este modo asumir que los cambios de posición de un elemento en las dos imágenes vienen dados por rotaciones del ojo con respecto a los tres ejes

perpendiculares de un sistema de coordenadas cuyo origen se encuentra en el centro del ojo. A estos ángulos los denotaremos por (α, θ, γ) . Además del movimiento de la escena, se tiene que tener en cuenta la proyección perspectiva ([Gon87, HS92b, Sch89]) de la escena en el plano imagen, que depende del sistema óptico utilizado en el proceso de captura.

La transformación perspectiva juega un papel central en el procesamiento de imágenes puesto que proporciona una aproximación al modo en el que se forma una imagen a partir de una escena tridimensional. Como hemos mencionado, este tipo de transformaciones no son lineales ya que involucran divisiones por alguna de las variables. Definimos las coordenadas de la cámara (X, Y, Z) de forma que el plano de la imagen coincida con el plano XY y que el eje óptico del sistema (centro de la lente) coincide con el eje Z . El centro del plano de la imagen está situado en el origen de coordenadas y el centro de la lente en la coordenada $(0, 0, F)$. Esta distancia F es la distancia focal de la lente si la cámara está enfocada a objetos distantes. Expresaremos esa distancia como f_0 .

Sea (x, y, z) las coordenadas de un punto cualquiera de la escena respecto al sistema de coordenadas que se muestra en la figura 5.1. Queremos encontrar la relación entre un punto (x, y, z) de la escena y su proyección en el plano de la imagen (x_p, y_p) . Esto se puede deducir fácilmente de la figura por semejanza de triángulos y viene dado por la siguiente ecuación:

$$x_p = \frac{-f_0 x}{z_r + z} \quad (5.1)$$

$$y_p = \frac{-f_0 y}{z_r + z} \quad (5.2)$$

Asumimos que el movimiento de la escena se debe a rotaciones alrededor de los tres ejes de coordenadas de un sistema centrado en el centro del globo ocular. De esta manera la relación entre un punto (x, y, z) y ese mismo punto cuando el objeto se ha girado según unos ángulos se obtiene aplicando la matriz producto $T = T_\alpha T_\theta T_\phi$ donde las matrices T_α , T_θ , y T_ϕ , representan las transformaciones de un punto cuando se gira un ángulo α en torno al eje X , θ alrededor del eje Y , y γ en torno al eje Z respectivamente. La matriz de rotaciones se obtiene de este modo como producto de

las tres matrices individuales $\mathbf{T} = \mathbf{T}_\alpha \mathbf{T}_\theta \mathbf{T}_\gamma$:

$$\mathbf{T} = \begin{pmatrix} \cos \theta \cos \phi & \sin \alpha \sin \theta \cos \phi + \cos \alpha \sin \phi & -\cos \alpha \sin \theta \cos \phi + \sin \alpha \sin \phi \\ -\cos \theta \sin \phi & -\sin \alpha \sin \theta \sin \phi + \cos \alpha \cos \phi & \cos \alpha \sin \theta \sin \phi + \sin \alpha \cos \phi \\ \sin \theta & -\sin \alpha \cos \theta & \cos \alpha \cos \theta \end{pmatrix}$$

Así, la relación entre la posición de un punto (x, y, z) y ese mismo punto cuando un objeto rígido es rotado un cierto ángulo se puede descomponer como rotaciones alrededor de los tres ejes perpendiculares de un sistema de coordenadas y se expresa como un producto,

$$\vec{w}' = \mathbf{T}\vec{w} \quad (5.3)$$

donde $\vec{w} = (x, y, z)$ y $\vec{w}' = (x', y', z')$ son los vectores de posición de un punto en la escena antes y después de la rotación. Esta ecuación expresa el movimiento en la escena de un determinado punto de la retina. Sin embargo, buscamos la misma relación pero en el plano imagen, que es la relación entre un punto en una imagen I_i y su correspondiente en otra imagen I_j en la que el ojo ha sufrido una rotación. Utilizando las ecuaciones de la transformación de perspectiva que relacionan un punto en la escena y en la imagen y las anteriores se llega a la siguiente relación entre un punto (x_p, y_p) en I_i y otro punto (x_s, y_s) en I_j de una secuencia dada.

$$x_p = \frac{x_s \cos \theta \cos \gamma + y_s (\sin \alpha \sin \theta \cos \gamma + \cos \alpha \sin \gamma) + \frac{f_0 z}{z_r + z} (\cos \alpha \sin \theta \cos \gamma - \sin \alpha \sin \gamma)}{\frac{z_r}{z_r + z} - \frac{x_s}{f_0} \sin \theta + \frac{y_s}{f_0} + \frac{z}{z_r + z} \cos \alpha \cos \theta} \quad (5.4)$$

$$y_p = \frac{-x_s \cos \theta \sin \gamma + y_s (\cos \alpha \cos \gamma - \sin \alpha \sin \theta \sin \gamma) - \frac{f_0 z}{z_r + z} (\cos \alpha \sin \theta \sin \gamma + \sin \alpha \cos \gamma)}{\frac{z_r}{z_r + z} - \frac{x_s}{f_0} \sin \theta + \frac{y_s}{f_0} + \frac{z}{z_r + z} \cos \alpha \cos \theta} \quad (5.5)$$

En estas dos ecuaciones todavía aparece una dependencia con la coordenada z de cada punto en la escena real, una dependencia con la profundidad del punto proyectado. Las imágenes con las que trabajamos son imágenes de fondo de ojo y suponemos que el ojo es idealmente esférico. Por lo tanto, las coordenadas de un punto (x, y, z) en la escena real están sujetos a una relación geométrica puesto que estos puntos están situados sobre una superficie esférica $R_0^2 = x^2 + y^2 + z^2$ donde R_0 es el radio del ojo.

Así, podemos escribir la coordenada z de un punto en función de R_0 , x e y . Además de esto, estas dos variables están relacionadas con la proyección perspectiva en el plano imagen, $x = x(x_p, f_0, z_r)$, $y = y(y_p, f_0, z_r)$. Podemos reemplazar z por la siguiente expresión:

$$z = -z_r \left(\frac{x_p^2 + y_p^2}{f_0^2} + \sqrt{\left(\frac{x_p^2 + y_p^2}{f_0^2} - \frac{x_p^2 + y_p^2 + f_0^2}{f_0^2} \right) \left(\frac{x_p^2 + y_p^2}{f_0^2} - \frac{R_0^2}{f_0^2} \right)} \right) \quad (5.6)$$

Las ecuaciones 5.4 y 5.5 constituyen la transformación geométrica entre imágenes de fondo de ojo y las denotamos por $\Phi = (\Phi_x, \Phi_y)$ que depende de seis parámetros. Es importante destacar que puesto que hemos supuesto un modelo simplificado del sistema óptico de una cámara de fondo de ojo real (una descripción detallada del sistema de cámaras de fondo aparece en el capítulo 2) los parámetros relacionados con el sistema óptico no se podrán entender como los parámetros reales de ese sistema.

5.3 Estimación de parámetros

Aparte de los parámetros relacionados con el movimiento real del ojo, que son los tres ángulos de rotación, el resto de parámetros que aparecen en la transformación geométrica se pueden medir experimentalmente: el radio R_0 del ojo, la distancia focal f_0 que es una característica intrínseca de la lente de la cámara y la distancia z_r entre la cámara y el centro del ojo. Esto permite tener una buena estimación de estos tres parámetros. Sabemos, por ejemplo, el radio promedio del globo ocular y que la distancia focal tomará valores en un intervalo no muy amplio.

Las ecuaciones de la transformación tienen en cuenta los movimientos de rotación del ojo. Sin embargo, según los estudios llevados a cabo por Tweed y colaboradores ([TV90]), el espacio de rotaciones permitidas durante los movimientos conocidos como sacadas (movimientos rápidos del ojo) está restringido a un subespacio de dos dimensiones. Este resultado, conocido como la ley de Listing, se puede describir de forma sencilla si expresamos las posiciones del ojo en términos de los ejes de los desplazamientos angulares desde una posición particular conocida como posición primaria. Entonces la ley de Listing establece que el ojo puede moverse solamente a aquellas posiciones

que pueden alcanzarse desde la posición primaria por una única rotación sobre un eje situado en el plano de Listing, que es ortogonal a la dirección de la mirada en la posición primaria. Si suponemos que el paciente está mirando en una dirección alineada con la cámara en el momento de la captura de las imágenes, podemos reducir la dimensión del conjunto de parámetros que deben ser estimados en el modelo de transformación no lineal, reemplazando el ángulo γ por el valor $\gamma = 0$. Así, tenemos que estimar cinco parámetros y tres de ellos son aproximadamente conocidos lo que nos permite trabajar con unos buenos estimadores iniciales.

5.4 El proceso de alineamiento

En el capítulo anterior se describió un algoritmo para el encaje de imágenes utilizando grafos relacionales como estructuras básicas para el proceso de encaje.

En el capítulo presente, se pretende desarrollar un algoritmo que nos permita encontrar un conjunto de encajes correctos así como estimar los parámetros de la transformación de coordenadas entre diferentes imágenes de una secuencia. El algoritmo parte de un conjunto de encajes iniciales, correctos o no, y proporciona como resultado aquellos encajes que son compatibles con la transformación propuesta.

Utilizaremos los grafos relaciones con los que trabajabamos en el capítulo anterior extraídos de las imagenes originales. Un grafo relacional consiste en un conjunto de nodos caracterizados por su posición de imagen (x_p, y_p) que corresponden a los cruces de vasos en la imagen. Dos nodos del grafo están relacionados si están conectados por un vaso sanguíneo en la imagen original. Así, hemos construido un grafo donde los nodos son cruces de venas y las relaciones corresponden a los nodos unidos por el mismo vaso sanguíneo.

Una vez que los grafos se extraen de dos imágenes diferentes, procedemos a realizar el encaje. Hemos elegido un algoritmo de relajación para encaje de grafos propuesto por Wilson y Hancock ([WH96]). Cuando aplicamos este algoritmo a angiografías fluoresceínicas, todos los encajes obtenidos no son correctos. Esto se produce porque los cambios de intensidad en las angiografías son bastante fuertes de una imagen a la

siguiente en una secuencia. Estos cambios son debidos al líquido de contraste inyectado para la observación de los vasos sanguíneos en la retina que hará que el nivel de intensidad en unos vasos se incremente en ciertas zonas y se decremente en otras. Este efecto provocará una extracción de los grafos peor de la esperada y, en consecuencia, un gran porcentaje de encajes erróneos.

Así nos proponemos desarrollar un algoritmo que permita eliminar todos los encajes falsos obtenidos por el proceso de relajación a la vez que estimamos la transformación geométrica entre imágenes.

Sea $A = \{\vec{a}_i : i = 0 \dots N\}$ el conjunto de nodos de un grafo extraído de la primera imagen y $B = \{\vec{b}_i : i = 0 \dots N\}$ el conjunto de nodos que pertenecen al grafo que representa a la segunda. Empezamos con un encaje inicial entre estos dos conjuntos A y B , que llamamos $C = \{(\vec{a}_i, \vec{b}_i) : i = 0 \dots N\}$ que se ha obtenido por medio de un paso anterior de preprocesado. Sea $W = \{w_i : i = 0 \dots N\}$ un conjunto de variables reales en el intervalo $0 \leq w_i \leq 1$. Estas variables indican si un encaje i debería borrarse del conjunto de encajes C . Cuanto más cerca está w_i de 1, mejores se consideran las correspondencias. Iniciaremos todas estas variables w_i , que llamaremos pesos, a 1, puesto que no tenemos ninguna evidencia de que alguno de los encajes sea incorrecto.

La estimación de los parámetros se basa en encontrar un valor para cada parámetro en la transformación que maximice o minimice una determinada función objetivo. Hemos elegido como función objetivo en el proceso de optimización el error cuadrático ponderado. La función coste será el sumatorio extendido a todos los residuos cuadráticos entre las posiciones de los nodos en una imagen y las correspondientes posiciones transformadas $\|\vec{a}_i - \Phi(\vec{b}_i)\|^2$ ponderado con las variables w_i , que pueden interpretarse como una medida de cuán correcto es un encaje para una transformación dada.

$$\Phi = \arg_{\varphi} \min \sum (w_i \|\vec{a}_i - \varphi(\vec{b}_i)\|^2) \quad (5.7)$$

Nos proponemos dos objetivos diferentes. El primero es borrar del conjunto de correspondencias aquellos encajes que son incorrectos. El segundo es estimar los parámetros de la transformación geométrica. El método que utilizaremos consiste en un algoritmo iterativo con dos pasos fundamentales: la estimación de los parámetros por medio de la

minimización de una función objetivo (función coste) y la actualización de las variables w_i utilizando una función cuyo resultado se aproxima a 1 cuando $\|\vec{a}_i - \varphi(\vec{b}_i)\|^2$ es muy pequeño.

Estimación de los parámetros de la transformación geométrica: El proceso de optimización puede ser enfocado de múltiples formas. La función coste a minimizar (el error cuadrático ponderado) no es lineal por lo que necesitamos un método numérico que nos permita encontrar los parámetros para los que la función toma el valor mínimo. Existen varios algoritmos para este propósito. En general, los métodos de optimización se puede clasificar en dos grandes grupos: los que necesitan conocer el gradiente de la función objetivo y los que no.

El método del gradiente conjugado ([PTVB92]) pertenece al primer grupo. Este algoritmo, cuyo comportamiento es teóricamente mejor que el clásico descenso por gradiente, utiliza el gradiente para obtener un conjunto de direcciones ortogonales entre sí en las que se aplicará una minimización lineal. La obtención de tales vectores de direcciones está restringida a la dimensión del espacio de la función coste, puesto que sólo podremos construir una secuencia de n elementos de vectores ortogonales entre si en un espacio n -dimensional. De este modo, si la dimensión de la función que deseamos optimizar no es suficientemente grande, el conjugado del gradiente puede obtener resultados erróneos. Para la minimización del error cuadrático ponderado utilizamos, en primer lugar, el método del descenso conjugado del gradiente. Sin embargo, en algunos casos, tal método parece caer en un mínimo local en lugar del mínimo global que estamos buscando. Como dijimos antes, este comportamiento puede deberse a la pequeña dimensión de la función objetivo. Otro problema al que nos enfrentamos es el pequeño valor de algunas de las componentes del vector gradiente relativas al resto por lo que el movimiento en esas direcciones en las sucesivas iteraciones es muy pequeño y practicamente el valor de los parámetros correspondientes no se modifica con respecto a los valores de partida.

El método de Powell ([PTVB92]) para optimización se clasifica en el grupo de los que no se sirven del gradiente para obtener el mínimo. Este método parte de un conjunto de n direcciones ortogonales $\mathbf{e}_1 \cdot \mathbf{e}_n$ entre sí, donde n corresponde a la dimensión del espacio de búsqueda. El método realiza una optimización lineal

en cada una de las direcciones ortogonales. Después de n pasos los vectores de direcciones $e_1 \dots e_n$ se actualizan según algún criterio inteligente. Este método se ha aplicado a la minimización de nuestra función objetivo con resultados muy satisfactorios. El método es capaz de encontrar el mínimo de la función de modo independiente del punto de partida.

Actualización de los pesos: Estos pesos w_i se actualizan por medio de una función que se acerca a cero cuando el argumento tiende a infinito. Esta función debe cumplir varias condiciones: su dominio imagen tiene que estar en el intervalo real $[0, 1]$, su valor máximo se alcanza cuando el error cuadrático es cero y su valor mínimo cuando el error tiende a infinito (función decreciente). Existen múltiples funciones que cumplen estos requisitos. Por ejemplo, podríamos utilizar una función exponencial negativa $\exp(-c)$ o una función de la forma $\frac{1}{1+c^2}$. En esta implementación del algoritmo, hemos optado por una función gaussiana en la cual el parámetro σ nos permite modificar el intervalo de encajes correctos: cuanto más alto es el valor de σ , más alta es la probabilidad de ser un encaje correcto. La función gaussiana debe normalizarse para ajustar el dominio al intervalo $[0, 1]$. Así la expresión de la función que actualiza los pesos w_i es:

$$w_i = \exp -(\|x_i - \Phi w_i\|^2 / \sigma^2) \quad (5.8)$$

El algoritmo utilizado es el siguiente:

ALGORITMO:

ENTRADAS: Partimos de un conjunto de correspondencias iniciales $C = \{(\vec{a}_i, \vec{b}_i) : i = 0..N\}$

PRINCIPIO: $w_i = 1.0 : i = 0..N$

HACER:

1 PASO: Encontrar la transformación Φ que:

$$\Phi = \arg_{\varphi} \min \sum w_i \|(\vec{a}_i) - \varphi(\vec{b}_i)\|^2 \quad \forall \varphi$$

2 PASO: Actualizar pesos w_i

$$w_i = \exp -(\|x_i - \phi w_i\|^2 / \sigma^2)$$

MIENTRAS $\delta \leq w_i \leq \delta$

FIN.

El parámetro δ que aparece en el procedimiento anterior es un valor umbral que hace que el algoritmo acabe cuando todos los w_i tomen valores en el intervalo $[1 - \delta, 1]$ o bien en $[0, +\delta]$, lo cual indica cuales son las correspondencias correctas. Este algoritmo proporciona como resultado un subconjunto de C , conjunto de correspondencias iniciales, al que llamaremos C_c definido como $C_c = \{(\vec{a}_i, \vec{b}_i) \in C | w_i > (1 - \delta)\}$. Además, se obtienen los parámetros de la transformación geométrica.

5.5 Resultados experimentales

El método propuesto se ha aplicado a una conjunto de secuencias de imágenes sintéticas. La utilización de imágenes sintéticas en lugar de utilizar directamente imágenes angiográficas se debe a que la transformación geométrica es bastante compleja por lo necesitamos diseñar un experimento controlado que nos permita detectar posibles errores de implementación de los algoritmos propuestos.

Estas imágenes nos van a permitir detectar posibles errores en la implementación de los algoritmos. Estas secuencias corresponden a fotografías de un modelo de la retina humana hecho con papel-maché, utilizando un balón de plástico como molde. Los vasos sanguíneos se dibujaron sobre la cara interior del modelo siguiendo un patrón real. Algunas de estas imágenes se muestran en la figura 5.2. Cada una de estas secuencias representa el movimiento de la escena cuando ésta sufre rotaciones incrementales alrededor de un eje que yace sobre el plano ortogonal a la línea que une el centro del modelo esférico y el centro del sistema óptico. Se midieron los parámetros desconocidos que aparecen en la expresión de la transformación geométrica de forma que es posible estudiar las condiciones que hacen que el algoritmo funcione correctamente.

Para visualizar la complejidad de la función objetivo para el caso de la transformación no lineal, se muestran varias gráficas 5.3 de la función objetivo fijando todos los

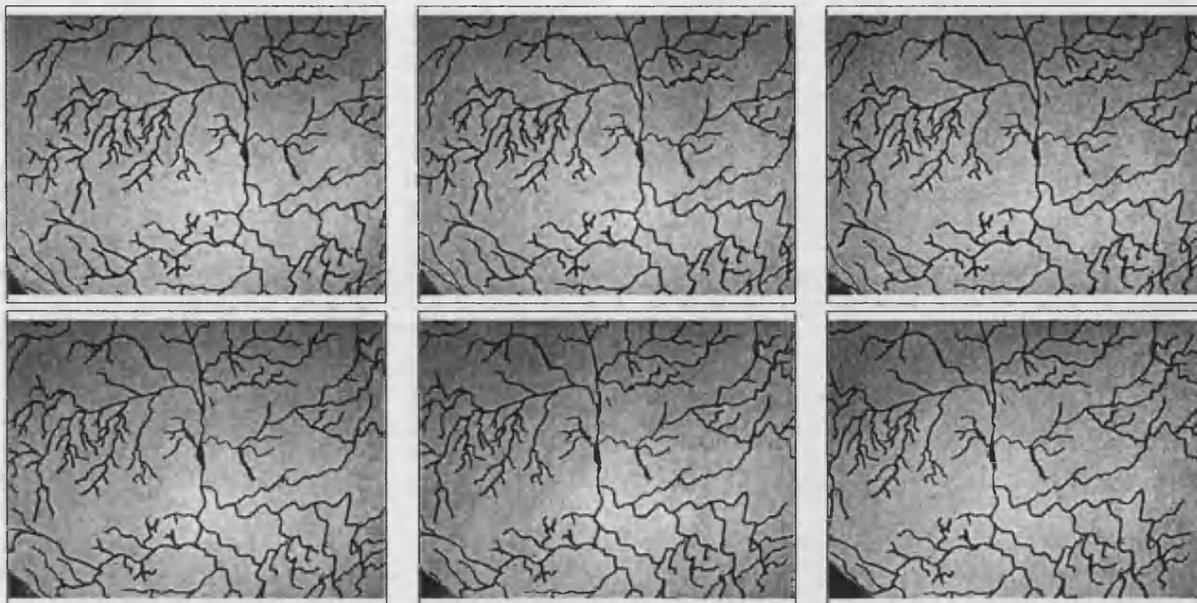


Figura 5.2: Varias imágenes de la secuencia sintética

parámetros excepto uno, parámetro en función del cual representamos la función coste. Esta representación muestra como los parámetros f_0 , z_r y R_0 afectan menos a la función coste que los parámetros relacionados con el movimiento de rotación, es decir, θ y α . Las componentes del gradiente correspondientes a estos parámetros son mucho mayores que las del resto lo que hace que el método del gradiente conjugado no minimice en las otras direcciones.

