



UNIVERSIDADE DA CORUÑA

Facultade de Informática
Departamento de Computación

**Cálculo de variables hemodinámicas a
partir del alineamiento de secuencias de
imágenes de SLO**

Doctorando: Cástor Mariño Pérez

Directores: Manuel F. González Penedo

María José Carreira Nouche

A Coruña, 16 de junio de 2006

Dña. María José Carreira Nouche, Profesora Titular de Universidad del Departamento de Electrónica y Computación de la Universidade de Santiago de Compostela

y

D. Manuel Fco. González Penedo, Profesor Titular de Universidad del Departamento de Computación de la Universidade da Coruña.

CERTIFICAN que:

La memoria que se presenta, titulada *Cálculo de variables hemodinámicas a partir del alineamiento de secuencias de imágenes de SLO* ha sido realizada por Dn. Cástor Mariño Pérez bajo nuestra dirección en el Departamento de Computación de la Universidad de A Coruña y constituye la Tesis que presenta para optar al grado de Doctor.

Y para que así conste firmamos la presente en A Coruña, a 16 de junio de 2006.

Fdo.: María José Carreira Nouche

Fdo.: Manuel Fco. González Penedo

Agradecimientos

Tras un largo camino, al fin este trabajo ve la luz. Han sido muchas horas de dedicación, dolores de cabeza y esfuerzo. Pero también muchos buenos momentos inolvidables, gracias sobre todo a un grupo de personas que han recorrido la ruta conmigo. Quisiera expresar a todos ellos mi más sincero agradecimiento por su ayuda, por su paciencia, y sobre todo por su compañía.

A mis directores María José Carreira Nouche y a Manuel F. González Penedo, por su infinita paciencia e inestimable apoyo. Sin su dirección este barco nunca habría llegado a puerto.

A los miembros del Centro de Visión por Computador de la Universidad Autónoma de Barcelona David Lloret, Joan Serrat, Antonio López y Juanjo Villanueva, por su ayuda y apoyo.

A los miembros del Departamento de Computación de la Universidad de A Coruña y del laboratorio de Fundamentos de Computación e Inteligencia Artificial, especialmente a Víctor Gulías, por su atención y palabras de ánimo, además de sus consejos.

A todos los compañeros del grupo de Visión Artificial y Reconocimiento de patrones, especialmente a Marta y Noelia, por sus aportaciones en los momentos de duda y grata compañía en los de ocio.

A Antonio Benito, que me hizo ver como con esfuerzo y constancia podría llegar hasta donde me lo propusiera.

Y para terminar, a todos mis amigos de la Consellería de Medio Ambiente de

la Xunta de Galicia, Pilar, Alberto, Juan, Nieves, Jaime, Carmen, Marta, ... , a mi familia, y a Ana, cuyo tiempo robado ha permitido fraguar este trabajo.

A M.Fé, a Rafael, a Xe e a Flori, en resumo, á miña familia

A Ana

Prólogo

En esta tesis se aborda el alineamiento automático de imágenes médicas. Cuando se dispone de imágenes de un paciente obtenidas en distintos instantes de tiempo o capturadas utilizando distintos métodos (resonancia magnética, tomografía computarizada, ecografía, etc.) suele ser necesario alinear todas las imágenes en un mismo eje de referencia para así poder extraer medidas a partir de las mismas. Llevar a cabo este proceso de forma manual es una tarea tediosa y propensa a errores debido a la baja calidad de imagen que proporcionan algunos de estos métodos de adquisición. Es por esto que los algoritmos de alineado de imágenes son una herramienta imprescindible en determinados campos de acción.

El sistema que hemos desarrollado permite alinear secuencias completas de *SLO*¹ por medio de la combinación de dos métodos distintos de ajuste, uno de ellos basado en la extracción de características, y otro basado en la intensidad de la imagen.

Este trabajo surgió como respuesta a la necesidad de los oftalmólogos del Complejo Hospitalario Universitario de Santiago (CHUS) de una herramienta que les permita llevar a cabo el alineamiento de secuencias de imágenes que obtienen con un *Oftalmoscopio de Barrido Láser (SLO)*. Este dispositivo, que describiremos en su momento, permite la captura de secuencias de imágenes, habitualmente compuestas por entre 3000 y 4000 fotogramas. Sin embargo, resulta imposible que el paciente mantenga fijo el ojo durante el proceso de captura de toda la secuencia (unos 35-40 segundos) por lo que, para que sea posible analizar el contenido de dicha secuencia, es necesario ajustar todos los fotogramas a un mismo eje de coordenadas. Para esto se utilizan los denominados *algoritmos de alineado de imágenes*.

¹Scanning Laser Ophthalmoscope

Para cubrir esta necesidad se comienza el desarrollo de una aplicación que permita de forma cómoda, rápida y fiable, alinear los fotogramas de una secuencia y extraer de la misma determinadas medidas hemodinámicas que permitan evaluar la evolución de un paciente ante un determinado tratamiento².

Para ello se desarrolló una técnica, en colaboración con el *Centre de Visió per Computador* de la Universidad Autónoma de Barcelona, basada en la extracción de marcadores, que alinea los fotogramas. Esta técnica basa su funcionamiento en la segmentación de unas estructuras denominadas *curvas cresta y curvas valle* [López Peña y col., 1999], que se utilizarán como referencia en el proceso de ajuste.