El procedimiento de alineamiento ha sido aplicado a una serie completa de imágenes sintéticas de las que se conocen todos los parámetros de la transformación puesto que han sido medidos experimentalmente. Este algoritmo debe ser capaz de, a partir de un conjunto de correspondencias, seleccionar aquellas que son correctas a la vez que se estima la transformación. El método, como se puede observar en la tabla , permite extraer los encajes correctos incluso cuando existe un 75% de encajes iniciales erróneos y la rotación entre imágenes es de 54° . Este resultado es de gran importancia ya que el ojo humano siempre gira ángulos inferiores a 40° lo que nos asegura el funcionamiento correcto del algoritmo en todos los casos.

Los resultados se muestran en la figura 5.4 donde aparecen en la parte superior pares de imágenes que han sido encajados mediante un procedimiento de relajación probabilística, y en la parte inferior el resultado cuando se utiliza un filtrado posterior.

	θ experimental	θ estimado
$I_1 - I_2$	5.0	5.1
$I_1 - I_3$	10.0	10.2
$I_1 - I_4$	15.0	15.6
$I_1 - I_5$	20.0	19.9
$I_1 - I_6$	25.0	25.8
$I_1 - I_7$	30.0	31.85
$I_1 - I_8$	0.0	-1.83
$I_1 - I_9$	-10.0	-11.7
$I_1 - I_{10}$	-20.0	-19.2
$I_1 - I_{12}$	-30.0	-35.0
$I_1 - I_{13}$	-40.0	-41.3

Tabla 5.1: Angulos medidos y estimados para varios pares de imágenes

Par de imagenes	ne	nc	R1	nf	R2
$I_1 - I_2$	48	36	0.64	36	1.0
$I_1 - I_3$	36	28	0.66	28	1.0
$I_1 - I_4$	35	29	0.74	29	1.0
$I_1 - I_5$	34	16	0.47	13	0.81
$I_1 - I_6$	26	16	0.61	6	0.4
$I_1 - I_7$	32	19	0.59	18	0.95
$I_1 - I_9$	24	16	0.66	12	0.75

Tabla 5.2: Esta tabla muestra el número de encajes obtenidos con el algoritmo de preprocesamiento y el número de encajes correctos. R1 y R2 son respectivamente $R1 = \frac{nc}{ne}$ y $R2 = \frac{nf}{nc}$ donde ne es el número de encajes iniciales, nc es el número de encajes correctos del total y nf es el número de encajes correctos obtenidos por el procedimiento propuesto

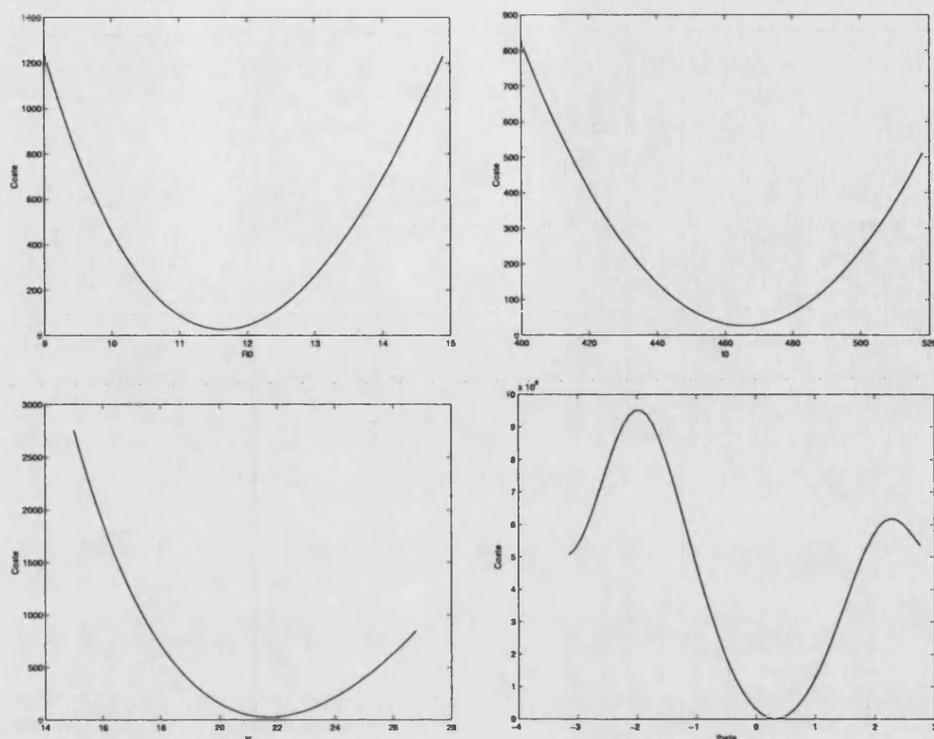


Figura 5.3: Representación de la función objetivo variando uno de los parámetros en cada caso. El resto de parámetros se han fijado a los valores donde se obtiene el mínimo.

La mayoría de los encajes en la parte superior no son correctos. Por el contrario, la parte inferior muestra un funcionamiento satisfactorio del método propuesto para varios pares de imágenes.

La figura 5.5 muestra el resultado de aplicar el procedimiento propuesto a un par de angiogramas reales. En la parte derecha aparecen los encajes obtenidos mediante un procedimiento de encaje de grafos que no tiene en cuenta la coherencia con una transformación geométrica. Por el contrario, en la parte izquierda solamente aparecen aquellos encajes que son compatibles con el modelo de transformación asumido. Como se puede observar en estas imágenes todos los encajes establecidos incorrectamente son eliminados por el algoritmo propuesto. El error cuadrático medio obtenido en este encaje es de 2.24 unidades. Los ángulos de giro en torno al eje X e Y respectivamente son de 0.1° y -0.1° . El desarrollo en serie de Taylor de la transformación geométrica en este caso particular aparece en la tabla 5.3 en comparación con la transformación afín obtenida

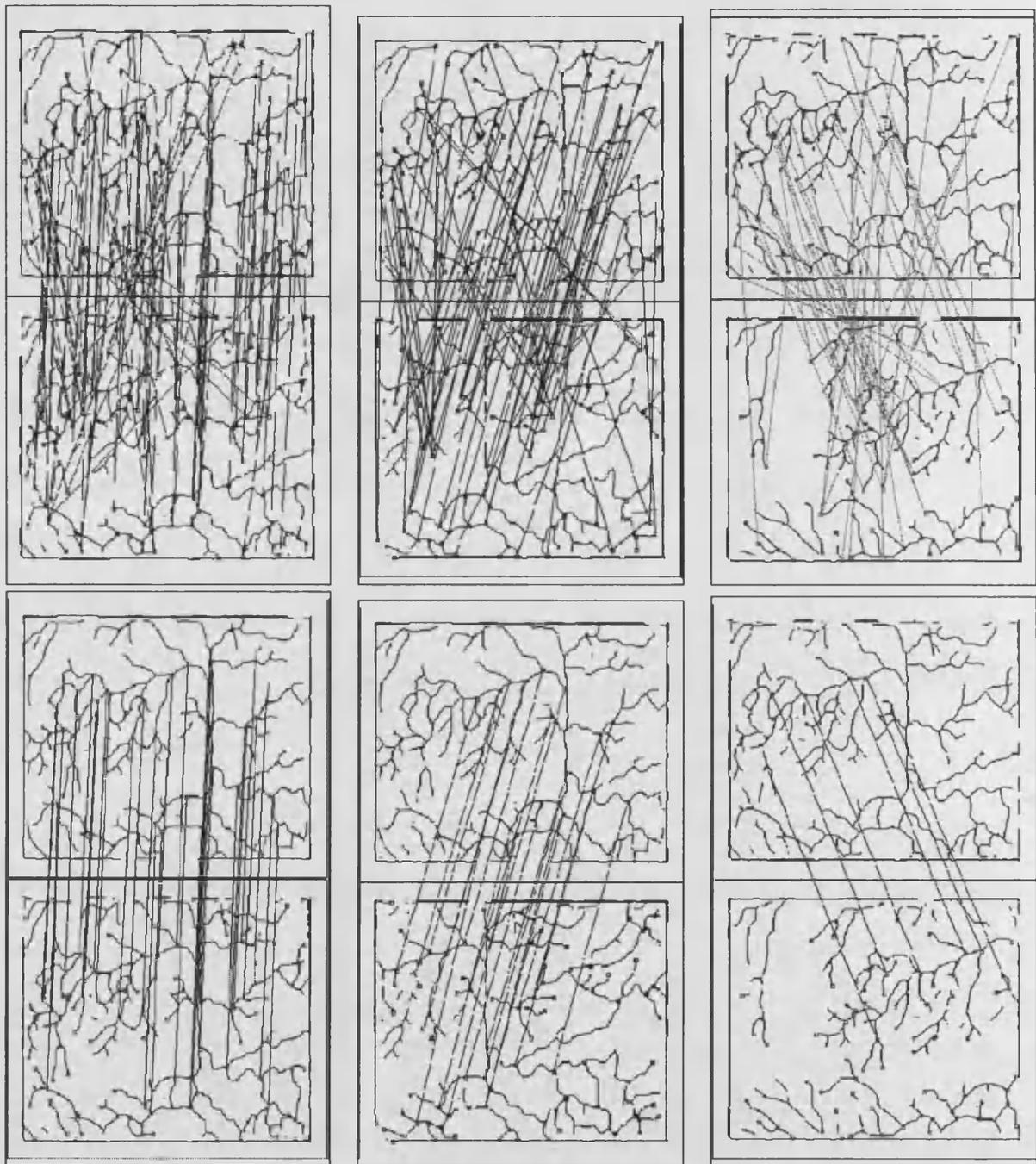


Figura 5.4: Correspondencias entre diferentes pares de imágenes. Primera fila: Correspondencias obtenidas mediante un algoritmo de relajación probabilística. Última fila: correspondencias utilizando información de la transformación. Izquierda: imágenes 1 y 2 de la secuencia. Centro: imágenes 1 y 7. Derecha: imágenes 1 y 13

	$f_x(x, y)$	$f_y(x, y)$
Afín	$x - 0.008y - 14.636$	$1.002y - 24.070$
Taylor	$x + 0.000002y - 16.319$	$y - 21.34$

Tabla 5.3: Comparación de la transformación afín y el desarrollo en serie de Taylor de la transformación propuesta para el angiograma de la figura 5.5

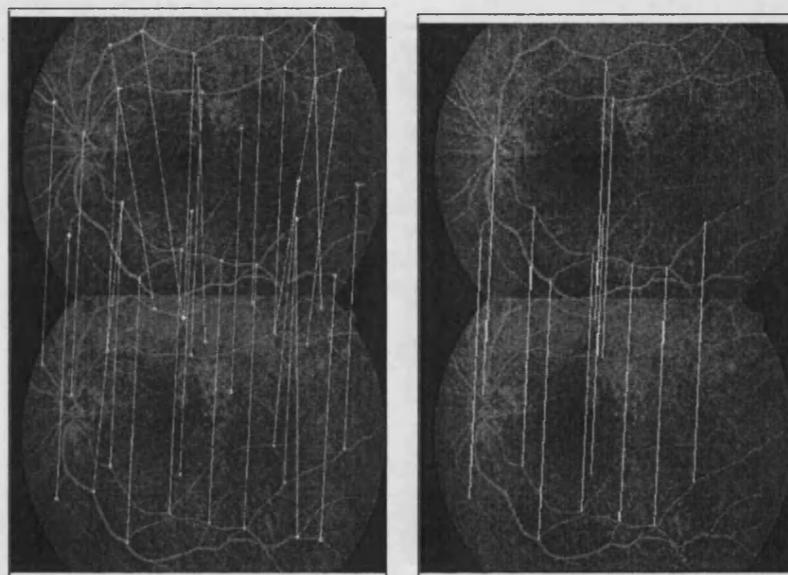


Figura 5.5: Correspondencias establecidas entre un par de angiografía

mediante el mismo método. De estos resultados observamos que no aparecen términos de grado superior al lineal por lo que una transformación afín es suficientemente adecuada para el alineamiento de estas imágenes. En el análisis de otros angiogramas se han obtenido resultados similares, es decir, los términos de grado superior al lineal o bien no aparecían o eran mínimos.

En la tabla 5.1 aparecen los ángulos medidos experimentalmente y los obtenidos mediante el proceso de optimización para las imágenes del experimento sintético donde la función coste es el error cuadrático medio ponderado. Los ángulos estimados son muy cercanos a los medidos. En el peor de los casos la diferencia entre el ángulo medido y estimado es de 5 grados que coincide con el error asociado al instrumento de medida utilizado en el que la división más pequeña era de 5° .

En la tabla 5.2 podemos ver el número de encajes (ne), el número de encajes co-

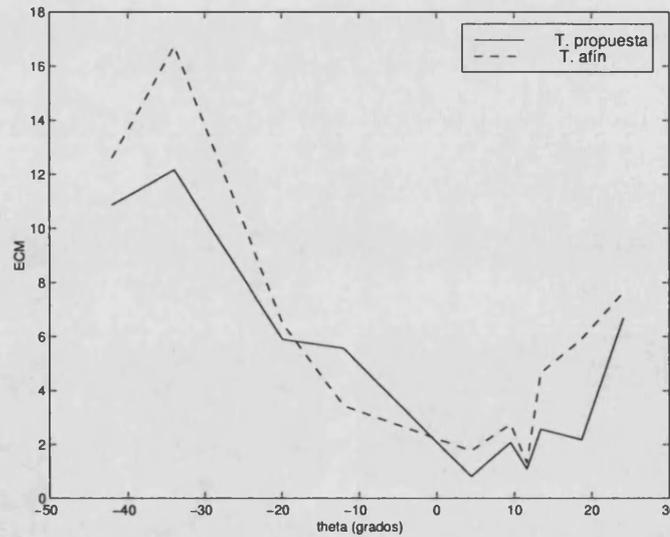


Figura 5.6: Comparación del ECM de una transformación afín y del modelo propuesto

rectos (nc), la fracción de encajes correctos ($R1 = \frac{nc}{ne}$), el número de correspondencias correctas extraídas del total (nf) y la fracción de encajes correctos recuperados del total de encajes correctos ($R2 = \frac{nf}{nc}$).

Puesto que lo que pretendemos es encontrar una transformación entre imágenes que reduzca el error cuadrático, entendido como la diferencia entre un punto y su correspondiente transformado mediante $\|x_i - \Phi w_i\|^2$, hemos realizado una comparación del error cuadrático medio (ECM) obtenido aplicando una transformación afín y la nueva transformación. Hemos partido de un conjunto de correspondencias correctas marcadas manualmente. Con estas correspondencias fijas se obtiene por un lado la transformación afín mediante un ajuste de mínimos cuadrados y por otro la transformación no lineal, utilizando para ello el método de Powell para optimización descrito en la sección anterior (5.4). La gráfica 5.6 muestra los ECM obtenidos para la transformación afín en función del ángulo de giro. Se espera que la transformación afín se comporte peor para valores altos de los ángulos de rotación. El ECM es menor para nuestra transformación que para la afín en todos los ángulos excepto cuando toma el valor de -11 grados. Esto muestra que la transformación propuesta da cuenta de las distorsiones que aparecen en la imagen como consecuencia de que la superficie es curva mucho mejor que lo hace una transformación afín.

Se ha realizado un experimento de las mismas características con imágenes angiográficas reales. El ECM obtenido con secuencias reales no es tan optimista. El error cuadrático medio calculado en el mejor de los casos es $ECM_{\min} = 0.5$ y el peor es $ECM_{\max} = 8.2$ siempre superior a los obtenidos con una transformación afín. No hay que olvidar que en el primer experimento analizábamos imágenes sintéticas en las que el objeto fotografiado era una superficie completamente esférica, y el ojo no lo es. Además, el sistema óptico de una cámara de fondo de ojo es mucho más complejo que el utilizado en este experimento.

5.6 Discusión

En este trabajo, hemos propuesto un algoritmo de encaje que estima los parámetros de una transformación geométrica no lineal así como los encajes correctos de entre un conjunto de correspondencias iniciales obtenidas mediante el proceso de relajación probabilística para el encaje de grafos propuesto por Wilson [WH96] entre un par de imágenes de una secuencia.

El funcionamiento de este algoritmo ha sido probado en una serie de secuencias sintéticas. Los resultados conseguidos son claramente satisfactorios puesto que el método es capaz de encontrar las correspondencias correctas incluso con más del 50 de encajes erróneos. Por otro lado, el ECM que se comete con la transformación propuesta es menor que con una transformación afín. Sin embargo, los experimentos con imágenes reales apuntan a que una transformación afín es suficiente para el encaje de las imágenes puesto que el ECM no se mejora considerablemente con la aplicación de una transformación más compleja. Una transformación que parece razonable según los resultados obtenidos sería una transformación polinómica que disminuiría ligeramente el ECM obtenido y ofrecería como ventaja su fácil cálculo pues permitiría que la obtención del mínimo de la función coste se abordara mediante el método de mínimos cuadrados.

Capítulo 6

Segmentación de angiografías fluoresceínicas

6.1 Introducción

Un paso previo y esencial en el análisis de angiografías es la segmentación de la imagen en sus componentes más relevantes. Estos componentes de interés son los vasos sanguíneos (arterias y venas), la fovea y el fondo. La segmentación es importante por muchas razones. En primer lugar, el nivel de fluoresceína no es el mismo en diferentes instantes de tiempo en estas partes, por ejemplo, el incremento de intensidad debido al proceso fluoresceínico aparece más tarde en las venas que en las arterias. En segundo lugar, el tipo y la seriedad de la lesión depende de su localización; en nuestro caso particular, cuanto más cerca de la fovea se encuentra una lesión más grave se considera. Y en último lugar, algunas manifestaciones de las enfermedades de la retina se descubren por la detección de un patrón anormal de los vasos sanguíneos (demasiado delgados, demasiado gruesos o vasos tortuosos).

Para obtener la segmentación hemos utilizado imágenes tomadas cuando la fluoresceína aparece en la circulación coroidal y la difusión no ha aparecido todavía (de 8 a 20 segundos después de la inyección). La figura 6.1 muestra dos angiografías fluoresceínicas. La segmentación se realiza sobre este tipo de imágenes para evitar que aparezcan otras estructuras extrañas debidas al efecto de la difusión. La primera de

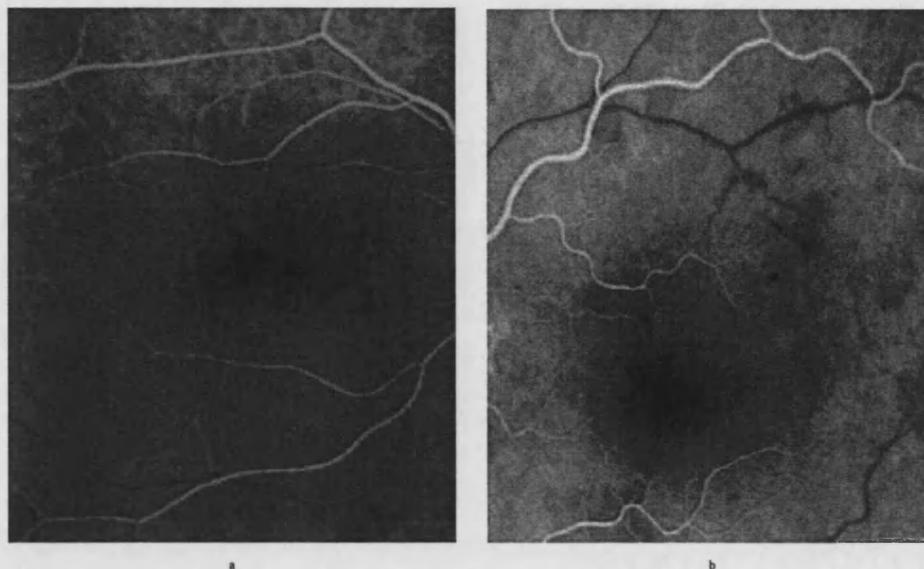


Figura 6.1: Dos imágenes angiográficas de fondo de ojo. (a) Paciente con retinopatía diabética. (b) Paciente con oclusión venosa

ellas fue tomada de un paciente con retinopatía diabética y la segunda de un paciente con oclusión venosa.

Aunque el problema de la segmentación ha sido ampliamente tratado en la literatura de la Visión por Computador sus aplicaciones al campo de la Oftalmología son escasas. M.H. Goldbaum y colaboradores [MNSM89] propusieron el diseño de filtros de encaje específicos para la segmentación de las imágenes de fondo ocular. Estos filtros son diferentes para cada tipo de elemento relevante en la retina. Un filtro gaussiano invertido rotado sobre ángulos discretos se utiliza para segmentar e identificar segmentos lineales de vasos sanguíneos. De la misma forma, una máscara de convolución gaussiana circularmente simétrica se utiliza para detectar el nervio óptico. La fovea se segmenta mediante filtros aplicados solamente en ciertas posiciones adecuadas con respecto a la localización del disco óptico. Existe una importante desventaja en esta técnica: el proceso de segmentación se divide en varios pasos de preprocesado, uno para cada elemento simple en lugar de utilizar un único procedimiento para la segmentación de la imagen total.

Respecto a la segmentación de vasos sanguíneos en imágenes médicas existen otros enfoques diferentes. Liu y Sun [LS93] sugieren un algoritmo de segmentación de las ar-

terias. Está basado en un procedimiento de seguimiento bajo un esquema de Detección-Eliminación. Kutka and Stier [KS96] proponen un procedimiento para la extracción de vasos sanguíneos que utiliza el filtro de Canny para una detección inicial de bordes. Una vez obtenidos estos bordes se sugieren ciertos parámetros asociados a cada uno de los puntos de la imagen de bordes como dirección y amplitud. Estos parámetros permiten hacer un seguimiento de las venas basándose en la similitud de cada uno de los puntos pertenecientes a la imagen de bordes con sus vecinos. Los puntos con parámetros suficientemente parecidos se conectarán formando venas. Como estamos interesados en separar venas y arterias, esta última técnica no es apropiada, puesto que no proporciona tal distinción. S. Chaudhuri ([CC89]), propone un algoritmo de detección de vasos sanguíneos en imágenes anéritas donde todos los vasos se visualizan más oscuros que el fondo. Este método asume un perfil de niveles de gris en los vasos aproximado por una gaussiana. Se basa en la construcción de diferentes filtros que sirven para la búsqueda de segmentos de venas en todas las direcciones posibles. Tolia ([TP98]) propone un algoritmo para la detección de venas basado en el estudio de los perfiles perpendiculares a la dirección del borde. Los puntos de este perfil se clasifican como pertenecientes a vasos o no vasos mediante el algoritmo de agrupamiento difuso de las C-medias. Este algoritmo se va aplicando sucesivamente a una nueva sección donde el punto medio se calcula mediante un desplazamiento en la dirección perpendicular al perfil actual. Este proceso permite realizar un seguimiento de todos los vasos sanguíneos de la retina siendo necesario únicamente un punto de inicio. Zana ([ZJ97]) presenta un algoritmo basado en morfología matemática para el reconocimiento de los vasos sanguíneos en angiografías. Se define un modelo geométrico de los patrones no deseados para separar los vasos de su entorno. Se utilizan operadores morfológicos para la detección de estructuras lineales en la imagen. De estas estructuras lineales se extraerán los vasos sanguíneos mediante cálculo de las curvaturas.

La detección de la fovea es de gran interés debido a la importancia, como ya hemos mencionado, de la posición relativa a la fovea de ciertas lesiones. Zana y colaboradores ([ZMJ98]) proponen un algoritmo de crecimiento de regiones utilizando morfología matemática para la detección de la fovea. El algoritmo necesita un punto inicial interior a la fovea para su inicialización. La estimación de su contorno se calcula por medio de un criterio que tiene en cuenta el nivel de intensidad que provoca la parada del algoritmo

cuando se encuentran puntos que posiblemente pertenecen a la zona vascular cercana a la fovea. Ibáñez y Simó [IS99] han aplicado métodos bayesianos para la detección de la fovea. En este trabajo, se plantea un campo aleatorio de Markov unidimensional en un sistema de coordenadas polares que les permite describir el contorno de la fovea. El resultado es un contorno bien definido que encierra la zona que se pretendía segmentar.

Por nuestra parte, hemos intentado en este trabajo, como aproximación alternativa a las ya indicadas, un procedimiento de segmentación basado en técnicas bayesianas.

Los métodos bayesianos son útiles en el análisis de imágenes médicas porque en ellas siempre está disponible un conocimiento previo sobre la estructura espacial de la escena que será segmentada. La anatomía humana y el uso de protocolos bien definidos para la adquisición garantiza una estructura de imagen común para cada tipo particular de examen. Los métodos bayesianos proporcionan una manera de utilizar este conocimiento en el diseño del proceso de segmentación.

En este marco de trabajo, la tarea de segmentación se reformula como un problema de estimación de parámetros considerando la imagen observada y como una versión con ruido de la imagen segmentada x . Es decir, la imagen segmentada es considerada un parámetro de la distribución de y y nuestro objetivo es la estimación de x . El conocimiento previo de x se incorpora por medio de la distribución *a priori*. Los modelos naturales para la distribución de probabilidad a priori de x son los campos aleatorios de Markov puesto que pueden modelizar información contextual. La información contextual juega un importantísimo papel en los problemas de segmentación porque la etiqueta verdadera de un punto debe ser compatible con las etiquetas de los puntos de la vecindad. El contexto representa nuestras hipótesis previas *a priori* sobre el mundo físico; en concreto, en el tipo de imágenes que analizamos esperamos que sean continuas y suaves.

Los métodos bayesianos han sido ampliamente utilizados en algunos campos del análisis de imágenes médicas tales como la reconstrucción de datos de tomografía computerizada de la emisión de fotones ([GM87, AG91, Gre90, WG94, Wei97] son ejemplos), la detección de contornos ventriculares ([Tel92]), la segmentación de tomografías y la detección de microcalcificaciones en mamografías ([Kar92]).

Este capítulo se organiza de la siguiente forma: la sección 6.2 describe las diferentes aproximaciones bayesianas en la segmentación de imágenes, en la sección 6.3 se propone el modelo estocástico, la sección 6.4 presenta el modelo a priori de la imagen segmentada, la sección 6.5 describe la forma de llevar a cabo la estimación de la imagen segmentada, la sección 6.6 trata de la estimación de los parámetros y, finalmente, se muestran algunos resultados obtenidos aplicando esta metodología en la sección 6.7.

6.2 Aproximaciones bayesianas

Esta sección presenta una breve revisión de las diferentes aproximaciones bayesianas en la segmentación de imágenes. Para un estudio detallado de este tema pueden consultarse [Bes86, GG84, Cre93, Guy95, Win96].

Sea $y = (y(s))_{s \in S}$ la imagen observada, siendo s la posición de cada punto y S el conjunto de todos ellos, y sea $x = (x(s))_{s \in S}$ la imagen etiquetada donde $x(s)$ denota el verdadero segmento o región (no conocido) al que pertenece el punto s .

Los valores observados $y(s)$ (niveles de gris en nuestras imágenes) se interpretan como una realización de una variable aleatoria cuyos parámetros son una función de $x(s)$. La distribución de probabilidad de la imagen observada y tiene densidad $f(y/x)$. Esta densidad describe la formación de la imagen, incluyendo tanto al ruido estocástico inherente a la observación de y , como a la influencia determinista de x en y .

Dada una distribución *a priori* sobre x con función de densidad π , la distribución *a posteriori* de x dado y será

$$P(x/y) \propto f(y/x)\pi(x), \quad (6.1)$$

y la segmentación x se obtiene habitualmente eligiendo \hat{x} como la moda de la distribución *a posteriori*. Este estimador se conoce como MAP. Es importante notar que la imagen x se estima como un todo. En particular, requerimientos contextuales incorporados en la distribución *a priori* (como la homogeneidad de las regiones) son inherentes en la distribución *a posteriori*, y así, influirán en el estimador \hat{x} . De esta manera, algunos puntos pueden ser mal clasificados debido a la aceptación del criterio de homogeneidad. Este es el estimador utilizado en nuestros experimentos.

Si no estuviéramos interesados en estructuras regulares sino solo en obtener pequeñas tasas de error, otro estimador razonable podría ser la moda marginal *a posteriori* (MP-ME) Por último en otras aplicaciones tales como la reconstrucción de tomografías, el valor medio o la esperanza de la distribución *a posteriori* es un estimador más adecuado (MMSE).

Como se ha dicho anteriormente, utilizaremos un modelo *a priori* para la verdadera imagen etiquetada x el cual expresa el conocimiento *a priori* que tenemos sobre el mundo físico, en nuestro caso, la interacción espacial entre puntos y la anatomía humana. Los campos aleatorios de Markov (a partir de ahora CAM) han mostrado ser herramientas apropiadas para modelizar información espacial (vease, Besag [Bes75] y Cressie [Cre93]) y por lo tanto son modelos naturales para la distribución *a priori* de x . Para definir un CAM es necesario introducir un sistema de vecindades en S , el conjunto de puntos. La noción de vecindad se establece, bien mediante una expresión analítica que involucra las coordenadas de los *pixels*, bien mediante una regla correctamente definida cualquiera. Un subconjunto de S se llama *clique* si cualquier par de elementos de c son vecinos. La introducción de las nociones de vecindad y *clique* en imágenes es importante porque permite que el conocimiento de información exclusivamente local (relativa a los puntos de un *clique* dado) pueda combinarse para dar información global (relativa a toda la imagen). Ello es posible en virtud del teorema de Hammersley-Clifford. Este teorema (ver Cressie [Cre93]) muestra que la distribución de probabilidad correspondiente a un CAM puede expresarse como

$$\pi(x) = \frac{1}{Z} \exp(Q(x)) = \frac{1}{Z} \exp\left(-\sum_{c \in C} V_c(x_c)\right), \quad (6.2)$$

donde $Q(x)$ es la función de potencial, c es un *clique*, $x_c = (x(s) : s \in c)$, y el sumatorio se extiende al conjunto de todos los *cliques* C . Los V_c se llaman potenciales; los potenciales son funciones (arbitrariamente establecidas) de manera que tomen valores bajos para aquellas configuraciones (asignaciones de etiquetas a cada punto del *clique*) más probables, y viceversa. Una revisión detallada de los CAM y la demostración del teorema de Hammersley-Clifford se encuentran en el apéndice A de la presente tesis.

Como hemos indicado, los estimadores bayesianos de x se basan en las distribuciones *a posteriori*. Aparte de algunos casos particulares, tal distribución no puede ser conocida completamente. Por esta razón, para obtener estos estimadores, se ha adoptado la

metodología MCMC (*Markov Chain Monte Carlo*) [Gre96]. En esta metodología podemos simular muestras de la distribución *a posteriori* utilizando cadenas de Markov cuya distribución estacionaria sea $P(x/y)$. Los métodos ampliamente aceptados utilizados para este propósito son los muestreadores de Gibbs y el método dinámico de Metropolis. Si estamos interesados en el estimador MAP, podemos utilizar el algoritmo conocido como *Simulated Annealing* [GG84] (SA) con actualizaciones de Gibbs o de Metropolis.

Otra opción para estimar x es el uso del algoritmo de estimación *Iterated Conditional Modes* o ICM [Bes86]. El estimador ICM es un máximo local de $P(x/y)$. Aunque la convergencia de este algoritmo es más rápida que el algoritmo SA, el ICM podría caer en un máximo local sin posibilidades de escapar de él, mientras que el SA se comporta mejor en estas situaciones saltando a una solución elegida aleatoriamente lo que evita la caída en extremos locales.

6.3 Un modelo estocástico

Para aplicar los métodos estocásticos del modelo descrito en la sección anterior deberemos asumir un modelo estocástico para las intensidades observadas dadas las etiquetas. Hemos considerado un modelo simple con un proceso de ruido Gaussiano estocásticamente independiente entre puntos.

Veamos previamente la notación utilizada. Sea S el conjunto de puntos de la imagen y $n = |S|$, el número de ellos. En cada punto o posición $s \in S$ denotamos:

- $y(s)$ las intensidades observadas en el punto $s \in S$.
- $X(s) \in \{0, 1, 2, 3\}$ los etiquetas de cada tipo de región, convencionalmente: 0 para el fondo, 1 para arterias, 2 para las venas y 3 para la fóvea.

Asumiremos que las intensidades observadas dadas x siguen el siguiente modelo:

$$y(s) = \alpha(s) + \epsilon(s). \quad (6.3)$$

donde $\{\alpha(s)\}_{s \in S}$ son constantes para cada tipo de región, es decir, $\alpha(s) = \alpha_{x(s)}$ y $\epsilon(s)$ son variables normales (condicionalmente) independientes con media nula y varianza

constante para cada región, $\sigma_{x(s)}^2$ (ruido Gaussiano). La verosimilitud, bajo las hipótesis anteriores, tiene la siguiente expresión:

$$f(y/\alpha) = \prod_{s \in S} f(y/\alpha(s)) = \prod_{s \in S} f(y(s)/x(s)) = \prod_{s \in S} \frac{1}{\sqrt{2\pi\sigma_{x(s)}^2}} \exp\left\{-\frac{1}{\sigma_{x(s)}^2} (y(s) - \alpha_{x(s)})^2\right\}$$

En esta función de verosimilitud la única información que se maneja se basa en la intensidad del punto, que hemos comprobado que no es suficiente para la segmentación de nuestras imágenes, puesto que es muy sensible a las condiciones de iluminación en el proceso de captura. Principalmente, los vasos sanguíneos son los que más se ven afectados, debido a sus pequeñas anchuras. Para resolver este problema podemos incorporar más conocimiento sobre los tipos de objetos que se encuentran en las angiografías fluoresceínicas. Así, se sabe que los vasos sanguíneos normalmente aparecen como estructuras curvilíneas que puede ser modeladas localmente como líneas.

La aproximación más obvia para modelar esta diferencia de forma sería a través del uso de una distribución *a priori* más compleja que la que se considera en la siguiente sección. Eso conduciría a considerar una vecindad de un orden mayor, lo que, desde el punto de vista de la computación, sería bastante ineficiente.

En su lugar, nuestra aproximación ha consistido en extraer de la imagen original un parámetro de linealidad asociado a cada punto e incorporar esta información en la función de verosimilitud.

En la literatura se pueden encontrar varios algoritmos para conseguir una buena medida de linealidad en cada punto, entendida como una medida de la certeza de que una línea recta cruza un punto dado. La mayoría de ellos se basan en el módulo del vector gradiente que indica si un punto pertenece a un borde (altos valores del resultado) o no, y, en caso positivo, la dirección angular de la línea asociada con él. En un trabajo reciente, Díaz, Domingo y Ayala [DDA97] proponen un método para detectar esquinas basado en un análisis estadístico de las orientaciones del gradiente en una ventana circular del punto considerado como un posible candidato a ser esquina. Una de las ideas usadas en este trabajo puede servirnos para caracterizar cuantitativamente la linealidad en un punto. Concretamente, un punto del borde s en la imagen puede

definirse como un punto cuyo gradiente no es nulo y las orientaciones del gradiente observadas para otros puntos de su vecindad son suficientemente similares. De hecho, estas se agrupan en torno a una orientación privilegiada μ . El ruido provoca que los ángulos observados varíen alrededor de este valor μ desconocido. Se asume que las orientaciones θ del gradiente siguen una distribución de *von Mises* ([Mar72]) que juega el mismo papel en la estadística circular que la distribución gaussiana en la estadística no circular.

Hemos aprovechado estas ideas para obtener el parámetro de linealidad requerido. Así, necesitamos calcular el gradiente en cada punto individual y umbralizarlo con un valor dado E . Se estudian todos los puntos que tienen un valor del módulo del gradiente mayor que E mediante el cálculo del histograma de orientaciones. De este modo, para cada punto s en la imagen analizada, los parámetros (μ, κ) de la distribución de von Mises para las orientaciones en su vecindad son estimados y se prueba la hipótesis nula: "La distribución de orientaciones en una vecindad de s es una distribución de von Mises con parámetros (μ, κ) " utilizando un test de Watson-Stephens [Wat62]. Los p-valores obtenidos, $l(s)$, se utilizarán como un parámetro de linealidad puesto que denota la similitud de nuestros datos en el modelo asumido, la distribución de von Mises.

Admitiremos independencia condicional entre las intensidades y y el parámetro de linealidad l . Esta suposición debe ser validada porque, como hemos mencionado, el cálculo del parámetro de linealidad se basa en el gradiente de intensidad. Utilizando datos de prueba, la hipótesis de independencia se ha probado, encontrando que con un nivel de significación $\alpha = 0.01$, no puede ser rechazada. De esta forma, hemos probado la hipótesis de ruido gaussiano utilizando datos de entrenamiento. Los coeficientes de correlación resultantes y su significación se muestran en la tabla 6.1 que indican que nuestra hipótesis no puede ser rechazada.

Finalmente, la verosimilitud adopta la siguiente expresión:

$$f(y, l/\alpha) = \prod_{s \in S} \frac{1}{\sqrt{2\pi\sigma_{x(s)}^2}} \exp\left\{-\frac{1}{\sigma_{x(s)}^2}(y(s) - \alpha_{x(s)})^2\right\} f(l(s)/x(s)) \quad (6.4)$$

donde f denota la distribución de probabilidad del parámetro de linealidad dada una etiqueta. Esta distribución no es conocida y se estimará en la sección 6.6.2 utilizando la estimación de densidad.

Etiqueta	p-valores del test de normalidad	coeficiente de correlación	p-valor
Fondo	0.0827	0.0661	0.2978
Fóvea	0.0319	0.0697	0.0347
Venas	0.0907	0.1474	0.0293
Arterias	0.0339	-0.0673	0.2891

Tabla 6.1: p-valores correspondientes a los tests de normalidad y no correlación

6.4 Modelo *a priori* para las etiquetas

El proceso que siguen las etiquetas x se ha modelizado como un campo aleatorio de Markov (véase apéndice A). Conocemos previamente los tipos de objetos que se encuentran en el fondo de ojo; por lo tanto las características de estos objetos se pueden incorporar en el modelo.

Los vasos sanguíneos normalmente aparecen como estructuras delgadas y curvilíneas que se ramifican o cruzan. La intensidad de las arterias es más alta que la retina por efecto de la fluorescencia. La intensidad de las venas es casi siempre más baja que la retina adyacente. La fóvea (sobre $500 \mu\text{m}$ de diámetro) es una zona avascular y generalmente aparece más oscura que los puntos circundantes, y cerca del centro de la imagen.

Más concretamente queremos que el modelo *a priori* exprese tres condiciones:

1. Los puntos vecinos pertenecen probablemente a la misma región, es decir, nuestra imagen es suave.
2. Los puntos del centro de la imagen pertenecen con más probabilidad a la fóvea, y puntos en la periferia pertenecen con mayor probabilidad a vasos sanguíneos, es decir, la distribución no es homogénea espacialmente.
3. Los puntos que pertenecen a los vasos sanguíneos se organizan en estructuras grandes y delgadas, mientras que la fóvea es una región compacta y convexa.

Para modelar la interacción entre puntos de forma que expresen las condiciones antes vistas, hemos definido un sistema de vecindad como la unión de dos tipos básicos de vecindades:

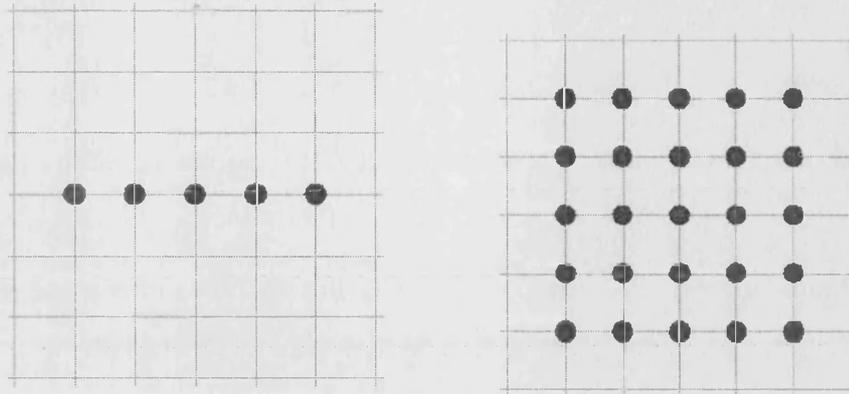


Figura 6.2: *Cliques* que contribuyen a la función negpotencial

- Vecindad de segundo orden: los vecinos de un punto s son aquellos puntos a una distancia menor o igual que 2.
- Vecindad de cuarto orden: los vecinos de un punto s son aquellos puntos cuyas coordenadas están a una distancia menor o igual que 4 de s .