Sin embargo, para algunas zonas de las secuencias en las que el contraste es muy bajo, el método de extracción de características falla, ya que en estas imágenes prácticamente no existen características que extraer, por lo que resulta necesario aplicar un método alternativo de alineamiento. Por ello se recurrió a una técnica de alineamiento basada en el cálculo de una propiedad de las imágenes, la *Información Mutua* [Vajda, 1989], que se obtiene a partir de los valores de la intensidad de los píxeles.

Una vez obtenidos dichos métodos y ajustados a nuestro dominio de aplicación, se combinaron convenientemente para que el proceso fuese totalmente automático, y por último se realizó su validación en un entorno hospitalario.

La memoria se encuentra dividida en varios capítulos: en el capítulo 1 realizaremos una introducción al problema del alineamiento de imágenes, así como una revisión de las distintas técnicas que se pueden encontrar en la literatura sobre el tema. En el capítulo 2 nos centraremos en la primera de las técnicas de alineado que se han desarrollado para ajustar las secuencias de *SLO*: el alineado basado en curvas cresta y curvas valle. Al final de este apartado veremos cómo para aquellas imágenes en las que no existen características señalables, esta técnica no es capaz de llevar a cabo el ajuste de manera precisa, por lo que surge la necesidad de utilizar otro método. El capítulo 3 presenta el ajuste basado en el cálculo de una medida estadística denominada *Información Mutua*, con una descripción del cálculo de la misma y cómo se utiliza para alinear los fotogramas. En este punto tendremos ya alineada la secuencia

²Este trabajo se desarrolló dentro del proyecto financiado con fondos FEDER cuyo título es *Implementación de un interfaz y una base de datos para sistemas de interpretación automática de imágenes de angiografías retinianas*

completa, pero con zonas discontinuas, debido a la utilización de los dos métodos. En el capítulo 4 se detalla la metodología utilizada para combinar las dos técnicas en un algoritmo que permite ajustar secuencias completas de *SLO* de manera automática. Finalmente en el capítulo 5 se exponen los resultados obtenidos con nuestro método, y en el capítulo 6 se enumeran las conclusiones extraídas del desarrollo de este trabajo así como las líneas futuras de investigación.

Índice general

1. Introducción	1
1.1. Introducción al alineamiento de imágenes	5
1.1.1. Espacio de características	8
1.1.2. Espacio de búsqueda	9
1.1.3. Estrategia de búsqueda	12
1.1.4. Medida de similitud	14
1.2. Alineamiento de imágenes	16
1.2.1. Técnicas basadas en la extracción de características	17
1.2.2. Técnicas basadas en la la intensidad de las imágenes	20
2. Alineamiento de imágenes de <i>SLO</i>	27
2.1. Registro basado en extracción de características	31
2.1.1. Método de extracción de crestas y valles	32
2.1.2. Alineamiento de las imágenes	39
2.1.3. Evaluación del método de alineamiento basado en crestas	56
3. Método de alineamiento basado en Información Mutua	63

3.1. Estimación basada en el método <i>de la ventana de Parzen</i>	64
3.2. Estimación basada en frecuencias	68
3.3. Algoritmo de Maximización de la <i>información mutua</i>	73
3.3.1. Comparativa de los métodos de cálculo de la <i>información mutua</i>	75
4. Metodología de alineamiento de secuencias	85
4.1. Selección de imágenes de referencia	85
5. Análisis y evaluación de resultados	97
5.1. Validación de las técnicas alineamiento	98
5.1.1. Experimento 1	99
5.1.2. Experimento 2	101
5.1.3. Experimento 3	102
5.2. Validación clínica	105
5.2.1. Cálculo de las curvas de dilución	106
5.2.2. Validación de los resultados obtenidos	115
6. Conclusiones y líneas futuras de investigación	123
A. Fundamentos de teoría de la información	131
A.1. Entropía	131
A.2. Entropía condicional, entropía conjunta e información mutua	132
A.3. Información mutua	133
B. Modelización de transformaciones afines	135

B.1. Coordenadas homogéneas	135
B.2. Transformaciones rígidas	137
B.2.1. Componentes de una transformación afín	139
C. Algoritmos de optimización	143
C.1. Downhill Simplex	143
C.2. Temple Simulado	146
D. Publicaciones seleccionadas	153

Índice de Figuras

1.1. Esquema del ojo con sus diferentes partes, entre ellas la retina, objeto de nuestro estudio.	2
1.2. Adquisición de imágenes de fondo de ojo	3
1.3. Evolución del contraste inyectado a un paciente en una secuencia <i>SLO</i>	4
1.4. Esquema general del proceso de alineamiento	7
1.5. Representación de los distintos tipos de transformaciones	11
1.6. Simplex en 3D y tipos de transformaciones.	13
1.7. Esquema del proceso de alineamiento basado en la extracción de características	18
2.1. Esquema de funcionamiento del dispositivo de adquisición de secuencias <i>SLO</i>	28
2.2. Fotografía de un equipo de <i>SLO</i> y un fotograma de un paciente	29
2.3. Estimación del nivel de ruido de fondo en imágenes de secuencias de <i>SLO</i>	30
2.4. Esquema de la obtención de las crestas y valles en un fotograma <i>SLO</i>	33
2.5. Interpretación gráfica de los términos relacionados con las curvas de nivel	34
2.6. Geometría implicada en la definición de la divergencia del campo vectorial	36
2.7. Proceso de obtención de las crestas y valles sobre dos imágenes extraídas de dos secuencias <i>SLO</i>	40
2.8. Representación del proceso de alineamiento basado en la extracción de crestas.	42
2.9. Representación del proceso de generación de las imágenes que constituyen los distintos niveles de la pirámide multiresolución.	43