Así utilizaremos el primer tipo de vecindad para expresar la condición (1) y el segundo para la condición (3), teniendo en cuenta que las venas y las arterias tienen normalmente una anchura menor de ocho *pixels*. La no homogeneidad se expresará por medio de un campo externo espacialmente dependiente.

Por motivos de simplicidad, solamente cuatro tipos diferentes de *cliques* contribuirán a la función de negpotencial, nombrados como: *cliques* con un único punto (que llamaremos *cliques* de tipo 1), *cliques* con dos vecinos de segundo orden (*cliques* de tipo 2), *cliques* con cinco vecinos de cuarto orden como los de la figura 6.2 (a) y sus rotaciones (*cliques* de tipo tres) y finalmente *cliques* con 25 vecinos de cuarto orden como los de la figura 6.2 (b) (*cliques* de tipo 4).

Como se ha comentado previamente, intentamos modelar la estructura espacial concreta de las imágenes de fondo de ojo mediante un campo externo dependiente de un modelo de CAM. Para simplificar, hemos considerado un campo externo espacialmente dependiente muy sencillo:

Hemos definido cuatro particiones de S_i , una partición para cada tipo de región:

(S_k, S'_k) $k = 0, 1, 2, 3$.

$S_0 = S$, porque el fondo se extiende por toda la imagen.

$S_1 = S_2$ son aquellos puntos que están a una distancia mayor que r_1 del centro de la imagen, porque venas y arterias se sitúan en la periferia.

y S_3 corresponde a aquellos puntos situados a una distancia menor que r_2 del centro de la imagen, porque la fovea se encuentra en el centro de la imagen.

El campo externo que hemos asumido, para un k fijo, es constante para cada miembro de la partición.

Todas estas consideraciones nos conducen a asumir la siguiente distribución *a priori* para la imagen segmentada x :

$$\pi(x) \propto \exp\{\sum_{s \in S} \sum_{k=0}^3 a_k^s 1_k(x(s)) + \quad (6.5)$$

$$\sum_{0 \leq k \leq l \leq 3} b_{kl} n_{kl} + \sum_l d_l n_l + e_l m_l\},$$

donde n_{kl} es el número de pares de puntos con una vecindad de segundo orden etiquetados como k y l respectivamente, n_l es el número de *cliques* de tipo 3 etiquetados como l , m_l es el número de *cliques* de tipo 4 etiquetados como l , y $a_k^s = a_k$ si $s \in S_k$ y a'_k en otro caso, siendo a_k y a'_k constantes, que fijaremos después.

Los parámetros b_{ij} están pensados para favorecer que los *cliques* de tipo 2 tengan la misma etiqueta de forma que la imagen segmentada sea suave. Los parámetros d_l intentan ayudar a los *pixels* que forman *cliques* de tipo 3 a tener la misma etiqueta y, por el contrario, los parámetros e_l se abren a que todos los puntos de los *cliques* de tipo 4 tengan etiquetas de arterias o venas. Por medio de estos parámetros obtendremos que las venas y las arterias sea largas y finas y la fovea sea compacta.

Tendremos que imponer restricciones en los parámetros para hacer que el modelo sea identificable. La restricción más común es considerar $V_c(x_c) = 0$ si $x(s) = 0$ para algún $s \in c$. Por lo tanto, hemos fijado $a_0 = d_0 = e_0 = 0$ y $b_{0l} = 0 \forall l$.

También hemos fijado $a_l = 0.1a'_l$ para $l = 1, 2$ y $a_3 = 0.01a'_3$.

Esta distribución condicional resulta en la siguiente posterior conjunta:

$$P(x/y, l) \propto \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left\{\sum_{s \in S} \left(\frac{1}{2\sigma^2} (y(s) - \alpha_{l(s)})^2\right)\right\} \quad (6.6)$$

$$\sum_{s \in S} \sum_{k=0}^3 a_k^s 1_k(x(s)) + \sum_{0 \leq k \leq l \leq 3} b_{kl} n_{kl} +$$

$$\sum_l d_l n_l + e_l m_l \} \prod_{s \in S} f(l(s)/x(s)).$$

La distribución *a posteriori* de x es también un Campo Aleatorio de Markov con respecto al mismo sistema de vecindad.

6.5 Estimación de x

Nuestro objetivo es encontrar la estimación de la imagen de etiquetas x . Para hacer inferencias sobre estos parámetros deberíamos trabajar con la distribución *a posteriori* $P(x/y, l)$. Esta distribución tiene un número adicional de parámetros desconocidos (hiperparámetros) necesarios para la definición total del modelo. Estos parámetros se pondrán fijos y conocidos. Hablaremos sobre ellos en la sección siguiente.

En nuestros experimentos, hemos utilizado dos métodos para la estimación de x , el algoritmo *Simulated Annealing* y el ICM. Estos métodos requieren solamente acceder a expresiones que sean proporcionales a la distribuciones condicionales *a posteriori*, dadas por:

$$P(x(s)/x(t) \ t \neq s, y, l) \propto \exp\{a_{x(s)}^s + \sum_{x(s) < k} b_{x(s)k} n_{s,k} + \quad (6.7)$$

$$\sum_{k < x(s)} b_{kx(s)} n_{s,k} + d_{x(s)} t_{s,x(s)} +$$

$$e_{x(s)} m_{s,x(s)} + \frac{1}{2\sigma^2} (y(s) - \alpha_{x(s)})^2\} f(l(s)/x(s))$$

donde $n_{s,k}$ es el número de vecinos de s de segundo orden con etiqueta k , $t_{s,k}$ es el número de *cliques* de tipo 3 en los cuales s está incluido teniendo todos los otros puntos etiquetados como k , $m_{s,k}$ es el número de *cliques* de tipo 4 en los que s está incluido con el resto de los puntos etiquetados como k .

El algoritmo que hemos utilizado es un algoritmo de relajación estocástica propuesto por Geman y Geman [GG84]. Se basa en *Simulated Annealing* y el muestreador de Gibbs. Se sabe que este algoritmo (Guyon [Guy95]) se adapta bien cuando la función que se minimiza es Markoviana.

El esquema de recorrido del *Gibbs Sampler* utilizado es un barrido fila a fila y el esquema de enfriamiento $T(k)$ es geométrico.

Este esquema es como sigue:

Paso 1: Inicialización. $k = 1$

Se introduce una estimación inicial de x . Se inicializa la temperatura $T(1)$.

Paso 2: Actualización completa. Ejecutar un barrido completo sobre toda la imagen x de tal forma que la nueva $x(s)$ sea elegida aleatoriamente con probabilidad $\exp\{\frac{1}{T(k)} \text{Ln}(P(x(s)/x(t) \ t \neq s, y, l))\}$.

Paso 3: Actualizar los parámetros de control. Se actualiza $k = k + 1, T(k)$.

Paso 4: Volver a 2 durante un número fijo de iteraciones o hasta la convergencia aproximada de x .

En la práctica, hemos obtenido resultados satisfactorios utilizando

$$\frac{1}{T(k)} = \frac{1}{1.01^k}, \text{ donde } k \text{ es el número de barridos completos.}$$

La convergencia del algoritmo *Simulted Annealing* es lenta. Como alternativa hemos probado un algoritmo de relajación determinista conocido como *Iterated Conditional Modes* propuesto por Besag [Bes86]. El esquema es el siguiente:

Paso 1: Iniciar. $k = 1$ Se introduce un estimador inicial para x .

Paso 2: Actualización completa. Se realiza un barrido completo sobre la imagen x de tal forma que la nueva $x(s)$ se elige para que maximice $P(x(s)/x(t) \ t \neq s, y, l)$.
Se actualiza $k=k+1$.

Paso 3: Se vuelve al paso 2 un número fijo de iteraciones o hasta una convergencia aproximada de x .

6.5.1 Estimadores iniciales

Para poder utilizar el procedimiento iterativo explicado previamente se necesitan algunos estimadores iniciales. El estimador inicial de x se ha obtenido utilizando un algoritmo de segmentación simple basado en umbrales locales.

El estimador inicial de las imágenes de la figura 6.1 se muestra en la figura 6.3

6.5.2 Convergencia

Vamos a seguir la convergencia mediante una gráfica del número de puntos para cada tipo de etiqueta frente al número de la iteración. Algunas de estas gráficas se muestran en la figura 6.8. En esta gráficas se observa como, a partir de un cierto número de iteraciones, el número de etiquetas de cualquier clase se estabiliza.

6.6 Parámetros e hiperparámetros

Como dijimos antes, el modelo tiene un conjunto numeroso de parámetros desconocidos. Estos parámetros corresponden a los parámetros de la distribución a priori (hiperparámetros), los parámetros de la distribución de intensidad y los parametros de linealidad. Si denotamos θ al vector de parámetros, existen varias formas de estimarlos en el paradigma Bayesiano en el que estamos trabajando:

- Aproximación empírica Bayesiana. En la forma empírica Bayesiana se considera que los parámetros son desconocidos, pero fijos.
- Aproximación Bayesiana completa. En esta aproximación los parámetros son muestras aleatorias extraídas de alguna distribución *a hiperpriori*. Este procedimiento se utiliza en Johnson [Joh94], Mollie [Mol96], Weir [Wei97] y Molina [Mol94]. Molina [Mol94] aconseja el uso de la aproximación Bayesiana cuando los modelos (verosimilitud y distribución *a priori*) no son una descripción muy correcta de los procesos. De esta forma la inferencia deseada sobre x debería basarse en la distribución *a posteriori* que surge de la integración sobre todas las variables en la conjunta *a posteriori*. La obtención de esta marginal es una tarea bastante difícil. Además, si

esto fuera posible nos quedaríamos todavía con una dimensionalidad alta, y, quizás, incluso una marginal más complicada.

En este trabajo, se ha adoptado una aproximación empírica Bayesiana y consideraremos que los parámetros son fijos y desconocidos y por lo tanto tienen que ser estimados.

Existen de nuevo dos formas alternativas: utilizar datos de entrenamiento si están disponibles, o no utilizarlos porque, o bien no están disponibles o su uso no es conveniente para nuestro tipo particular de problema. Como consecuencia de los diferentes tipos de parámetros utilizados en nuestro modelo tenemos que distinguir tres casos.

6.6.1 Parámetros en la distribución *a priori*

Los datos de entrenamiento pueden usarse para la estimación de parámetros en la distribución *a priori* porque todas nuestras imágenes son de fondo de ojo y por lo tanto comparten estructuras comunes. El método de máxima verosimilitud podría utilizarse pero la constante de normalización (distribución $\pi(x)$) no está disponible ni puede ser calculada. Se han propuesto algunos algoritmos para superar esta dificultad. Podemos utilizar, por ejemplo, el debido a Younes [You88], Geyer y Thompson [GT94], Geyer [Gey94] y la aproximación de *Monte Carlo Newton-Raphson* [Cre93] [Guy95].

Sin embargo, se ha utilizado el enfoque de pseudoverosimilitud aquí debido a su bajo coste computacional. La función de pseudoverosimilitud de la distribución *a priori* de x tiene la siguiente expresión:

$$PL(a_k, b_{kl}, d_l, e_l) = \frac{\exp\{\sum_{s \in S} (\sum_{k=0}^3 a_k^s 1_k(x(s))) + \sum_{0 < k < l < 3} 2b_{kl} n_{kl} + \sum_l 5d_l n_l + 25e_l m_l\}}{\prod_{s \in S} \sum_{x=0}^3 \exp\{a_x^s + \sum_{x < k} b_{xk} n_{s,k} + \sum_{k < x} b_{kx} n_{s,k} + d_x t_{s,x} + e_x m_{s,x}\}} \quad (6.8)$$

Para maximizar esta pseudoverosimilitud se ha utilizado el método del gradiente conjugado de Fletcher-Reeves-Polak-Ribiere (vease Press, Teulosky, Vetterling and Flannery [PTVB92]).

6.6.2 Distribución de los parámetros de linealidad

La función de densidad de la distribución de linealidad es completamente desconocida, por lo que tendremos que estimarla. Para obtener $\hat{f}(l(s)/x(s))$ se ha utilizado un *kernel* gaussiano a partir de datos de entrenamiento.

6.6.3 Parámetros de la distribución de intensidad

Los datos de entrenamiento no son útiles para estimar los parámetros de la distribución de intensidad porque existe una gran variabilidad en los niveles de gris entre imágenes.

Hemos elegido seguir un procedimiento iterativo (vease Besag [Bes86] y Qian y Titterington [QT93]). Dada una estimación actual de x , sea \hat{x} , se obtiene un nuevo valor para los parámetros de la distribución de y maximizando la densidad condicional de y dada por la ecuación 6.4:

$$\hat{\alpha}_k = \frac{1}{n_k} \sum_{s \in S: \hat{x}(s)=k} y(s)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum (y(s) - \hat{\alpha}_{\hat{x}(s)})^2.$$

donde $n_k = \#\{s \in S : \hat{x}(s) = k\}$. Con estos nuevos valores de los parámetros, se obtienen nuevos estimadores para x . Este proceso se repite hasta la convergencia.

6.7 Resultados

La metodología propuesta ha sido aplicada a varias imágenes de fondo de ojo. Todas estas imágenes han sido obtenidas con una cámara de fondo de ojo Canon CF-60U y Nikon NFC 50, transformadas en señal de video con una cámara de video Sony PHV-A7E y digitalizadas con una tarjeta Matrox MVP-AT con resolución de 512×480 puntos. Pertenecen a pacientes que sufren de diferentes lesiones tales como retinopatía diabética, oclusión venosa y edema macular. Dos imágenes de las utilizadas se muestran en la figura 6.1.

Como hemos explicado, los parámetros de la distribución *a priori* han sido estimados mediante un conjunto de datos de entrenamiento. Hemos considerado $d_1 = d_2$ y $e_1 = e_2$

$a_1 = -0.02$	$a_2 = 0.07$	$a_3 = 0.31$
$b_{11} = 1.28$	$b_{12} = -1.4$	$b_{13} = -0.25$
	$b_{22} = 1.7$	$b_{23} = -0.15$
		$b_{33} = 1.8$
$d_1 = 1.1$	$d_3 = 0.21$	
$e_1 = -0.98$	$e_3 = 0.1$	

Tabla 6.2: Parámetros estimados utilizando datos de entrenamiento.

por el comportamiento similar de las venas y las arterias. Los valores para los otros parámetros se muestran en la tabla 6.2.

Se han aplicado dos métodos diferentes de optimización: el *Simulated Annealing* y el ICM que han sido programados en VISTA ¹, conjunto de librerías de procesamiento de imágenes escritas en C de uso público.

Las imágenes obtenidas utilizando estos esquemas se muestran en las figuras 6.4 y 6.5, que corresponden respectivamente al SA después de 100 iteraciones y al ICM después de 50 iteraciones. En primer lugar, se observa que los resultados que corresponden a la primera imagen son peores debido a su pobre contraste.

Si comparamos los algoritmos SA y ICM podemos ver que los dos proporcionan buenos resultados globales con pequeñas diferencias en cuanto al tamaño de la fovea fundamentalmente.

Hemos encontrado que en los dos casos las imágenes obtenidas mediante el algoritmo SA dan una mejor estimación para las venas y las arterias. Por otro lado, la estimación de la fovea es ligeramente mayor y, en este caso, peor que la obtenida mediante el algoritmo ICM.

Las imagen binaria 6.6 muestra la diferencia entre los resultados obtenidos mediante el SA y el ICM. Se aprecia que la fovea se centra aproximadamente en la misma posición y que es claramente mayor en la imagen obtenida con el SA. Del mismo modo, en la imagen 6.7 se compara la segmentación que se obtiene con el método propuesto y una

¹Desarrollado principalmente por Art Pope en el departamento de Ciencias de la Computación, Universidad de British Columbia

segmentación manual. Como se ve la fovea segmentada manualmente es mucho más pequeña que la detectada por cualquiera de los procedimientos utilizados.

Con respecto a la convergencia, en la figura 6.8 vemos la gráfica del número de puntos frente al número de iteraciones (barridos completos de la imagen) correspondiente a la segunda imagen. Vemos que ICM alcanza la convergencia en aproximadamente 40 iteraciones mientras que SA en 90 iteraciones. Lo mismo ocurre en las otras gráficas.

De todo lo que se ha dicho, podemos concluir que no hay muchas diferencias entre los resultados obtenidos mediante el SA y aquellas obtenidas utilizando el ICM. Sin embargo, el tiempo de convergencia con el primero es tres veces el tiempo del último, lo cual, en principio, lo haría preferible.

El método propuesto consigue segmentar la imagen de forma global, es decir, extraer venas, arterias y fovea con un mismo procedimiento. Los vasos sanguíneos principales del fondo ocular se detectan correctamente. Sin embargo las pequeñas ramificaciones desaparecen debido a que en esos puntos los parámetros de homogeneidad, incluidos en el modelo *a priori*, tienen un mayor paso por ser estructuras muy delgadas. El método consigue una segmentación de la fovea como una zona compacta situada más o menos en el centro de la imagen, tal como imponíamos en el modelo, pero de un tamaño considerablemente mayor el obtenido manualmente .

La desventaja fundamental de este tipo de procedimientos es su gran coste computacional; la convergencia del algoritmo, tanto si se utiliza el SA o el ICM, es muy lenta.

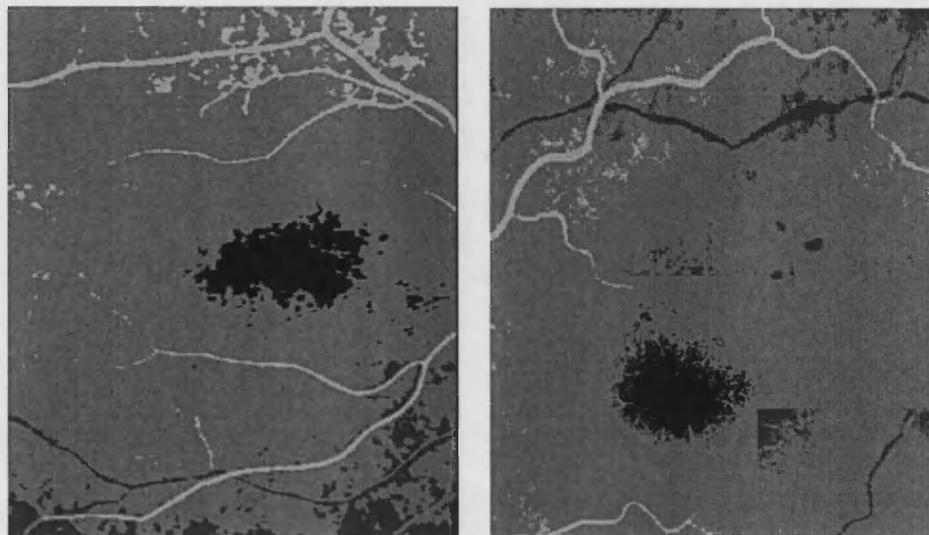


Figura 6.3: Estimadores iniciales para la segmentación de imágenes

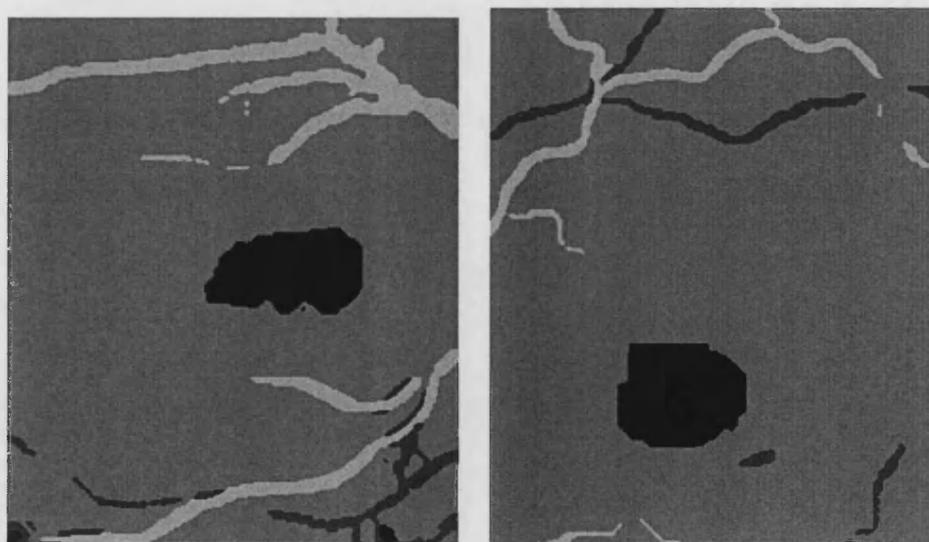


Figura 6.4: Resultados utilizando 50 iteraciones del algoritmo ICM

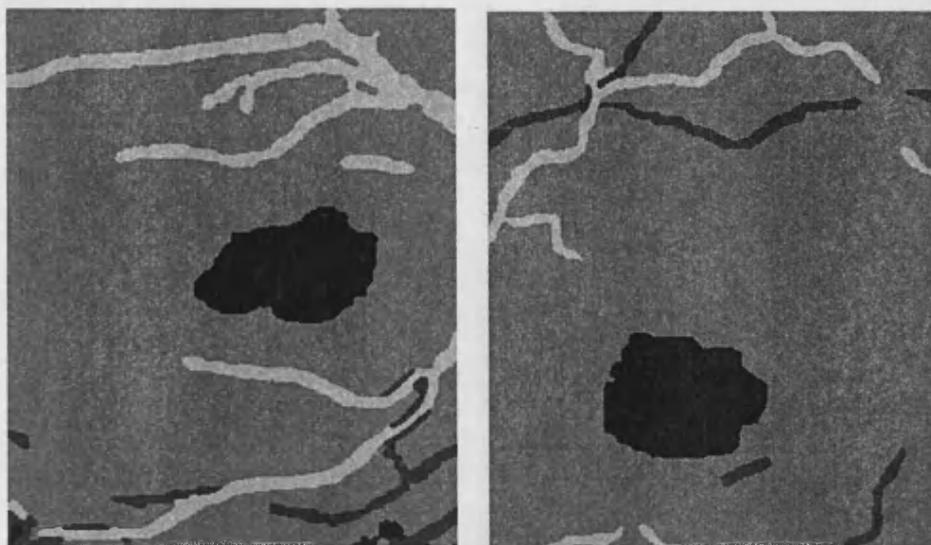


Figura 6.5: Resultados utilizando 100 iteraciones del algoritmo SA



Figura 6.6: Imagen diferencia de los resultados obtenidos con el ICM y SA

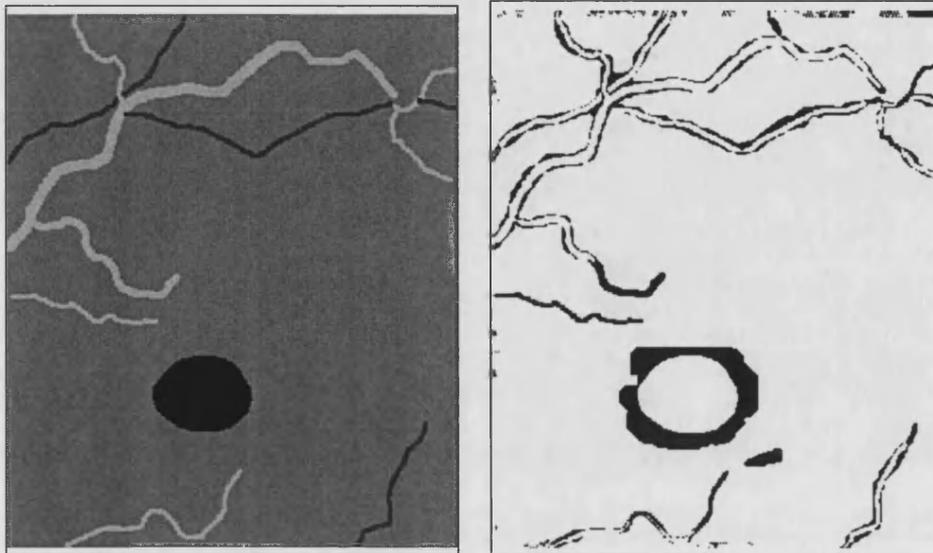


Figura 6.7: Comparación del resultado obtenido por el ICM y una segmentación manual

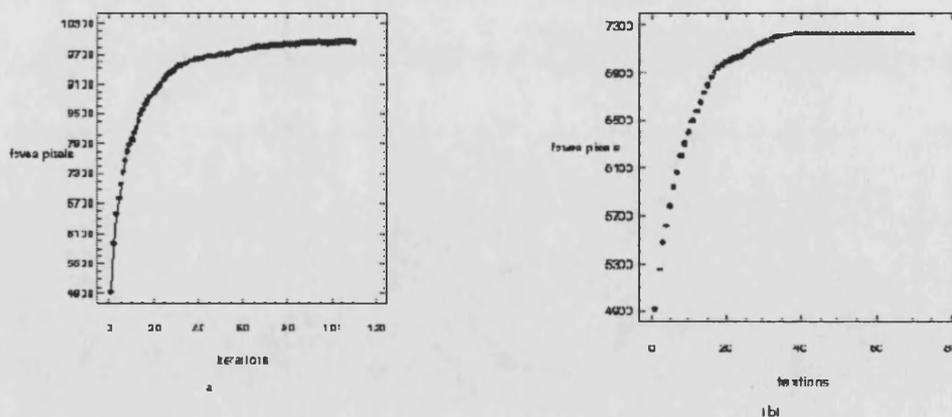


Figura 6.8: Gráficas de convergencia para la imagen (b) de la figura (estabilización del número de *pixels* pertenecientes a la fovea) 6.8. (a) Algoritmo SA. (b) Algoritmo ICM.

Capítulo 7

Evaluación del edema macular

7.1 Introducción

La oclusión de venas retinianas es una enfermedad común de los vasos sanguíneos de la retina que puede desarrollar complicaciones que lleven a la ceguera [Rea89].

Siguiendo a la retinopatía diabética, la oclusión de venas retinianas (incluyendo tanto la oclusión de la vena central como de sus ramificaciones) es la enfermedad vascular de la retina más común. Este capítulo se centra en la evaluación del área macular en un grupo de pacientes que padecen oclusión de venas. Las angiografías fluoresceínicas digitalizadas se emplean para cuantificar el área macular dañada. La macula es la zona donde terminan los vasos mayores de la retina. El centro de la macula es una pequeña área, cóncava, desprovista de capilares retinianos, la *Zona Avascular Foveal (FAZ)*. En las angiografías fluoresceínicas, la FAZ es una posición que ayuda a localizar el centro de la fovea. En la fovea recaen células especializadas que proporcionan la visión central, por lo tanto la severidad de las lesiones dependerán de la distancia al centro de la fovea.

El edema macular es la principal causa de la pérdida de visión en la oclusión tanto de la vena principal como de sus ramificaciones. La angiografía fluoresceínica es un valioso examen para evaluar el edema macular. En esta técnica, como se ha explicado anteriormente, se toman una serie de fotografías del fondo ocular una vez inyectado

Sodio Fluoresceínico y se observa la circulación de este colorante en los vasos de la retina y de la coroides. La fluoresceína puede difundir fuera de los capilares de la retina cuando ésta está dañada. En las zonas donde la difusión está presente la velocidad con la que la fluoresceína decrece se reduce, o incluso puede darse un incremento gradual a lo largo de la secuencia de imágenes. La apariencia angiográfica del edema macular es la aparición de difusión en los capilares de la zona parafoveal y paramacular ya en la primera fase de las angiografías y manchas más tardías en la retina.

En resumen, la severidad de la difusión fluoresceínica está asociada fuertemente con el estado de edema macular. Los factores de alto riesgo en la pérdida de visión en el edema macular están relacionados con la proximidad de las lesiones al centro de la fovea. La cuantificación de la difusión es una forma de evaluar la evolución del edema macular y planear el tratamiento laser.

Existen algunos trabajos previos donde se presentan el uso de imágenes de fondo de ojo para delimitar la difusión macular [MMB⁺94, RWK⁺97] en angiografías fluoresceínicas.

En este capítulo describiremos algunos métodos para la segmentación de la difusión macular, teniendo en cuenta varios aspectos mencionados anteriormente que podrían influir en el pronóstico: el tamaño y la intensidad de la difusión y su posición con respecto al centro de la fovea ([Gro91]). Este procedimiento puede ayudar a cuantificar la difusión y a evaluar objetivamente la progresión del daño macular para establecer el pronóstico visual.

7.2 Métodos

7.2.1 Preprocesamiento de la secuencia

La fijación inestable del ojo junto con el movimiento normal del ojo del paciente hace que se desalineen el sistema óptico de la cámara y el ojo. Los cambios en la iluminación y posición pueden alterar significativamente los resultados del análisis cuantitativo de imágenes. El alineamiento de la cámara es por tanto un paso difícil y crucial en cual-

quier análisis digital de las imágenes de fondo de ojo, para lo existen algunos métodos parcialmente o completamente automáticos de encaje de este tipo de imagen ([JB⁺90]). En el presente trabajo el alineamiento de imágenes se ha conseguido mediante el uso de un método basado en tres procedimientos fundamentales: (i) Detección de puntos relevantes en cada imagen de la secuencia. Se ha utilizado el operador de Förstner's para este propósito [HS92b]. (ii) Encaje de puntos pertenecientes a diferentes imágenes. (iii) Estimación de la transformación geométrica asumida como lineal. La descripción detallada de este método se puede encontrar en el capítulo 3 de la presente tesis doctoral. La figura 7.1 (derecha) muestra los puntos detectados en dos imágenes (primer paso descrito) y las correspondencias establecidas entre ellas (segundo paso).

7.2.2 Segmentación de la difusión

Recordemos que cuando el líquido fluoresceínico difunde fuera de los vasos retinianos en ciertas partes de la retina, la razón de decrecimiento del tinte en ese punto se reduce o incluso existe un incremento de la fluorescencia a lo largo de la secuencia con respecto a la retina adyacente. Un pixel se ha clasificado como *difusión* dependiendo del incremento de nivel de gris a lo largo de la secuencia. Los índices numéricos obtenidos se relacionan con sus distancias a la zona de la fovea.

El centro de la fovea se determinó manualmente para cada paciente. Nuestra región de interés (de ahora en adelante, *ROI*) contiene aquellos *pixels* en una distancia de unas 90 unidades del centro de la fovea. Es por lo tanto una región circular centrada en esta fovea (ver la figura 7.2). Cuanto más cerca está la difusión a la fovea, más peligrosa resulta puesto que la visión central está amenazada.

El criterio utilizado en este trabajo para considerar un *pixel* como perteneciente al área de difusión ha sido comprobado y comparado con otros trabajos previos [MC94, MCMA96].

Nuestra imagen de trabajo será la substracción de dos imágenes de la secuencia ya alineadas, D , como una forma de comparar temporalmente las dos imágenes encajadas. $D(i, j)$ denota el cambio de nivel de gris en el pixel (i, j) . $D(i, j)$ corresponde aproximadamente al incremento de intensidad en el *pixel* (i, j) debido a la difusión.

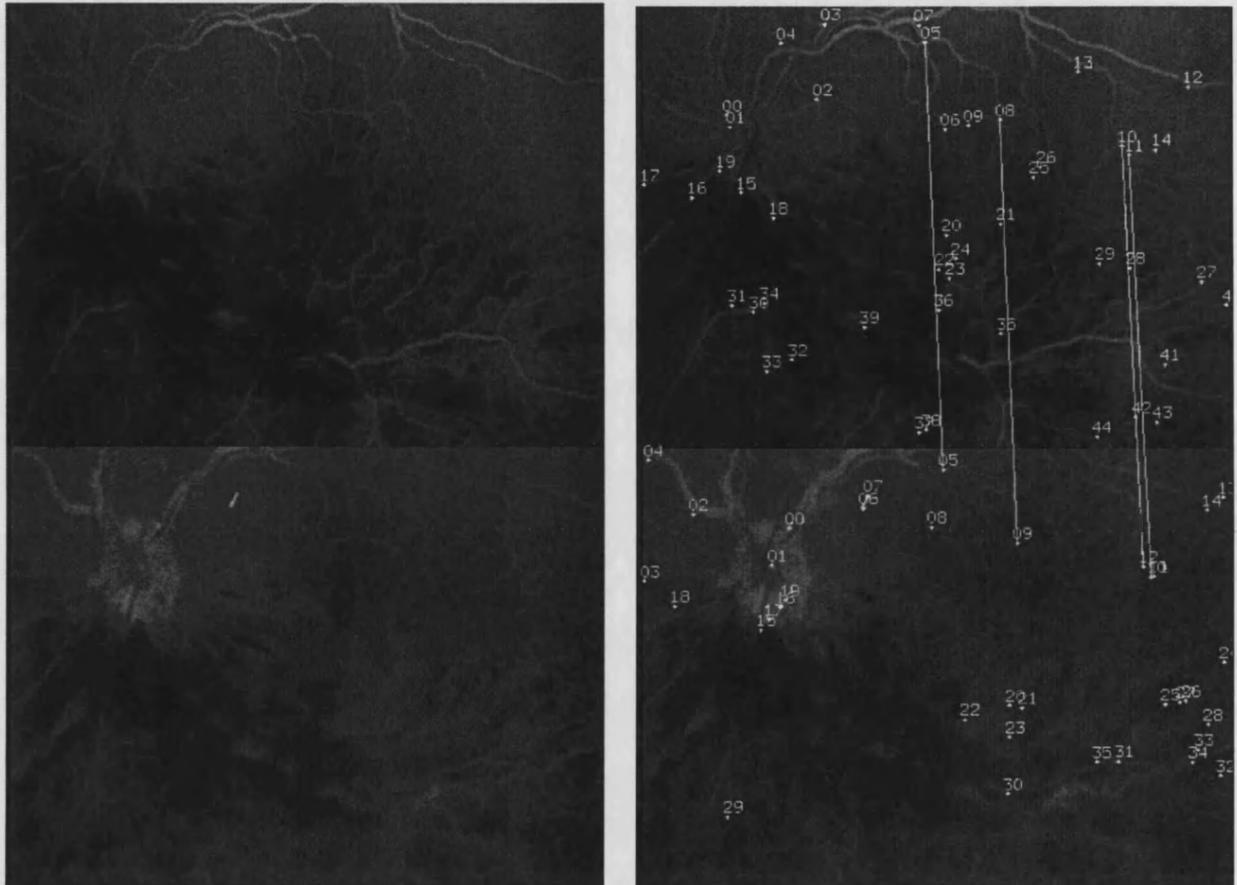


Figura 7.1: Izquierda: Imagen inicial (arriba) y final del caso 1 del angiograma a: oclusión de la ramificación de la vena retiniana de un paciente de 24 años de edad. Derecha: algunas correspondencias establecidas entre las dos imágenes para su posterior alineamiento.

Para determinar si el cambio de nivel de gris en el punto (i, j) ha sido suficiente para considerarlo como difusión, hemos utilizado un método estadístico para detectar valores extremos:

Sea C_0 el círculo centrado en el centro de la fovea con radio 30 de distancia *inter-pixel*, es decir, aquellos *pixels* a una distancia 30 del centro de la fovea. C_1 denota aquellos *pixels* a una distancia entre 30 y 60 y C_2 aquellos *pixels* entre 60 y 90. \bar{D}_{C_i} , S_{C_i} denota los cambios observados en la media y la desviación, $D(i, j)$, en el área C_i , es decir, $\bar{D}_{C_i} = \frac{\sum_{(i,j) \in C_i} D(i,j)}{\#(C_i)}$, $S_{C_i} = \left(\frac{1}{\#(C_i)} \sum_{(i,j) \in C_i} (D(i, j) - \bar{D}_{C_i})^2 \right)^{\frac{1}{2}}$, donde $\#$ denota el número de *pixels*. Un punto (i, j) localizado en la ROI pertenece a C_i . Este *pixel* (i, j) se considera como difusión macular en el nivel k si

$$D(i, j) \geq \bar{D}_{C_i} + kS_{C_i}, \quad (7.1)$$

donde k es una constante positiva.

La idea de base de este método es detectar aquellos *pixels* con un *alto incremento del nivel de gris en la zona más cercana al centro de la fovea* (mayores que el incremento medio más k veces la desviación típica). Cuanto más alto es el valor de k elegido, más severamente se segmentan las zonas (es decir, se extraen aquellas con incrementos extremos de fluorescencia). Trabajando con anillos concentricos alrededor de la fovea podemos obtener información cuantitativa sobre la difusión relacionada con la distancia al centro de la fovea y por lo tanto este método proporciona valiosa información sobre casos de alto riesgo.

En lugar de un valor prefijado de k , se han utilizado cuatro valores diferentes de k $k=(0, 0.5, 0.7, 1)$. Para cuantificar las lesiones, se calculan los siguientes índices numéricos para cada par de imágenes:

- *Cobertura esférica* o porcentaje de zona considerada como difusión en cada dilatación.
- *Cobertura direccional*: para obtener esta medida, *ROI* se ha partido en 20 sectores alrededor del centro de la fovea, y la difusión se ha segmentado para cada uno de ellos. De esta forma, obtenemos un porcentaje de área segmentada como difusión en cada sector.

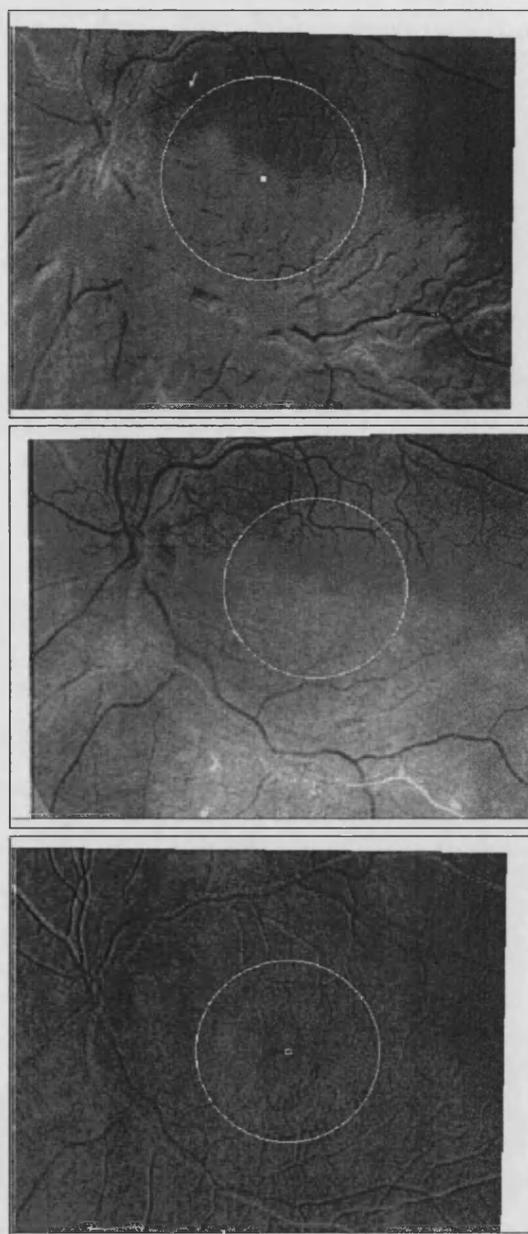


Figura 7.2: Imágenes diferencia correspondientes a los angiogramas a, b y c. Las imágenes están reescaladas al rango de intensidad $[0, 255]$. Los centros foveales y la región de interés aparecen resaltados.

7.2.3 Concordancia con observadores externos

Para evaluar la precisión y utilidad del método, se debe realizar una comparación con expertos. Dos observadores llegaron a un acuerdo para detectar manualmente las áreas de difusión en cada angiograma. Estas áreas se delimitaron manualmente utilizando un *software* comercial (VISILOG¹): se obtiene una imagen binaria que representa el diagnóstico de los observadores y puede compararse con las lesiones binarias obtenidas automáticamente. Se obtienen tres parámetros: falsos positivos, falsos negativos y sensibilidad. El primero de ellos expresa el número de *pixels* en el área de interés marcados como difusión por el programa pero no por los observadores, el segundo expresa el número de *pixels* marcados por los observadores pero no por el programa. Estas dos medidas son dadas como un porcentaje con respecto al número total de *pixels* del área de difusión real (marcada por los observadores). La suma de los dos parámetros da una idea de la proporción del área total que ha sido clasificado erróneamente por el programa. Por último, se ha obtenido además el parámetro de sensibilidad (porcentaje de *pixels* marcados por los expertos que son también detectados por el programa).

7.3 Material

Los negativos de las angiografías fluoresceínicas tomadas con una cámara de fondo de ojo *Canon CF-60U* se transformaron en señal de vídeo con una foto-vídeo cámara *PHV-A7E, Sony*. Estas imágenes se digitalizaron con una MATROX-PIP-EZ (384 x 512 *pixels*; 256 niveles de gris) instalada en un PC. Las librerías de dominio público VISTA² han sido utilizadas para la implementación del software. Los algoritmos de análisis de imagen han sido desarrollados mediante el uso de una workstation Hewlett-Packard 725.