2.10. Esquema que muestra la técnica empleada para la búsqueda de semillas iniciales en el nivel superior de la pirámide multiresolución	45
2.11. Muestra de las imágenes en los niveles superiores de la pirámide con 6 niveles.	47
2.12. Esquema de obtención de los valores que optimizan la correlación a través de los niveles de la pirámide	52
2.13. Esquema de obtención de los valores que optimizan la correlación en el nivel 0 de la pirámide, incluyendo el proceso de refinamiento de la solución	53
2.14. Imágenes de crestas y valores de correlación obtenidos con las mismas	54
2.15. Las secuencias <i>SLO</i> se dividen en tres secciones, caracterizadas por el nivel de fluoresceína presente en el ojo en cada una de ellas	57
2.16. Ejemplo de fotogramas y las crestas extraídas de los mismos en función de la sección a la que pertenecen	58
3.1. Esquema del alineamiento empleando <i>información mutua</i> (IM) como medida del ajuste.	65
3.2. Estimación de una gaussiana de media 0 y varianza 1 a partir de 100 muestras (líneas verticales negras) mediante una función de densidad ventana de Parzen.	66
3.3. Cálculo del histograma conjunto de dos imágenes	69
3.4. Área de superposición empleada en el cálculo del histograma conjunto	71
3.5. Ejemplo de mapa de la <i>información mutua</i>	72
3.6. Influencia del muestreo sobre el valor de la <i>información mutua</i>	74
3.7. Influencia del número de muestras en el tiempo de cálculo de la infor- mación mutua	78
3.8. Análisis de los métodos de maximización de la <i>información mutua</i> . .	81
3.9. Fotogramas del paciente C	82
4.1. Fotogramas de referencia F_1 y F_2 seleccionados por el oftalmólogo que se usarán para alinear la secuencia de <i>SLO</i>	86
4.2. Ejemplos de fotogramas con errores, y que no podrán ser utilizados como fotogramas de referencia.	87

4.3. Representación del método de alineamiento basado en <i>información mutua</i> y crestas.	88
4.4. Imágenes para el cálculo de los valores de correlación	92
4.5. Evolución del nivel de gris a lo largo de una secuencia <i>SLO</i>	93
4.6. Evolución del nivel de gris a lo largo de una secuencia <i>SLO</i> registrada.	94
4.7. Esquema de construcción de las imágenes 3D para la evaluación visual del resultado del alineamiento	95
4.8. Imágenes con los fotogramas de la secuencia apilados	96
5.1. Resultados del primer experimento de alineamiento	101
5.2. Resultados del segundo experimento de alineamiento	103
5.3. Resultados del tercer experimento de alineamiento	105
5.4. Representación de la variabilidad de las muestras tomadas en un punto de una secuencia a lo largo del tiempo.	107
5.5. Ejemplo de fotogramas no válidos.	109
5.6. Ilustración del proceso de ajuste para estimar la curva de dilución a partir de la distribución de puntos muestreados.	113
5.7. Representación de los valores atípicos eliminados.	114
5.8. Comparativa de las nubes de puntos obtenidas manual y automáticamente.	116
5.9. Comparativa de las nubes de puntos obtenidas manual y automáticamente.	117
5.10. Ilustración del método de cálculo del tiempo arteria-vena.	118
5.11. Curvas de dilución obtenidas por los oftalmólogos y automáticamente, para cinco secuencias <i>SLO</i>	120
A.1. Esquematización de la relación entre entropía e <i>información mutua</i> utilizando diagramas de Venn.	134
B.1. Representación de la transformación en 2D del punto <i>P</i> en el punto <i>Q</i> mediante la función <i>T</i>	138
B.2. Representación de la composición de dos transformaciones.	142
C.1. Operación de reflexión del simplex en 2D.	144
C.2. Operaciones de reflexión y expansión del simplex en 2D.	145

C.3. Operación de contracción en un vértice del simplex en 2D.	145
C.4. Contracción total del simplex en 2D.	146
C.5. Forma de la distribución de Cauchy.	150

Índice de Tablas

2.1. Tiempos de registro obtenidos al variar el número de niveles de la pirámide multiresolución	46
2.3. Evolución del proceso de alineamiento en el método basado en crestas	55
2.4. Estadísticas del experimento realizado para la evaluación de la alineación basada en crestas y valles para cinco secuencias de <i>SLO</i>	59
3.1. Análisis de los tiempos medios de cálculo de la función de densidad de probabilidad para cinco pacientes distintos empleando el método de estimación basado en el histograma de la imagen.	79
3.2. Análisis de los métodos de maximización de la <i>información mutua</i> para cinco pacientes distintos.	80
3.3. Resultados del alineamiento de fotogramas de la sección 1 utilizando el método basado en <i>información mutua</i>	83
4.1. Resultados de la correlación para imágenes de diferentes secuencias . .	90
5.1. Rango de valores que puede alcanzar cada parámetro de una transformación afín para las imágenes de las secuencias <i>SLO</i>	98
5.2. Resultados del experimento 1: alineamiento de transformaciones con componente única de traslación.	100
5.3. Resultados del experimento 2: alineamiento de transformaciones con componente única de rotación.	102
5.4. Resultados del experimento 3: alineamiento de transformaciones con traslación y rotación.	104
5.5. Valores iniciales para los parámetros de la ecuación 5.2.	112