Se seleccionaron dos imágenes de cada serie, correspondiente a los siguientes momentos: i) la fase de pico de la angiografía fluoresceínica (20 – 25 s)[Rea89] y ii) la fase que comienza 5 minutos después de la inyección. En la primera imagen los capilares

¹VISILOG es una marca registrada de Noesis

²Desarrolladas principalmente por Art Pope en el Departamento de Ciencias de la Computación, Universidad de British Columbia.

paramaculares de la retina se llenan de fluoresceína y la difusión comienza. En la segunda imagen, los vasos sanguíneos se vacían y la difusión tiene lugar produciéndose áreas de hiperfluorescencia. El campo fotográfico se centró en la macula con un campo de visión de 60°.

La figura 7.1 izquierda muestra una imagen temprana (arriba-izquierda) y una tardía (abajo-izquierda) de una secuencia de angiografías fluoresceínicas.

7.4 Resultados

Se han digitalizado 21 secuencias de angiografías fluoresceínicas de 16 pacientes diferentes que padecían trombosis en las ramas de las vasos retinianos tal y como se explicó anteriormente. El caso 1 (figura 7.1, arriba izquierda) corresponde a una mujer de 24 años afectada de oclusión de las ramificaciones de venas junto con un síndrome antifosfolipídico. Estas imágenes presentan hemorragias múltiples cerca a la vena ocluida (más oscura en la imagen). Los capilares inferiores paramaculares dañados difunden la fluoresceína produciendo un edema macular (ver figura 7.1, abajo izquierda). Tres angiogramas (de ahora en adelante, angiogramas a, b y c) de esta paciente se han digitalizado y analizado para apreciar la utilidad del método en el seguimiento de los cambios de la difusión y la efectividad del tratamiento laser. Hemos utilizado este caso como ejemplo.

Cada secuencia se ha alineado utilizando el método de encaje descrito en el capítulo tres de la presente tesis. De esta forma, obtenemos dos imágenes alineadas y su correspondiente sustracción. La figura 7.2 muestra la sustracción de las imágenes correspondiente a los angiogramas a, b y c.

Las imágenes sustraídas se analizan mediante otro programa de ordenador para segmentar los puntos con difusión alrededor del centro de la fovea. La lesión hiperfluorescente detectada manualmente correspondiente al angiograma a se muestra en la figura 7.3 (arriba derecha). La figura 7.3 (abajo izquierda) muestra la difusión segmentada automáticamente.

Los falsos positivos, falsos negativos y la sensibilidad del método se calculan en-

tonces para cada secuencia. La tabla 7.1 muestra las razones de concordancia de un observador (entre un 75 y un 98%) cuando se utiliza un valor intermedio de k ($k = 0.5$). Consideramos este valor como el más similar al criterio de los observadores con una sensibilidad media de 91%.

Los resultados de la cuantificación del edema macular aparecen en la figura 7.4. La *cobertura esférica* da información global sobre los cambios en el nivel de fluorescencia en la *ROI*. Para cuantificar la distribución angular de la difusión, la *ROI* ha sido dividida en 20 sectores alrededor del centro de la fovea y así, cada línea que parte del centro refleja la difusión en cada sector siendo su longitud proporcional a la intensidad de la difusión. Cuando miramos las gráficas correspondientes al caso 1 podemos seguir 3 series de angiogramas, tomados cuatro meses después, para obtener datos numéricos precisos sobre la resorción del edema macular.

En el angiograma a (ver figura 7.4), la gráfica de la cobertura esférica (arriba izquierda) muestra que la difusión es máxima alrededor del centro de la fovea, representando el 100 % del primer anillo de ditalatación. En una distancia de 20 *pixels* alrededor del centro de la fovea, la difusión ocupa el 40% del anillo. La gráfica de la cobertura direccional (arriba derecha) representa la distribución de la difusión relativa. La intensidad de la difusión máxima corresponde al vector de 270 grados (inferior al centro de la fovea).

En el angiograma b (ver figura 7.4), la gráfica de *la cobertura esférica* (centro izquierda) muestra que la difusión ha decrecido globalmente, y se localiza a una mayor distancia de la fovea. Comienza a incrementar a una distancia de 10 *pixels* alrededor del centro de la fovea, y es máxima entre 20 y 30, cuando alcanza el 30%. Como muestra la gráfica de *la cobertura direccional* (centro derecha), los sectores afectados predominantemente se localizan (entre 270 y 300 grados alrededor del centro de la fovea).

Las gráficas de cuantificación del angiograma c (figura 7.4, abajo) revelan que la difusión ha decrecido dramáticamente en severidad (no más de un 6% en el anillo más afectado), y la gráfica es casi plana. Los observadores no encuentran ninguna difusión en este par de imágenes. La detección automática mostró solamente tres puntos fluoresceínicos mínimos, correspondientes a cicatrices de laser.

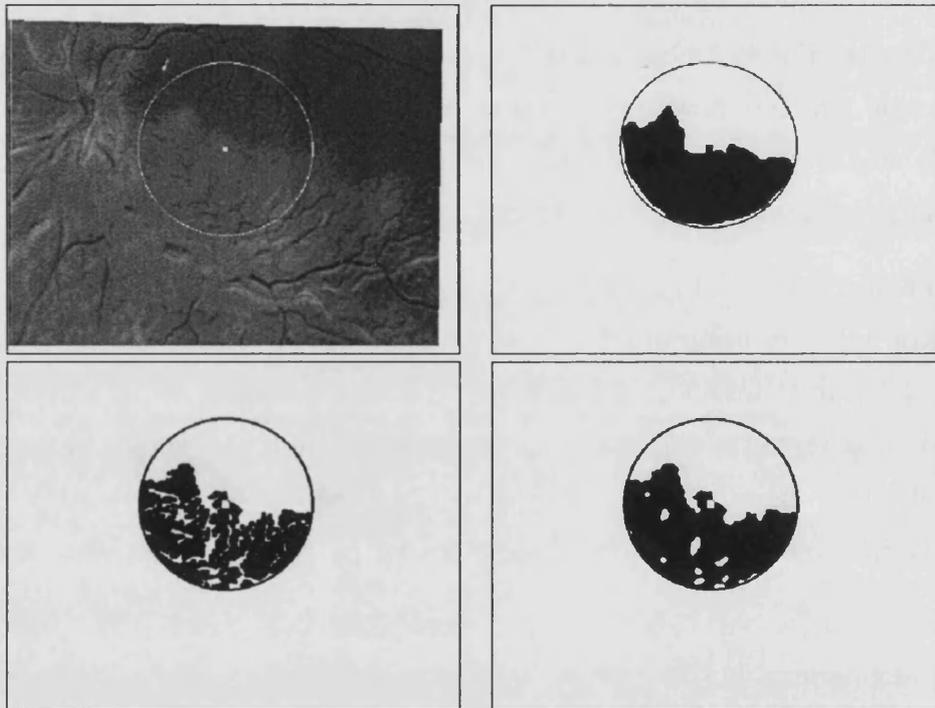


Figura 7.3: Angiogramas a, comenzando por la fila superior y de izquierda a derecha: imagen diferencia, difusión segmentada por los expertos, difusión segmentada automáticamente y difusión detectada automáticamente después de una apertura.

7.5 Discusión

Imágenes digitales del fondo ocular se han convertido en un examen usual en los últimos años [CFG92, MSea91, NSR85, Sha97]. Se han hecho esfuerzos para detectar y cuantificar automáticamente varias estructuras anatómicas y patológicas del fondo ocular (exudados ([WTT89]), red de capilares ([JAB⁺92, RJ⁺93, RWK⁺97]), isquemia macular ([MMB⁺94]), microaneurismas ([SOMH⁺96])). Phillips y colaboradores [PSR⁺91, PRT⁺91] han trabajado en este campo y calculado valores umbral del gradiente para evaluar la difusión macular diabética.

Como las imágenes han mostrado claramente, la angiografía fluoresceínica es una exploración dinámica complicada; los elementos anatómicos y patológicos experimentan cambios a veces drásticos de una imagen a otra a lo largo de la evolución de la enfermedad evaluada. Las diferencias individuales en anatomía y la pigmentación del fondo de ojo, variaciones del flujo sanguíneo y movimientos oculares entre imágenes interfieren con la detección automática de las lesiones.

En este capítulo, se ha propuesto un nuevo método para la segmentación de la difusión macular el cual ha sido probado con 21 secuencias de Angiografías Fluoresceínicas.

El área central alrededor de la fóvea se ha considerado como una de las características más importantes relacionadas con la función visual. Las áreas de difusión no se segmentan considerando valores absolutos de hiperfluorescencia sino incrementos de niveles de gris entre dos imágenes de cada caso, para evitar falsos positivos debidos a una pigmentación pobre de algunos pacientes.

Según Spencer [SPSF92] es difícil obtener un estandar universal para las áreas de difusión no fácilmente diferenciadas de la fluorescencia del fondo, vasos o cualquier otra característica adyacente a las lesiones. Para minimizar el problema, utilizamos un enfoque estadístico para detectar incrementos fluoresceínicos grandes a lo largo de la secuencia. Las variaciones del flujo de sangre en la intensidad fluoresceínica son pequeñas comparadas con el incremento en los niveles de gris debidos a la difusión, por lo que puede obviarse eligiendo valores adecuados de k (ver sección 7.2.2).

En lugar de elegir un área cuadrada de la zona de interés centrada en la fóvea como

previamente presentaron [PRT⁺91], el centro foveal se resalta y dilata automáticamente mediante métodos de morfología matemática [Ser82]. De esta forma, se obtiene la información numérica sobre las imágenes relacionada con el parámetro más importante: la distancia al centro de la fovea.

Los oftalmólogos pueden observar las gráficas de cuantificación para determinar donde se sitúa la principal distribución de difusión alrededor de la fovea. De hecho, en la mayoría de los casos que son considerados como edema difusivo, nuestro método detecta direcciones predominantes de difusión no perceptibles por análisis visual. El tamaño de *ROI* para cada paciente puede ser elegido por el usuario, así como el grado de hiperfluorescencia en la zona segmentada (valor del incremento k), por lo que el especialista actúa interactivamente con el ordenador, en un proceso de diagnóstico asistido más que en un proceso de diagnóstico automático. Cuanto más alto es el valor de k , mayor es el porcentaje de falsos negativos y el menor el porcentaje de falsos positivos obtenidos. La sensibilidad varía de acuerdo con esto.

Los rangos de concordancia con el observador están comprendidos entre un 75 y 98% para los 21 angiogramas analizados cuando se compara la difusión detectada manualmente con las áreas segmentadas por el ordenador mediante el uso de los valores de k más sensibles. Un valor de k de 0.5 da una segmentación más similar a la del observador (una media de sensibilidad de 91%). Estamos de acuerdo con otros autores que existe una mayor sensibilidad del análisis por ordenador que cuando se realizan por un observador y una mayor reproducibilidad que en la detección manual. De hecho, el cansancio del observador y su variabilidad son factores importantes ([SPSF92]).

La *ROI* se analiza automáticamente punto a punto para segmentar las lesiones, mientras que el oftalmólogo es propenso a delimitar áreas conectadas como difusión. Si se prefiere obtener una detección de la difusión más similar a la de los observadores, es decir áreas conectadas, podemos de nuevo aprovechar los métodos de morfología matemática haciendo una apertura de la difusión segmentada automáticamente (ver figura 7.3, abajo derecha).

En resumen, consideramos que este procedimiento ayuda a cuantificar la difusión y evaluar objetivamente la progresión del daño macular, para establecer un pronóstico visual. En cualquier caso, una decisión terapéutica debe ser basada obviamente en



datos clínicos más que en información de una imagen para obtener conclusiones clínicas juiciosas. El tiempo necesario para analizar un par de imágenes es 120-130 segundos (114-117 segundos para el encaje de imágenes y unos pocos segundos para la detección y cuantificar la difusión). El procedimiento es completamente automático, excepto para la localización del centro de la fovea. La detección de la difusión fuera de la región macular se podría realizar con este programa seleccionando otro punto de la retina como centro de la *ROI*.

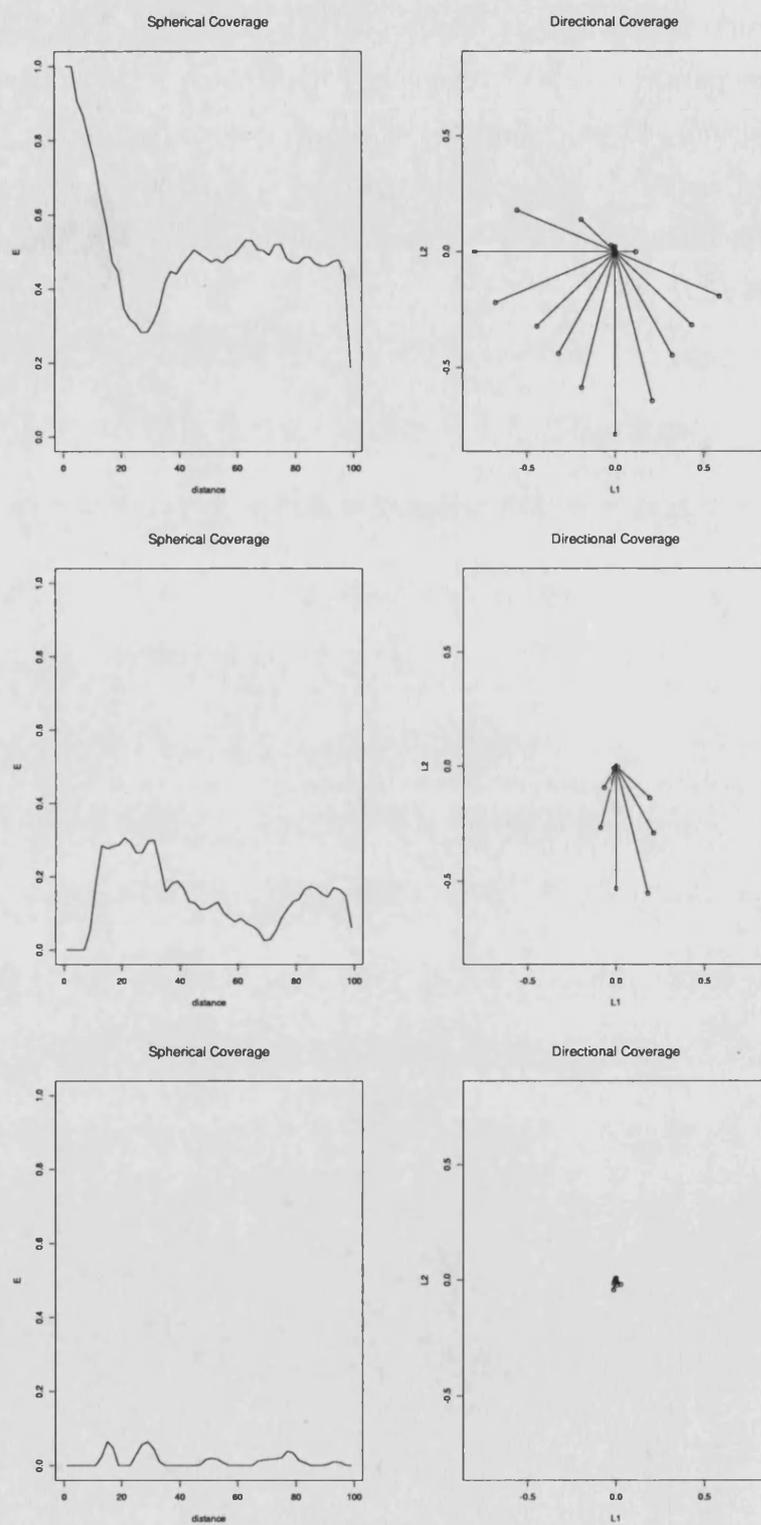


Figura 7.4: Cuantificación correspondientes a los angiogramas a, b y c

Caso	Falsos positivos	Falsos negativos	Sensibilidad
1	2 %	22 %	75 %
1b	3 %	15 %	81 %
1c	0.06 %	1.7 %	98 %
2	5 %	6%	89 %
3	1 %	6 %	93 %
4	4 %	9 %	86 %

Tabla 7.1: Comparación entre la detección realizada por los observadores y la automática Resultados experimentales correspondientes a seis angiogramas con un valor intermedio de k , 0.5.

Capítulo 8

Conclusiones, aportaciones y trabajo futuro

8.1 Resultados y aportaciones

Desde el punto de vista de la Visión por Computador y concretamente del análisis de imágenes médicas los resultados y aportaciones que se obtienen del trabajo realizado son las que se resumen a continuación:

- Dado el gran interés que tiene el alineamiento de una secuencia (cualquier estudio temporal de lesiones dependerá de él) se han propuesto y analizado nuevos métodos para llevarlo a cabo. Se diferencian principalmente en el tipo de estructuras anatómicas utilizadas para realizar el encaje. Dos de los métodos desarrollados, descritos en el capítulo 3, se basan en el mismo modelo estocástico y en la detección y posterior encaje de puntos de interés mediante un procedimiento iterativo. Difieren en el modo de actualizar las correspondencias iniciales. Únicamente se necesita un conjunto de puntos junto con un vector de características locales asociadas a cada uno de ellos. La detección de los puntos relevantes se ha realizado mediante el detector de Förstner. Este operador ha dado muy buenos resultados en su aplicación a nuestro problema. Sin embargo, parece interesante hacer un estudio del comportamiento del algoritmo según otros detectores que aparecen en la literatura. Se ha analizado el funcionamiento de estos métodos observando una serie de factores: la

distribución espacial de los puntos en la imagen, el número de correspondencias iniciales encontradas y el error cuadrático medio asociado. Los experimentos muestran como la imposición de un número muy pequeño de correspondencias iniciales puede hacer que el algoritmo falle. Por otro lado, un número de correspondencias iniciales numeroso hace que aumente el coste computacional del algoritmo puesto que el árbol de búsqueda tiene una complejidad exponencial. Además, la precisión del algoritmo en cuanto al error cuadrático medio obtenido depende de la distribución espacial de los puntos. Si todos los puntos están localizados en una zona, el error cuadrático obtenido para puntos alejados de esa área será grande. Se observa un mejor funcionamiento del segundo método en cuanto al error cuadrático obtenido. Esto se debe a que es capaz de obtener un conjunto de correspondencias finales más numeroso. En general podemos decir que estos métodos son fiables y robustos para el alineamiento de imágenes en las que aparecen cambios bruscos de intensidad y contraste.

- Se ha propuesto un procedimiento para alineamiento inspirado en el algoritmo EM generalizado, que, a diferencia de los métodos anteriores, utiliza grafos relacionales para el encaje. Con este nuevo método se ha mostrado la utilidad de estructuras más complejas en el alineamiento de imágenes. Los grafos permiten introducir información contextual en el proceso de encaje y representan en nuestras imágenes la estructura vascular de las imágenes de fondo de ojo. Los nodos del grafo corresponden a cruces de vasos sanguíneos y existe una arista entre nodos si están conectados por un vaso. Este método es capaz de obtener un gran número de correspondencias entre las sucesivas imágenes, lo que hace disminuir considerablemente el error cuadrático asociado al método. Sin embargo, en muchos casos el algoritmo cae en mínimos locales y se obtienen algunas correspondencias incorrectas.
- La mayoría de los algoritmos de alineamiento de imagen de fondo de ojo utilizan como modelo de transformación geométrica una transformación afín, que da cuenta de las rotaciones, traslaciones y cambios de escala entre dos imágenes. Puesto que sabemos que el desalineamiento de las imágenes se ha producido por las rotaciones del ojo, hemos modelizado la transformación geométrica mediante una transformación no lineal que tiene en cuenta el movimiento real del ojo, aproximadamente esférico, y su proyección sobre una superficie plana. De esta forma, se obtiene un modelo

de transformación que depende de dos tipos de parámetros: los del movimiento del ojo (ángulos de rotación) y los del sistema óptico utilizado. Se ha analizado el error cuadrático medio cometido con esta transformación, en un estudio comparativo con el obtenido mediante el de una transformación afín. En este estudio se han utilizado imágenes sintéticas tomadas de un modelo de ojo fabricado con este propósito. La conclusión de este análisis es que la transformación geométrica propuesta da, en este caso, errores cuadráticos menores que una transformación afín. Sin embargo, cuando el análisis se realiza sobre imágenes angiográficas reales, los resultados no son tan satisfactorios. La transformación no lineal da errores mayores que los obtenidos por la transformación afín. Hay que tener en cuenta que la transformación propuesta se ha deducido suponiendo un sistema óptico de la cámara muy sencillo, y además asumiendo que el ojo es esférico. Además, el uso de esta transformación implica la utilización de métodos numéricos para su estimación, con lo que no es exacta. Este análisis nos conduce a rechazar la transformación no lineal en favor de una transformación afín para el alineamiento de imágenes de fondo de ojo.

- Se ha desarrollado un nuevo método de segmentación de las angiografías fluoresceínicas en sus elementos característicos. Este procedimiento permite obtener una segmentación global que, a diferencia del resto de los algoritmos propuestos con este fin, clasifica los *pixels* de la imagen como pertenecientes a venas, arterias, fovea y fondo de ojo. Está basado en técnicas bayesianas que permiten incorporar el conocimiento previo de las estructuras que aparecen normalmente en este tipo de imágenes. Esta información se incorpora en el método a través de la distribución de probabilidad *a priori*. La tarea de segmentación se reformula como un problema de estimación de parámetros en el que la imagen observada se considera una versión ruidosa de la segmentada. El estimador utilizado en este trabajo es el máximo *a posteriori* (MAP). En este contexto, los modelos apropiados para la distribución *a priori* son los campos aleatorios de Markov puesto que son capaces de modelizar información contextual. En este caso particular, se ha necesitado un campo aleatorio de Markov no homogéneo. Este procedimiento nos proporciona una imagen segmentada que se ajusta a los requisitos impuestos a través de la distribución *a priori* de la imagen. La fovea aparece como una zona compacta y situada más o menos en el centro de la imagen. Las venas y arterias principales aparecen en el

resto de la imagen como estructuras localmente lineales. Sin embargo, estos vasos sanguíneos se fragmentan en algunos puntos de la imagen y las pequeñas ramificaciones de vasos no se detectan. A pesar de sus limitaciones, (no hay que olvidar que hemos propuesto un método global que debe detectar diferentes estructuras de las imágenes donde los cambios de intensidad y contraste son frecuentes) los resultados son bastante alentadores y se pueden incorporar mejoras en el método que nos permitan solventar el problema de ruptura de los vasos sanguíneos. Una desventaja clara del método propuesto es su alto coste computacional. El estimador MAP para la imagen segmentada se realiza mediante dos métodos diferentes: el recocido simulado (*Simulated Annealing*) y las modas condicionales iterativas *Iterative Conditional Modes*. Los dos algoritmos proporcionan resultados similares. Sin embargo, el coste computacional del segundo de ellos es tres veces menor, lo que lo hace preferible.

- Se ha desarrollado y analizado un método para la segmentación de la difusión macular teniendo en cuenta varios aspectos que pueden influir en el diagnóstico: el tamaño, la intensidad de la difusión y su posición con respecto al centro de la fóvea. Este procedimiento parte de una imagen diferencia obtenida a partir de dos imágenes de la secuencia alineada. La idea clave del procedimiento es detectar aquellos *pixels* con un alto incremento del nivel de gris en la zona más cercana al centro de la fóvea. El método se ha evaluado comparándolo con el resultado dado por un observador experto. Los rangos de concordancia con el experto están comprendidos entre un 75% y un 98% que da al método una alta fiabilidad. Las ventajas de la segmentación automática de la difusión son claras: el tiempo necesario y la reproducibilidad de resultados. Este procedimiento analiza punto a punto el área de interés para la segmentación de lesiones y no tiende, como hacen los expertos, a delimitar áreas conectadas como difusión lo que lo convierte en un método muy objetivo. De este modo, se ha desarrollado un procedimiento totalmente automático que ayuda a cuantificar la difusión macular y evaluar objetivamente la progresión del daño macular.

Los resultados de este trabajo de investigación se podrían resumir en el desarrollo de algoritmos específicos para el alineamiento, segmentación y detección de lesiones oculares que permitan automatizar la extracción de datos con los que los expertos

pueden emitir un diagnóstico. Estos procedimientos, aunque se han desarrollado con un propósito claro, creemos que son extensibles a otros problemas similares sin grandes modificaciones.

8.2 Trabajo futuro

Muchas son las líneas de trabajo que quedan abiertas y que se resumen a continuación.

En cuanto a los métodos de alineamiento propuestos, pensamos que son suficientemente fiables para que el alineamiento se puede realizar de forma completamente automática. No parece, pues, necesario realizar ninguna modificación de los métodos de alineamiento que mejoren directamente los resultados en su aplicación al campo de la medicina, y concretamente al diagnóstico asistido de la retinopatía diabética. La única mejora que nos planteamos corresponde a la transformación geométrica utilizada que podría ser modificada por una aproximación polinómica de orden superior al lineal que de cuenta de pequeñas distorsiones y que a su vez pueda ser estimada de forma exacta. También tiene interés, como ya se dijo, la realización de un estudio del comportamiento de estos algoritmos según los diferentes detectores de puntos relevantes.

Con respecto a la segmentación de angiografías, se ha visto que, aunque se ha conseguido una segmentación global de la imagen en sus componentes de interés, el método propuesto no es capaz de detectar las pequeñas ramificaciones de los vasos sanguíneos. Una línea clara de trabajo sería la incorporación de otro tipo de información en el modelo utilizado que nos permita segmentar este tipo de estructuras de pequeño tamaño. No obstante, métodos alternativos que detecten individualmente cada tipo de estructura en la imagen de fondo de ojo parecen más fiables por lo que creemos interesante desarrollar otros tipos de procedimientos. Se ha desarrollado un método que permite estimar las orientaciones de los bordes que confluyen en un punto unión (en nuestras imágenes, cruces), así como el número de bordes [SdVD⁺99]. Este procedimiento nos permitiría realizar un seguimiento de las venas a partir de estos puntos.

En cuanto al diseño de algoritmos concretos para la detección y cuantificación de las lesiones es, probablemente, el campo con mayor potencial futuro. En primer lu-

gar y siguiendo con la detección del edema macular, pensamos desarrollar algoritmos que nos permitan realizar un seguimiento de la difusión a lo largo de un angiograma completo. Con este propósito estamos trabajando en nuevos descriptores de formas que nos permitan caracterizarla de una forma adecuada ([dVDS⁺99]). En segundo lugar, sabemos que uno de los síntomas claves en la retinopatía diabética es la aparición de microaneurismas. Estos microaneurismas se manifiestan como pequeños puntos de forma aproximadamente circular por lo que su detección puede realizarse mediante el desarrollo de algún tipo de procedimiento de detección de aspectos circulares. Algunos estudios preliminares sobre ello han sido ya realizados con resultados interesantes.

La angiografía es un examen clínico que conlleva algunos riesgos por lo que no es recomendable su realización a grandes grupos de población. Una línea de trabajo directa sería la detección de microaneurismas utilizando otro tipo de imágenes de fondo de ojo como son las aneritras y las de color. Un estudio de la fiabilidad de realizar exploraciones con este tipo de imágenes tiene un evidente interés, tanto clínico como económico.

Apéndice A

Campos aleatorios

A.1 La imagen como un campo aleatorio

La utilización de campos aleatorios en el análisis de imágenes digitales surge de considerar la imagen como un conjunto finito de datos almacenados en una matriz rectangular. Cada una de las posiciones de la matriz lleva asociada información sobre el nivel de gris existente en un determinado punto de la imagen. De este modo se trabaja con un rejilla discreta cuyas posiciones vienen determinadas por posiciones enteras (i, j) . Desde un punto de vista estadístico, cada uno de los puntos (i, j) de la imagen tiene asociada una variable aleatoria $X(i, j)$ donde el conjunto finito de estados corresponde con los distintos niveles de gris. Una imagen se verá como una realización de un campo aleatorio con un conjunto de estados finitos que corresponden a los posibles niveles de gris de la imagen.

A.2 Campos aleatorios

Sea S el conjunto de todas las posiciones de la rejilla y para cada $s = (i, j) \in S$ sea X_s el conjunto finito de posibles valores. El producto $\mathbf{X} = \prod_{s \in S} X_s$ será el espacio finito de todas las posibles configuraciones $x = (x_s)_{s \in S}$. Podemos considerar distribuciones de probabilidad P sobre \mathbf{X} de forma que $P(x) \geq 0$ y $\sum_{x \in \mathbf{X}} P(x) = 1$. **Campo aleatorio**

o **proceso estocástico espacial** es la colección de variables $\{X(s) : s \in S\}$ con la función de probabilidad P .

Si un campo aleatorio cumple que $P(x) > 0 \forall x \in X$, se dice que cumple la condición de positividad. Definimos el conjunto Ω como el conjunto de configuraciones que cumple dicha condición, es decir, $\Omega = \{x \in X : P(x) > 0\}$.

Definición. Diremos que la colección de conjuntos $\delta = \{\delta(s) : s \in S\}$ es un sistema de vecindad si: (i) $s \notin \delta(s)$ y (ii) $s \in \delta(t)$ si y solo si $t \in \delta(s)$. $\delta(s)$ es, pues, el conjunto de vecinos de s en el sistema de vecindad $\delta(s)$

Definición. Cualquier subconjunto C de S diremos que es un clique si tiene un único elemento o si cualquier par de elementos de C son vecinos.

Una vez definido el concepto de vecindad y clique, muy importantes en el contexto de los campos aleatorios, definiremos los campos aleatorios de Markov y los de Gibbs por cumplir ciertas propiedades que los hacen especialmente útiles.

Definición. Diremos que un campo aleatorio es de Markov para el sistema de vecindad δ si:

$$P(x(s_i)/x(s_j) : j \neq i) = P(x(s_i)/x(s_j) : s_j \in \delta(s_i)). \quad (\text{A.1})$$

Según esta definición cualquier campo aleatorio será de Markov para el sistema de vecindad en la que todos los diferentes puntos de la rejilla son vecinos. Sin embargo, estamos interesados en campo aleatorio que son de Markov para pequeñas vecindades lo que implica independencia de puntos de las variables aleatorias $X(s)$ y $X(s')$, si s y s' están suficientemente alejados espacialmente.

Definición. Diremos que un CA es un campo aleatorio de Gibbs si se puede expresar como:

$$P(x) = \frac{\exp(Q(x))}{\sum_{z \in X} \exp(Q(z))}. \quad (\text{A.2})$$

La función $Q(x)$ se llama función negpotencial, y $H(x) = -Q(x)$ recibe el nombre de función de energía.

La terminología de campos de Gibbs se ha tomado de la mecánica estadística, donde se utilizan como modelos para los estados de equilibrio en sistemas físicos.

Es obvio que cualquier campo aleatorio puede expresarse como un campo de Gibbs tomando una determinada configuración como referencia, por ejemplo, tomando el 0 como una configuración posible para cualquier posición, es decir, que $0 \in \mathbf{X}$. Dado un campo aleatorio, la función negpotencial necesaria para que el campo aleatorio tome la forma de un campo de Gibbs se expresa como sigue:

$$Q(x) = \log \frac{P(x)}{P(0)}, x \in \mathbf{X} \quad (\text{A.3})$$

De este modo el conocimiento de $P(x)$ equivale al conocimiento de $Q(x)$ o viceversa.

A.3 El teorema Hammersley-Clifford

La equivalencia entre campo aleatorio de Markov y campo de Gibbs la proporciona el teorema de Hammersley-Clifford, también conocido como teorema de equivalencia. Este teorema responde a la pregunta de como obtener la distribución conjunta dada una distribución condicional. Esto es de gran importancia ya que es la probabilidad conjunta la que contiene la representación completa de la imagen.

Asumimos que cada variable $X(s)$, nivel de gris en pixel s , toma valores de un conjunto discreto de valores, a los que pertenece el cero, y que todas las configuraciones tienen una probabilidad no nula, es decir, $P(x) > 0$. También definimos $Q(x) = \log(\frac{p(x)}{p(0)})$ y sea $\mathbf{x}(s_i)$ la configuración $(x(s_1), \dots, x(s_{i-1}), 0, x(s_{i+1}), \dots, x(s_n))$. La forma de la función negpotencial $Q(x)$ nos da la forma de la estructura de la probabilidad condicional, puesto que

$$\exp(Q(x) - Q(\mathbf{x}(s_i))) = \frac{P(x(s_i)/x_{s_j} : j \neq i)}{P(0(s_i)/x(s_j))} \quad (\text{A.4})$$

La función energía $Q(x)$ se puede expandir sobre \mathbf{X} de la siguiente forma:

$$Q(x) = \sum_{s_1 \leq s_i \leq s_n} x(s_i)G_{s_i}(x(s_i)) + \sum_{s_1 \leq s_i < s_j \leq s_n} \sum x(s_i)x(s_j)G_{s_i s_j}(x(s_i), x(s_j)) + \dots + x(s_1) \dots x(s_n)G_{s_1 \dots s_n}(x(s_1), \dots, x(s_n)) x \in \mathbf{X} \quad (\text{A.5})$$

donde las funciones G son arbitrarias. De la ecuación A.4 la especificación de la probabilidad condicional dice que $Q(x) - Q(x_i)$ solamente puede depender de los valores de $x(s_i)$ y sus vecinos. Para la posición 1 tenemos

$$Q(x) - Q(x_{s_1}) = x(s_1)(G_{s_1}(x(s_1)) + \sum_{s_2 \leq s_j \leq s_n} x_{s_j} G_{s_1, s_j}(x(s_1), x(s_j)) + \sum_{s_2 \leq s_j < s_k \leq s_n} \sum x(s_j)x(s_k) G_{s_1, s_j, s_k}(x(s_1), x(s_j), x(s_k)) + \dots + x(s_2)x(s_3) \dots x(s_n) G_{s_1, s_2, \dots, s_n}(x(s_1), x(s_2), \dots, x(s_n))) \quad (\text{A.6})$$

Si la posición s_l no es vecina de s_1 , $Q(x) - Q(x_{s_1})$ debe ser independiente de x_{s_l} . Consideremos $x_{s_i} = 0$ para $s_i \neq s_1, s_l$ y así $G_{s_1, s_l}(x_{s_1}, x_{s_l}) = 0$. Otras elecciones de ordenes mayores de las funciones G implicarían que x_{s_1} y x_{s_l} deben ser nulos. Consideraciones de parejas de posiciones no vecinas dan analogos resultados. De esta forma G_{s_i, s_j, \dots, s_k} solamente es no nulo si s_i, s_j, \dots, s_k forman un clique. Después de estos importantes resultados veamos el enunciado del teorema de equivalencia.

Teorema de Hammersley-Clifford *Supongamos que X se distribuye según un campo aleatorio de Markov sobre Ω que satisface la condición de positividad. Entonces, la función negpotencial Q verifica que si las posiciones s_i, s_j, \dots, s_k no forman un clique, entonces $G_{s_i, s_j, \dots, s_k}(\cdot) = 0$.*

Veamos el teorema de equivalencia de otro modo. Sea C un clique y definimos

$$x_C = (x_i : i \in C) \quad (\text{A.7})$$

$$V_C(x_C) = G_C(x_C) \prod_{i \in C} x_i \quad (\text{A.8})$$

Y entonces podemos escribir $Q(x)$ como:

$$Q(x) = \sum_{C \in \mathcal{C}} V_C(x_C) \quad (\text{A.9})$$

De las ecuaciones A.4 y A.9 concluimos que:

$$P(x_{s_i}/x_{s_j} : s_j \neq s_i) \propto \exp\left(\sum_{C:i \in C} V_C(x_c)\right), \forall i = 1, \dots, n \quad (\text{A.10})$$

Hemos encontrado una importante relación entre los campos de Gibbs y los campos de Markov. Un campo aleatorio de Markov se puede expresar como un campo de Gibbs en el que la función negpotencial $Q(x)$ se expresa como un sumatorio de potenciales que solo dependen de las variables que pertenecen a un clique C .

Apéndice B

El algoritmo EM

B.1 Descripción del algoritmo

El algoritmo conocido como Esperanza-Maximización (*Expectation-Maximization*) es una herramienta que ayuda a calcular los estimadores de máxima verosimilitud en un problema con datos no observados. Supongamos que tenemos un conjunto de datos X que han sido observados y otros datos Y que no han sido observados, y un vector de parámetros θ . El objetivo es encontrar el estimador de máxima verosimilitud de θ dados los datos observados X en una situación en la que la densidad conjunta $p(x, y; \theta)$ es conocida explícitamente, pero la densidad marginal de X , $p(x; \theta)$ se puede encontrar sólo numéricamente.

El algoritmo EM se puede aplicar a problemas en los que se tiene una serie de pares (X, Y) donde X son los datos observables y Y los datos faltantes. Entonces la log-verosimilitud se expresa de la siguiente forma

$$\mathcal{L}(\theta; (X_i)) = \sum_i \log \int p(x_i, y|\theta) dy \quad (\text{B.1})$$

El algoritmo fue propuesto con generalidad por Dempster et al. en 1977 [DRR77].

Sea $Q(\theta, \theta') = E[\log p(X, Y|\theta)|X, \theta']$ que es una función de θ ; la esperanza condicional se calcula sobre los valores de Y , y se evalúa como si θ' fueran los parámetros verdaderos. El algoritmo EM comienza con algunos valores iniciales $\theta^{(0)}$ y se basa en dos pasos fundamentales que se van alternando:

- **E:** Se calcula $Q(\theta, \theta^{(i-1)}) = E[\log p(X, Y|\theta)|X, \theta^{(i-1)}]$
- **M:** Se elige $\theta^{(i)}$ de forma que maximice la función $Q(\theta, \theta^{(i-1)})$.

En cada una de las iteraciones se incrementa la log-verosimilitud $\mathcal{L}(\theta|X) = \log p(X|\theta)$ dado por

$$\mathcal{L}(\theta; X) = \log p(X, Y|\theta) - \log p(Y|X, \theta) \quad (\text{B.2})$$

y se calcula la esperanza utilizando la densidad $p(Y|X, \theta')$ y se obtiene

$$\mathcal{L}(\theta|X) = Q(\theta, \theta') - E[\log p(Y|X, \theta)|X, \theta'] \quad (\text{B.3})$$

Consideremos ahora la esperanza E' con respecto a $p(Y|X, \theta')$ y sea $h(Y) = \frac{p(Y|X, \theta)}{p(Y|X, \theta')}$. Entonces $E' \log h(Y) \leq E' h(Y) - 1 = 0$ puesto que $\log x \leq x - 1$. Así

$$E' \log p(Y|X, \theta) \leq E' \log p(Y|X, \theta'). \quad (\text{B.4})$$

Supongamos ahora que $Q(\theta, \theta') > Q(\theta', \theta')$, por lo que

$$\begin{aligned} \mathcal{L}(\theta; X) &= Q(\theta, \theta') - E' \log p(Y|X, \theta) \\ &> Q(\theta, \theta') - E' \log p(Y|X, \theta') \end{aligned} \quad (\text{B.5})$$

$$= \mathcal{L}(\theta'; X)$$

(B.6)

El incremento de la verosimilitud será positivo en cada iteración del EM dado que $Q(\theta, \theta^{(i-1)}) > Q(\theta^{(i-1)}, \theta^{(i-1)})$, y esto es de este modo a menos que $\theta^{(i-1)}$ sea ya el máximo. Estos argumentos se aplican a lo que se llama algoritmo EM generalizado (GEM), en donde la cantidad $Q(\theta, \theta^{(i-1)})$ no es totalmente maximizada sino que $\theta^{(i)}$ se elige para que simplemente incremente su valor.

Las propiedades de convergencia del algoritmo EM y su variante GEM no se han establecido completamente.

Existe también una versión bayesiana del algoritmo EM como una forma de encontrar las modas *a posteriori* de un subconjunto de parámetros. Escribamos $\theta = (\phi, \psi)$ lo que en el paradigma bayesiano es un vector aleatorio, y supongamos que deseamos

encontrar la moda de la densidad *a posteriori* $p(\Phi|X)$. Si ahora tomamos ψ como los datos ocultos tenemos que,

$$\log p(\phi|X) = \log p(\phi, \psi|X) - \log p(\psi|\phi, X) \quad (\text{B.7})$$

y podemos aplicar los mismos argumentos para tomar $Q(\phi, \phi') = E[\log p(\psi|X)|\phi', X]$ y el algoritmo EM puede ayudar a encontrar el estimador máximo *a posteriori* de ϕ .