5.6. Comparativa del porcentaje de fotogramas alineados correctamente antes y después de la aplicación del proceso de eliminación de puntos atípicos.	115
5.7. Tiempos arteria-vena (en segundos) obtenidos utilizando el método manual y el método automático para las cinco secuencias analizadas. .	118
6.1. Parámetros del método de alineado basado en el cálculo de líneas cresta y líneas valle.	126

Capítulo 1

Introducción

En la actualidad los grandes avances alcanzados en las ciencias relacionadas con la informática, tanto a nivel físico (hardware) como a nivel lógico (software) han contribuido al avance de otras ciencias, como por ejemplo la medicina. En concreto, los nuevos instrumentos de adquisición de datos unidos a las nuevas técnicas de procesamiento de los mismos permiten el estudio, tratamiento y evaluación de enfermedades con unos tiempos de respuesta muy reducidos, hecho fundamental para que los diagnósticos sean lo más rápidos y eficientes posibles.

La retinopatía diabética es una de las enfermedades oculares más graves, ya que es la principal causa de pérdida de visión y ceguera en la población activa. Esta pérdida de visión proviene de la neovascularización de la retina¹ (figuras 1.1 y 1.2) o parte de la aparición de edemas en la mácula (área central de la retina encargada de la visión fina, figura 1.2(b)). Algunas anomalías en el riego sanguíneo de la retina aparecen previamente a la neovascularización de la misma, por lo que es importante cuantificar estas anomalías con el fin de prevenir la pérdida de visión.

La prevalencia² de la diabetes es de un 4% de la población española, (aproximadamente 1.6 millones de personas) [Bonafonte y García, 1996]. Existen dos tipos

¹Formación anormal de nuevos vasos sanguíneos, lo que ocurre con frecuencia en la retinopatía diabética proliferativa, oclusiones venosas de la retina, etc.

²proporción de individuos de una población que presentan el evento en un momento o periodo de tiempo determinado.

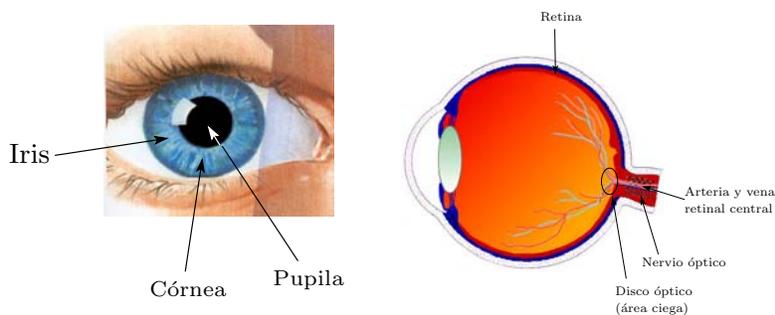


Figura 1.1: Esquema del ojo con sus diferentes partes, entre ellas la retina, objeto de nuestro estudio.

de diabetes: la diabetes Tipo I, que se denomina diabetes insulino-dependiente, también denominada diabetes juvenil, y la diabetes Tipo II, conocida como diabetes no insulino-dependiente y que también se denomina diabetes del adulto, siendo esta segunda menos grave, aunque también puede necesitar en algunos casos insulina.

Aproximadamente el 98 % de los diabéticos juveniles y el 78 % de los adultos presentan retinopatía diabética en los primeros 15 años posteriores al diagnóstico de la diabetes, quedando ciegos a los 20 años alrededor de un 10 % de ambos grupos.

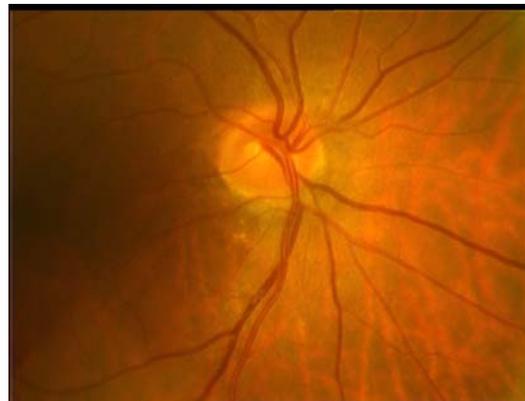
El riesgo de ceguera en personas diabéticas es 25 veces superior al resto de la población, por lo que es primordial el desarrollo de técnicas automáticas para la detección de esta enfermedad en etapas tempranas.

Existen muchos métodos para estudiar la retina: velocimetría Doppler, angiografía con fluoresceína, videoangiografía con fluoresceína o ultrasonografía Doppler. Cada una de estas técnicas permite obtener imágenes con unas determinadas características que permiten el estudio de las diferentes patologías.

Sin embargo, el método más habitual para el estudio de las enfermedades del ojo es la angiografía de retina (figura 1.2(b)) [Pinz y col., 1998, Zana y Klein, 1999, Stewart y col., 2003, Laliberté y col., 2003], que consiste en tomar una fotografía del fondo de la retina tras haber inyectado al paciente un contraste que resalta los vasos arterio-venosos.



(a)



(b)

Figura 1.2: (a) Ejemplo del proceso de captura, digitalización y almacenamiento en un ordenador de una secuencia SLO o de una angiografía. (b) Ejemplo de una angiografía de retina.

Aunque esta herramienta resulta muy útil para el examen ocular, se han desarrollado nuevas técnicas para la adquisición de dichos datos. El *Oftalmoscopio con Láser de Barrido* (*Scanning Laser Ophthalmoscope, SLO*) es una modalidad de captura de imágenes relativamente nueva y que permite el estudio cuantitativo de la circulación en la retina [Axer-Siegel y col., 2002, Fujii y col., 2003, Aitken, 2003]. Para ello se inyecta un contraste en el paciente y se captura una secuencia en la que se visualiza cómo este contraste llega a los vasos de la retina (figura 1.3). El *SLO* proyecta un rayo láser y un diodo láser infrarrojo simultáneamente en el fondo del ojo, con un tamaño de imagen de 33×21 grados. Esta modalidad permite capturar secuencias de imágenes del fondo del ojo tan grandes como se desee (sólo limitadas por la capacidad de almacenamiento del dispositivo que se destine a tal fin, por ejemplo un disco duro), con buena calidad.

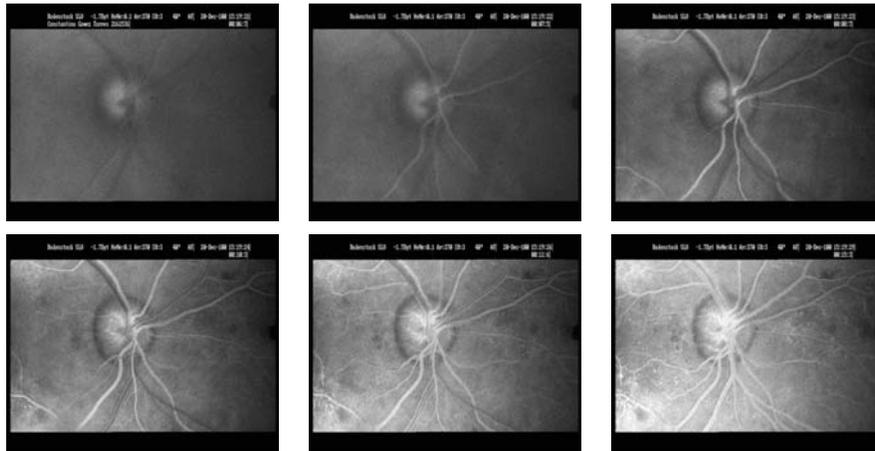


Figura 1.3: Evolución del contraste inyectado a un paciente en una secuencia SLO. En esta figura se muestran 6 fotogramas pertenecientes a distintos instantes de la captura de una secuencia en los que se aprecia claramente cómo el contraste avanza por las arterias y venas de la retina.

El problema de este método es que durante la fase de adquisición de la secuencia de vídeo, y debido al largo período de tiempo que el paciente está expuesto a la luz láser, se producirá inevitablemente algún movimiento del ojo. Estos movimientos son de una velocidad muy elevada (la velocidad máxima de rotación del ojo es de unos $600^\circ/s$ [Markov y col., 1993]), dando lugar a que fotogramas consecutivos en el vídeo se encuentren desalineados, con valores de transformación elevados. Por ello es

necesario un paso previo de alineamiento automático de los fotogramas antes de llevar a cabo cualquier tipo de análisis cuantitativo. Existen muchas técnicas de alineamiento (o registro) de imágenes, de las cuales citaremos las más relevantes en una breve revisión bibliográfica que llevaremos a cabo en la sección 1.2.

1.1. Introducción al alineamiento de imágenes

En esta sección realizaremos una introducción a los métodos de alineamiento de imágenes. Para ello definiremos una serie de conceptos que si bien aquí utilizaremos en el contexto de las imágenes médicas, pueden ser extrapolados a métodos de análisis de imagen en general.

Como ya hemos señalado, los métodos de alineamiento de imagen buscan poner en correspondencia imágenes de un mismo objeto cuyos contenidos no coinciden. Estas diferencias pueden deberse a que:

- las imágenes proceden de diferentes métodos de adquisición, por ejemplo imágenes de *Tomografía Computerizada* e imágenes de *Resonancia Magnética*.
- alguna parte del objeto se ha desplazado, como ocurre cuando un paciente mira a un punto distinto en dos instantes de tiempo. En este caso el nervio óptico aparecerá en dos sitios distintos en cada una de las imágenes.
- el objeto en sí ha cambiado, por ejemplo en el caso de la retina, la neovascularización provoca la visualización de nuevos vasos.
- alguna parte del objeto ha cambiado, como es el caso de los ventrículos o aurículas del corazón, según se analicen en la fase de diástole o en la de sístole.

Para llevar a cabo este ajuste, los métodos de alineamiento buscan la transformación geométrica que alinea los contenidos de las dos imágenes, a las que denominaremos *imagen de referencia* e *imagen flotante*, siendo esta última la que se busca poner en correspondencia con la primera. En el caso del alineamiento de secuencias se procede de dos en dos, por lo que estos términos siguen siendo aplicables.

Dadas dos imágenes en el plano bidimensional I_1 (imagen de referencia) e I_2 (imagen flotante), su alineamiento puede expresarse matemáticamente como una función que pone en correspondencia los píxeles de una imagen con los de la otra:

$$I_2(x, y) = T(I_1(F(x, y))) \quad (1.1)$$

donde la función F representa una transformación espacial, es decir, mapea las coordenadas (x, y) en otras coordenadas (x', y') :

$$(x', y') = F(x, y) \quad (1.2)$$

y T es la función que combina los cambios en la escena y variaciones de la cámara con posibles factores exógenos que puedan afectar a la imagen (cambios en intensidad, contraste, etc).

Por ello, a la hora de resolver un problema utilizando un algoritmo de alineamiento de imágenes deberemos considerar cuatro factores básicos [Brown, 1992]:

1. *Espacio de características*, que representa la información de las imágenes que se utilizará para el ajuste (por ejemplo, los vasos de la retina).
2. *Espacio de búsqueda*, que recoge el conjunto de transformaciones que se aplicarán a una imagen para corregir las diferencias existentes.
3. *Estrategia de búsqueda*, que define el algoritmo que se utilizará para obtener la transformación óptima que ponga en correspondencia las dos imágenes.
4. Finalmente, *medida de similitud o medida del ajuste*, M , que es una función que depende de las transformaciones aplicadas a la imagen flotante, y representa la bondad de las mismas.

Así, se puede definir el proceso de alineamiento como aquel que maximiza la medida de ajuste entre dos imágenes mediante la estrategia de búsqueda:

$$\text{Arg máx}_{T, F} M(I_1(x, y), I_2(x, y), T, F) = \hat{M} \quad (1.3)$$

donde I_1 e I_2 son las dos imágenes que se desea alinear, F representa la transformación espacial aplicada, T los cambios de escena y variaciones en la cámara con posibles factores exógenos y finalmente M denota la medida del ajuste entre las imágenes. La medida \hat{M} representa el valor óptimo (máximo o mínimo) de la medida de ajuste tras el proceso de búsqueda.

La figura 1.4 muestra el procedimiento general de una técnica de alineamiento. El procedimiento general es un proceso iterativo en el que se aplica una transformación a una de las imágenes (la denominada imagen dinámica o flotante). A continuación se calcula la medida del ajuste entre esta imagen transformada y la imagen original, y si se ha alcanzado el óptimo de esta medida, el proceso ha llegado a su fin. De lo contrario, se aplicará una nueva transformación del espacio de búsqueda considerado para tratar de mejorar el valor de la medida de ajuste.

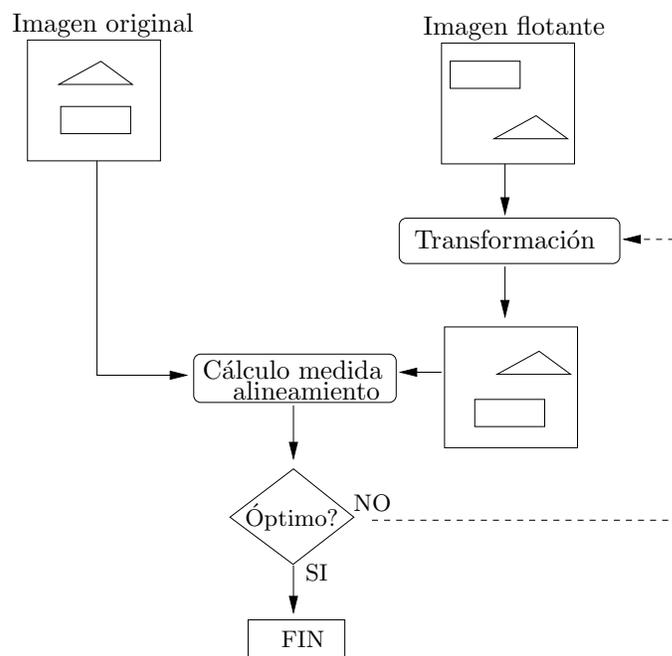


Figura 1.4: Esquema general del proceso de alineamiento. Se trata de un proceso iterativo en el cual se aplica una transformación a una de las imágenes para tratar de optimizar el valor de la función de ajuste considerada, que mide la similitud entre esta imagen y la de referencia.

A continuación veremos las distintas opciones existentes dentro de cada una de estas componentes.

1.1.1. Espacio de características

El espacio de características es el factor del proceso de ajuste que deberá ser considerado en primer lugar, y que además condicionará la elección de las demás componentes que se emplearán en el proceso de alineamiento. Las características más empleadas son:

- **Puntos de control:** marcan zonas características y son fácilmente identificables dentro de la imagen. Estos puntos se suelen seleccionar manualmente, lo cual resulta inviable para conjuntos de imágenes con un número elevado de elementos [Moseley y Munro, 1994, Savi y col., 1995, Fang y col., 1996]. Debido a esta limitación, han surgido algunas aproximaciones en las que la selección de los puntos de control es automática [Kanal y col., 1981, Hart y Goldbaum, 1994, Becker y col., 1998, Rohr, 1999].
- **Primitivas de la imagen:** la extracción de primitivas de la imagen o marcadores (*landmarks*) es una de las opciones más empleadas, ya que se reduce enormemente la cantidad de información a procesar y consecuentemente el tiempo necesario para procesarla. Las principales primitivas empleadas son los bordes [Nack, 1977, Pratt, 1991], líneas cresta y líneas valle [Maintz y col., 1996, Lloret y col., 2001, Mariño y col., 2001], segmentación de imágenes [Evans y col., 1996, Chua y Jarvis, 1996], los descriptores de Fourier [Kuhl y Giardina, 1982, Cheng y col., 1994, Belongie y col., 2001] o características estadísticas extraídas a partir de la imagen [Ettinger y col., 1994, Leventon, 2000].
- **Modelos:** Hay ocasiones en las que las variaciones que sufren las estructuras presentes en la imagen no son muy grandes y se puede construir un modelo de las mismas. Este modelo se amolda de tal forma que se adecúa en cada caso a las imágenes que se desea alinear. Existen dos aproximaciones: los *atlas* [Gee y col., 1993, Chen y col., 1998, Toga y Thompson, 2001] en los que se elabora un atlas de imágenes médicas que se utilizará como referencia, y el modelado

de objetos [Kass y col., 1988, McInerney y Terzopoulos, 1996, Cohen y Cohen, 1993, Jones y Metaxas, 1997], en los que se extrae la región de interés de la imagen de referencia y posteriormente se permite que esta región se deforme sobre la imagen flotante.

- **Propiedades de los píxeles de la imagen:** dentro de esta sección se enmarcan aquellos métodos que usan los niveles de gris (o color) de la imagen como espacio de propiedades, sin realizar ningún tipo de segmentación o preprocesado de las mismas, con lo que la medida de similitud que nos indica lo bueno que es el alineamiento de las imágenes se calcula en base a los valores de los propios píxeles de la imagen o medidas extraídas a partir de los mismos [Nagin y col., 1985, Viola y Wells III, 1995, Bro-Nielsen, 1998, Zitová y Flusser, 2003].

1.1.2. Espacio de búsqueda

El tipo de transformación que modela las variaciones de las imágenes es un factor muy importante en el proceso de alineamiento. Los tipos de transformaciones que podremos encontrar en una imagen (mostrados en la figura 1.5) son:

- **Transformaciones rígidas:** en ellas sólo pueden aparecer traslaciones, rotaciones y cambios de escala. Este tipo de transformaciones se aplican sobre todo a imágenes en movimiento en las que los objetos no se deforman ni cambian su tamaño relativo. En el apéndice B se describe en profundidad este tipo de transformaciones, ya que son las que hemos elegido para modelar el comportamiento del ojo en nuestro trabajo. Representando de forma matricial las transformaciones tendremos, dado un punto original $P = (P_x, P_y)$ al que aplicaremos la transformación para obtener el punto $Q = (Q_x, Q_y)$, para el caso de una traslación con componentes (t_x, t_y) (denotando traslación en los ejes X e Y respectivamente), la ecuación 1.4.

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} \quad (1.4)$$

Para la rotación con un ángulo ϕ centrada en el origen se utiliza la ecuación 1.5.

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} \quad (1.5)$$

Finalmente, para el escalado con componentes (s_x, s_y) (escalado en los ejes X e Y respectivamente) la ecuación 1.6:

$$\begin{pmatrix} Q_x \\ Q_y \\ 1 \end{pmatrix} = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ 1 \end{pmatrix} \quad (1.6)$$

- **Transformaciones afines:** en este tipo de transformaciones la característica principal es que se conserva el paralelismo de las líneas y los puntos equiespaciados, pero no los ángulos. Añade a las operaciones que pueden afectar a un objeto, además de las rígidas, la transformación proyectiva, que causa una distorsión de los píxeles a lo largo de un eje sin modificar los mismos a lo largo del otro eje. Su expresión en 2D viene dada por:

$$Sh_x = \begin{pmatrix} 1 & h_x \\ 0 & 1 \end{pmatrix} \quad Sh_y = \begin{pmatrix} 0 & 1 \\ h_y & 0 \end{pmatrix}$$

donde los factores h_x y h_y representan el grado de distorsión de la imagen en los ejes X e Y respectivamente (figura 1.5(a)).

- **Transformación perspectiva:** esta transformación lleva las coordenadas (x, y) en un plano a las coordenadas en la imagen (x', y') . La forma general de una transformación perspectiva es:

$$x' = \frac{a_1x + a_2y + a_3}{a_7x + a_8y + 1}$$

$$y' = \frac{a_4x + a_5y + a_6}{a_7x + a_8y + 1}$$

donde los distintos a_i son constantes que dependen de las ecuaciones de la escena y del plano de la imagen. Esta transformación mapea líneas en líneas, pero no conserva su paralelismo (figura 1.5(b)).

- **Transformaciones elásticas:** son el tipo más general de transformaciones. Estas transformaciones mapean una línea en una curva (figura 1.5(c)). Dentro

de las transformaciones elásticas las más empleadas son las transformaciones polinómicas bilineales, que tienen la siguiente expresión general:

$$x' = \sum_{i=0}^m \sum_{j=0}^i a_{ij} x^i y^{j-i}$$

$$y' = \sum_{i=0}^m \sum_{j=0}^i b_{ij} x^i y^{j-i}$$

donde (x, y) son los índices en la imagen de referencia, (x', y') son los índices en la imagen transformada y a_{ij} y b_{ij} son los coeficientes constantes del polinomio que estamos buscando. El orden del polinomio, m , depende del compromiso entre la precisión del ajuste y la velocidad del proceso de alineamiento.

Además de esto, las transformaciones pueden ser de tipo global o local, según afecten a toda la imagen o a zonas independientes de la misma.

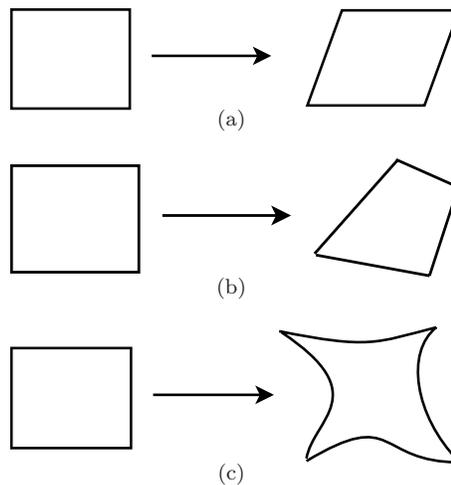


Figura 1.5: Representación de los distintos tipos de transformaciones. (a) Transformación afín. (b) Transformación perspectiva. (c) Transformación elástica.

En nuestro trabajo hemos seleccionado las transformaciones rígidas. Hemos asumido que la transformación perspectiva es muy leve si se tiene en cuenta que el fondo del ojo tiene una variación de la profundidad muy pequeña con respecto a la distancia del centro óptico de la cámara. Lo que causa el desalineamiento es el hecho de que la región capturada del fondo del ojo no es completamente plana, sino cóncava, y los movimientos del ojo no son traslaciones puras, sino que pueden incluir rotaciones del

globo ocular así como variaciones en la escala. Esta presunción se justifica teniendo en cuenta que el ojo es una estructura semi-rígida, y si bien puede aparecer algún tipo de distorsión local, su localización y ajuste supone un tiempo adicional de cálculo que provocaría que nuestro método no fuese viable para su implantación en un entorno hospitalario real (pues cada secuencia *SLO* está compuesta por entre 3000 y 4000 imágenes). Además estas pequeñas distorsiones no van a suponer, como veremos en el capítulo de resultados, grandes diferencias en el proceso de ajuste.

1.1.3. Estrategia de búsqueda

La estrategia de búsqueda viene determinada por el algoritmo de optimización (minimización o maximización según sea el caso) de la medida de semejanza. A su vez, esta medida condiciona la elección del algoritmo, pues su comportamiento será mejor o peor en función de factores como el número de óptimos locales que puedan aparecer. Los métodos de optimización más utilizados son [Press y col., 1992]:

- **Algoritmo de Powell:** se utiliza en la obtención de óptimos en funciones N -dimensionales. En él se parte de un punto P del espacio N -dimensional y un vector \mathbf{v} que nos da una dirección en el mismo espacio. La optimización de una función multidimensional \mathbf{f} se lleva a cabo realizando minimizaciones unidimensionales de cada parámetro del que depende la función. La optimización procede buscando un parámetro λ de tal forma que se optimice $\mathbf{f}(P + \lambda\mathbf{v})$. Luego se reemplaza P por $P + \lambda\mathbf{v}$ y se sigue iterando. Este método se aplica en funciones convexas cuadráticas.
- **Algoritmo Downhill-Simplex:** este método se utiliza también para la minimización de funciones multidimensionales. Requiere del cálculo del valor de la función en cada punto, por lo que es un algoritmo robusto, aunque más costoso desde el punto de vista computacional. Puede darse además una interpretación geométrica del algoritmo. Un *simplex* es una figura geométrica en un espacio N -dimensional formada por $N + 1$ vértices. Así, en 2D tendríamos un triángulo, en 3D tendríamos un tetraedro, y así sucesivamente (figura 1.6).

Los vértices representan los valores de la función, y el algoritmo evoluciona de tal forma que el vértice con un valor mayor sufre una serie de transformaciones:

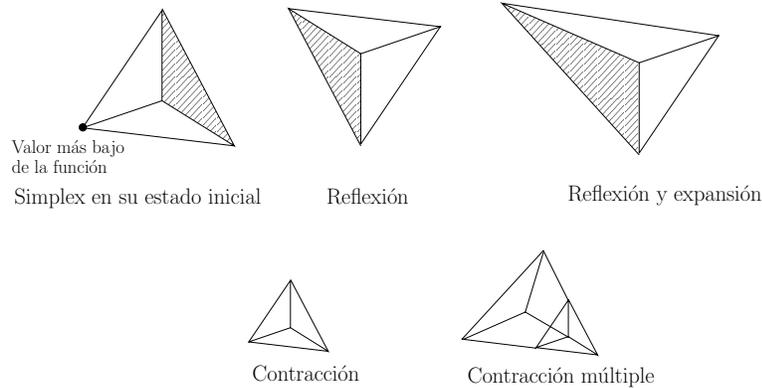


Figura 1.6: Simplex en 3D y tipos de transformaciones.

reflexión, contracción y expansión (figura 1.6), de tal forma que el vértice de mayor valor sea sustituido por uno de menor valor. El algoritmo sigue iterando hasta que se alcanza el máximo de la función o se supera un número determinado de iteraciones, incluido para el caso de que no se alcance un máximo global.

- **