B.1.1 Mixturas de densidades

Supongamos que tenemos una densidad de la forma

$$p(\mathbf{x}) = \sum_{\mathbf{i}} w_{\mathbf{i}} f_{\mathbf{i}}(\mathbf{x}|\phi_{\mathbf{i}}) \quad (\text{B.8})$$

donde los parámetros ϕ de las densidades pueden tener componentes comunes (por ejemplo una matriz de covarianza comun en una mixtura gaussiana), y los pesos de las mixturas ($w_{\mathbf{i}}$) toman valores no restringidos (aparte de formar una distribución para la componente I de la mixtura, son no negativos y suman la unidad). Podemos entender la componente I como los datos faltantes. Entonces el paso E del algoritmo se expresa de la siguiente forma

$$Q(\theta, \theta') = \sum_{\mathbf{x} \in F} \sum_{\mathbf{i}} \pi_{\mathbf{i}}(\mathbf{x}) [\log f_{\mathbf{i}}(\mathbf{x}|\phi_{\mathbf{i}}) + \log w_{\mathbf{i}}] \quad (\text{B.9})$$

donde

$$\pi_{\mathbf{i}}(\mathbf{x}) = E[I(I = \mathbf{i})|X = \mathbf{x}, \theta'] = \frac{w_{\mathbf{i}} f_{\mathbf{i}}(\mathbf{x}|\phi_{\mathbf{i}})}{\sum_{\mathbf{j}} w_{\mathbf{j}} f_{\mathbf{j}}(\mathbf{x}|\phi_{\mathbf{j}})}. \quad (\text{B.10})$$

Para maximizar esta función en $w_{\mathbf{i}}$ solamente necesitamos considerar el segundo término de $Q(\theta, \theta')$, por lo que se maximiza $\sum_{\mathbf{i}} \log \pi_{\mathbf{i}} w_{\mathbf{i}}$ donde $\pi_{\mathbf{i}}$ es el promedio de los $\pi_{\mathbf{i}}(\mathbf{x})$ sobre el conjunto de datos de entrenamiento. Así tenemos,

$$\sum_{\mathbf{i}} \pi_{\mathbf{i}} \log w_{\mathbf{i}} / \pi_{\mathbf{i}} \leq \sum_{\mathbf{i}} \pi_{\mathbf{i}} w_{\mathbf{i}} / \pi_{\mathbf{i}} - 1 = 0 \quad (\text{B.11})$$

puesto que $\sum \pi_{\mathbf{i}} \log w_{\mathbf{i}} \leq \sum \pi_{\mathbf{i}} \log \pi_{\mathbf{i}}$ y $\hat{w}_{\mathbf{i}} = \pi_{\mathbf{i}}$. Para los parámetros $\phi_{\mathbf{i}}$ maximizamos la verosimilitud ponderada

$$\sum_{\mathbf{i}} \sum_{\mathbf{x} \in F} \pi_{\mathbf{i}}(\mathbf{x}) \log f_{\mathbf{i}}(\mathbf{x}|\phi_{\mathbf{i}}). \quad (\text{B.12})$$

Si los parámetros no son comunes, cada una de las ϕ puede encontrarse por separado.

Bibliografía

- [AG91] R. G. Aykroyd y P.J. Green. Global and local priors, and the location of lesions using gamma-camera imagery. *Philosophical Transactions of the Royal Society of London, Ser. A*, 337:323–342, 1991.
- [AK96] Y. Amit y A. Kong. Graphical templates for model registration. *IEEE Transactions of Pattern Analysis and machine Intelligence*, 18:225–236, 1996.
- [Bay] K. Bayley. Ophthalmic fluorescein and indocyanine green angiography. <http://www.wrmny.com/angiography.htm>.
- [Bes74] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *J. Royal Statist. Soc. B*, 36:192–225, 1974.
- [Bes75] J.E. Besag. Statistical analysis of non lattice data. *The statistician*, 24:179–195, 1975.
- [Bes86] J. Besag. On the statistical analysis of dirty pictures. *J. Royal Statist. Soc. B*, 48:259–302, 1986.
- [Bez81] J.C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, 1981.
- [Bil89] P. Billingsley. *Probability*. Wiley and Sons, 1989.
- [Boo89] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trnsactions on Pattern analysis and machine intelligence*, 11:pp 567–585, 1989.
- [CC89] S. Chaudhuri y S. Chatterjee. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8(3), 1989.

- [CFG92] Thomas M Clark, William L Freeman, y Michael H Goldbaum. Digital overlay of fluorescein angiograms and fundus images for treatment of subretinal neovascularization. *Retina*, 12(2):118–126, 1992.
- [CH98] A. D. J. Cross y E. R. Hancock. Perspective pose recovery with a dual step em algorithm. *Advances in Neural Information Processing Systems*, Edited by M. Jordan, M. Kearns and S. Solla, MIT Press, 10, 1998.
- [Cid95] A. V. Cideciyan. Registration of ocular fundus images. *IEEE Engineering in Medicine and Biology*, páginas pp 52–58, 1995.
- [Cre93] Noel A.C. Cressie. *Statistics for Spatial Data. Revised Edition*. John Wiley and Sons, New York, 1993.
- [DAS+97] J. Domingo, G. Ayala, A. Simó, E. de Ves, L. Martínez-Costa, y P. Marco. Irregular motion recovery in fluorescein angiograms. *Pattern Recognition Letters*, 18:805–821, 1997.
- [DDA97] M.E. Diaz, J. Domingo, y G. Ayala. A gray-level 2-d feature detector using circular statistics. *Pattern Recognition*, 1997.
- [Dig83] P.J. Diggle. *Statistical Analysis of Spatial Point Patterns*. Academic Press, London, 1983.
- [DRR77] A. P. Dempster, N. M. Rubin, y D. B. Rubin. Maximum-likelihood from incomplete data via the em algorithm. *J. Royal Statistical Soc. Ser. B (methodological)*, 39:1–38, 1977.
- [DS81] N.R. Draper y H. Smith. *Applied Regression Analysis. Second Edition*. John Wiley and Sons, New York, 1981.
- [dVDS+99] E. de Ves, E. Díaz, A. Simó, G. Ayala, y J. Domingo. Robust summary of binary shapes with applications. *International Journal of Computer Vision*, Aceptado para su publicación, 1999.
- [Gey94] C.J. Geyer. On the convergence of monte carlo maximum likelihood calculations. *Journal of the Royal Statistical Society, Series B*, 56:261–274, 1994.

- [GG84] S. Geman y D. Geman. Stochastic relaxation, gibbs distribution and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-6:721–741, 1984.
- [GG87] D. Geman y S. Geman. Relaxation and annealing with constraints. Informe Técnico 35, Complex System Technical Report. Div of Applied Mathematics. Brown university, 1987.
- [GK⁺90] Michael H Goldbaum, Norman P Katz, y others. Digital image processing for ocular fundus images. *Ophthalmology clinics of North America*, 3(3), September 1990.
- [GM87] S. Geman y D.E. McClure. Statistical methods for tomographic image reconstruction. En *Proc. 46th Sess. Inst. Stat. Inst. Bulletin ISI*, volumen 52, 1987.
- [Gon87] R.C. Gonzalez. *Digital Image Processing*. Addison-Wesley, Reading, Massachusetts, 1987.
- [Gre90] P.J. Green. Bayesian reconstruction from emission tomography data using a modified em algorithm. *IEEE Transactions on Medical Imaging*, 9:84–93, 1990.
- [Gre96] P.J. Green. Mcmc in image analysis. En W.R. et al. Gilks, editor, *Markov Random Fields. Theory and Application*, páginas 381–400. 1996.
- [Gro91] Early Treatment Diabetic Retinopathy Study Group. Classification of diabetic retinopathy from fluorescein angiograms: Etdrs report number 11. *Ophthalmology*, 98:807–822, 1991.
- [GT94] C.J. Geyer y E. A. Thompson. Constrained monte carlo maximum likelihood for dependent data, (with discussion). *Journal of the American Statistical Association*, 56:261–274, 1994.
- [Guy95] X. Guyon. *Random fields on a network. Modeling statistics and applications*. Springer-Verlag, 1995.

- [GW87] R.C. Gonzalez y P. Wintz. *Digital Image Processing*. Addison-Wesley, second edición, 1987.
- [HG94] W.E. Hart y M.H. Goldbaum. Registering retinal images using automatically selected control point pairs. *Proceedings IEEE International Conference on Image Processing*, 1994.
- [HS92a] R.M. Haralick y L.G. Shapiro. *Computer and Robot Vision Volume I*. Addison-Wesley, Reading, Massachusetts, 1992.
- [HS92b] R.M. Haralick y L.G. Shapiro. *Computer and Robot Vision Volume II*. Addison-Wesley, Reading, Massachusetts, 1992.
- [IS99] M.V. Ibáñez y A Simó. Bayesian detection of the fovea in eye fundus angiographies. *Pattern Recognition Letters*, 20:229–240, 1999.
- [JAB⁺92] R. Jagoe, J. Arnold, C. Blauth, P.C.L. Smith, K.M. Taylor, y R. Wootton. Measurement of capillary dropout in retinal angiograms by computerised image analysis. *Pattern Recognition Letters*, 13:143–151, 1992.
- [JB⁺90] R. Jagoe, C. Blauth, y others. Automatic geometrical registration of fluorescein retinal angiograms. *Computers and Biomedical Research*, 23:403–409, 1990.
- [JBBL87] Z. Jericevic, D.M. Benson, J. Bryan, y Smith L.C. Geometric correction of digital images using orthonormal decomposition. *Journal of Microscopy*, 3:233–245, 1987.
- [JJ94] M. I. Jordan y R. A. Jacobs. Hierarchical mixtures of experts and the em algorithm. *Neural Computation*, 6:181–214, 1994.
- [Joh94] V.E. Johnson. A model for segmentation and analysis of noisy images. *Journal of the American Statistical Association*, 89:230–241, 1994.
- [Kar92] N. Karssemeijer. Application of bayesian methods to segmentation in medical images. En Frigessi Barone, P. y M. Piccioni, editores, *Lecture Notes in Statistics. Stochastic Models, Statistical Methods, and Algorithms in Image Analysis.*, volumen 74. Springer-Verlag, 1992.



- [KH89] J. Kittler y E. R. Hancock. Combining evidence in probabilistic relaxation. *International of Pattern recognition and Artificial Intelligence*, 3:29–51, 1989.
- [KRD80] D. J. Kahl, A. Rosenfeld, y A. Danker. Some experiments in point pattern matching. *IEEE Transactions on Systems, Man and Cibernetics*, 10:105–1015, 1980.
- [KS96] R. Kutka y S. Stier. Extraction of line properties based on direction fields. *IEEE Transactions on Medical Imaging*, 15(1):51–58, 1996.
- [LS93] I. Liu y Y. Sun. Recursive tracking of vascular networks in angiograms based on the detection-deletion scheme. *IEEE Transactions on Medical Imaging*, 12(2), 1993.
- [LTC93] A. Lanitis, C. J. Taylor, y T. F. Cootes. A generic system for classifying variable objects using flexible template matching. *Proceedings of the British Machine Vision Conference BMVA Press*, páginas 329–338, 1993.
- [Mar72] K. Mardia. *Statistics of Directional Data*. Academic Press, 1972.
- [MC94] L. Martínez-Costa. *Análisis cuantitativo computerizado de las imágenes obtenidas mediante angiografía fluoresceínica*. Tesis Doctoral, Universidad de Valencia, 1994.
- [MCMA96] L. Martínez-Costa, P. Marco, y G. Ayala. Computer-assisted classification of macular edema in diabetic retinopathy, 1996. The Royal College of Ophthalmologists. Annual meeting.
- [MH97a] S. Moss y E. R. Hancock. Multiple line-template matching with the em algorithm. *Pattern Recognition Letter*, 18:1283–1292, 1997.
- [MH97b] S. Moss y E. R. Hancock. Registering incomplete radar images with the em algorithm. *Image and Vision Computing*, 15:637–648, 1997.
- [MMB⁺94] J. Meunier, G. Mimoun, P. Bunel, G. Quentel, G. Soubrane, y G. Coscas. Automated quantification of diabetic macula non-perfusion on digitized angiography. *Invest Ophthalmol Vis Sci*, 35:1595, 1994.

- [MNSM89] Goldbaum MH, Katz NP, Chaudhuri S, y Nelson M. Image understanding for automated retinal diagnosis. En Kingsland L, editor, *Proceedings of the IEEE for the 13TH Annual Symposium for Computer Application in Clinical Medicine*. Computer Soc Press, páginas 756–760, 1989.
- [Mol94] R. Molina. On the hierarchical bayesian approach to image restoration: Applications to astronomical images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 16:1122–1128, 1994.
- [Mol96] A. Mollie. Bayesian mapping of disease. En W.R. Gilks, S. Richardson, y S. Spiegelhalter, editores, *Markov chain Monte Carlo in practice*, capítulo 20. Chapman and Hall, 1996.
- [MSea91] G. Mimoun, G. Soubrane, y et al. Clinical experience with digitalized fluorescein angiography. *International Ophthalmol*, 15:41–46, 1991.
- [NA61] HR. Novtny y DL. Alvis. A method of photographing fluorescence in circulating blood in the human retina. *Circulation*, 24, 1961.
- [NSR85] P. Nagin, B. Schwartz, y G. Reynolds. Measurement of fluorescein angiograms of the optic disc and retina using computerized image analysis. *Ophthalmology*, 4(92):547–552, 1985.
- [oM] National Library of Medicine. The visible human project. <http://www.nlm.nih.gov/research/visible/visiblehuman.html>.
- [PL86] E. Peli y M. Lahar. Drusen measurements from fundus photographs using computer image analysis. *Ophthalmology*, 93:1575–1580, 1986.
- [PPJ+99] J.C Pérez, R. Paredes, Valiente J.M., J. Arlandis, y J. Cano. An elastic registration method for quality control of textile printing. *VIII National Symposium on Pattern Recognition and Image Analysis*, 1:pp 379–394, 1999.
- [PRT+91] R.P. Phillips, P.G. Ross, M. Tyska, P.F. Sharp, y J.V. Forrester. Detection and quantification of hyperfluorescent leakage by computer analysis of fundus fluorescein angiograms. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 229:329–335, 1991.

- [PSR⁺91] R. Phillips, T. Spencer, P. Ross, P. Sharp, y J. Forrester. Quantification of diabetic maculopathy by digital imaging of the fundus. *Eye*, 5:130–137, 1991.
- [PTVB92] W. Press, S. Teukolsky, W. Vetterling, y Flannery B. *Numerical Recipes in C*. Cambridge University Press, second edition, 1992.
- [QT93] W. Qian y D. M. Titterington. Bayesian image restoration: An application to edge-preserving surface recovery. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 15:748–752, 1993.
- [Rea89] S.J. Ryan y et al. *Retina*. The C.V. Mosby Company, 1989.
- [RJ⁺93] Jagoe R, Arnold J, y others. Retinal vessel circulation patterns visualized from a sequence of computed-aligned angiograms. *Invest Ophthalmol Vis Sci*, 34:2881–2887, 1993.
- [RMP⁺96] A. Rangarajan, E. Mjolsness, S. Pappu, L. Davachi, P.S. Goldman-Rakic, y J.S. Duncan. A robust point matching algorithm for autoradiograph alignment. En *Proc. Visualization in Medical Imaging VBC'96*, volumen 1131 de *Lectures Notes in Computer Science*, páginas 277–286. Springer, 1996.
- [RWK⁺97] A. Remky, S. Wolf, H. Knabben, O. Arend, y M. Reim. Perifoveal capillary network in patients with acute central retinal vein occlusion. *Ophthalmology*, 104:33–37, 1997.
- [Sai] P. J. Saine. Fundus photography. <http://webey.ophth.uiowa.edu/ops/op-photo/fundus/funphot1.htm>.
- [Sch89] Robert J. Schalkoff. *Digital Image Processing and Computer Vision*. John Wilwy and sons, 1989.
- [SdVD⁺99] A Simó, E. de Ves, E. Díaz, G. Ayala, y J. Domingo. Bayesian estimation of edge orientations in junction. *Pattern Recognition, Aceptado para su publicación*, 1999.

- [Ser82] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, 1982.
- [Sha95] L.S. Shapiro. *Affine Analysis of Image Sequences*. Cambridge University Press, 1995.
- [Sha97] P.F. Sharp. Digital imaging in ophthalmology. fact or fiction?, 1997. The Royal College of Ophthalmologists. Annual meeting.
- [Sil86] B.W. Silverman. *Density Estimation*. Chapman and Hall, New York, 1986.
- [SKB80] G. C. Stockman, S. Kopstein, y S. Bennet. Matching images to models for registration and object detection via clustering. *IEEE Transactions on Systems, Man and Cibernetics*, 10:105–1015, 1980.
- [SOMH+96] T. Spencer, J.A. Olson, K.C. Mc Hardy, P.F. Sharp, y J.V. Forrester. An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. *Computers and Biomedical Research*, 29:284–302, 1996.
- [SP95] S. Sclaroff y A. P. Pentland. Modal matching for correspondence and recognition. *IEEE transactions on Pattern Analysis and Machine Intelligence No. 6*, 17:545–561, 1995.
- [SPSF92] T. Spencer, R. Phillips, P. Sharp, y J. Forrester. Automated detection and quantification of microaneurysms in fluorescein angiograms. *Graefe's Arch Clin Exp Ophthalmol*, 230:36–41, 1992.
- [SR92] P. Nagin, B. Schwartz, y G. Reynolds. Measurement of fluorescein angiograms of the optic disc and retina using computerized image analysis. *Ophthalmology*, 4(pp. 547-552), 1992.
- [Tel92] M. Teles. Bayesian estimation of ventricular contours in angiographic images. *IEEE Transaction on Medical Imaging*, 11(3):416–429, 1992.
- [TNSW87] A. Tuulonen, P. Nagin, B. Schwartz, y D-C. Wu. Increase in pallor and fluorescein-filling defect of the optic disc in the follow-up of ocular hyper-

- tensives measured by computerized image analysis. *Ophthalmology*, 94:558–563, 1987.
- [TPV98] Yannis A. Tolias y Stavros M. Panas. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Transactions on Medical Imaging*, 17(2), 1998.
- [TPV92] A. Todd-Pokropek y M. Viergever, editores. *Medical Images: Formation, Handling and Evaluation*, Berlin, 1992. Springer Verlag.
- [TV90] D. Tweed y T. Vilis. Geometric relations of eye position and velocity vectors during saccades. *Vision Research*, 30:pp 111–127, 1990.
- [Ume88] S. Umeyama. An eigendecomposition approach to weighted graph matching problems. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, No. 5, 10:695–703, 1988.
- [Wat62] G Watson. Goodness of fit tests on a circle. *Biometrika*, 49:57–63, 1962.
- [WB92] H. Wang y J.M. Brady. Corner detection: some new results. En *IEEE Colloquium Digest on Systems Aspects of Machine Perception and Vision*, páginas 1.1–1.4, 1992.
- [Wei97] I.S. Weir. Fully bayesian reconstruction from single photon emission computed tomography data. *Journal of the American Statistical Association*, 92(437):49–60, 1997.
- [WG94] I.S. Weir y P.J. Green. Modeling data from single-photon emission computed tomography. En K.V. Mardia, editor, *Statistics and Images*, páginas 313–338. 1994.
- [WH93] R.C Wilson y E.R. Hancock. Relaxation matching of road networks in aerial images using topological constraints. *Proceedings of SPIE Sensor Fusion VI.*, páginas 444–455, 1993.
- [WH96] R.C. Wilson y E. R. Hancock. A bayesian compatibility model for graph matching. *Pattern Recognition Letters*, 17(pp. 263-276), 1996.

- [Win96] G. Winkler. *Image Analysis, Random fields Models and Dynamic Monte Carlo Methods*. Springer-Verlag, 1996.
- [Wol90] G. Wolberg. *Digital Image Warping*. IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [WTT89] N.P. Ward, S. Tomlinson, y C.J. Taylor. The detection and measurement of exudate associated with diabetic retinopathy. *Ophthalmology*, 96:80–86, 1989.
- [You88] L. Younes. Estimation and annealing for gibbsian fields. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 24:269–294, 1988.
- [ZHRB96] X. Zhang, R. M. Haralick, V. Ramesh, y A.S. Bedekar. A bayesian corner detector: Theory and performance evaluation. Informe técnico, Intelligent Systems Laboratory. Department of Electrical Engineering, FT-10. University of Washington, Seattle, WA 98195, 1996.
- [ZJ97] F Zana y Klein J.C. Robust segmentation of vessels from retinal angiography. En *DSP97*, páginas 1087–1090, 1997.
- [ZMJ98] F Zana, I Meunier, y Klein J.C. A region merging algorithm using mathematical morphology: application to macula detection. En *ISMM98*, páginas 423–430, 1998.

UNIVERSITAT DE VALÈNCIA

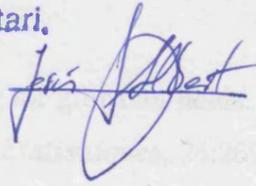
FACULTAD DE CIÈNCIES FÍSQUES

Reunit el Tribunal que subscriu, en el dia de la data,
acordà d'atorgar, per unanimitat, a aquesta Tesi Doctoral
d'En/ Na/ N' Esther de Ves Cuenca

la qualificació d' Sobresaliente "cum laude"

València a 17 d' Septiembre de 1979

El Secretari,



El President,

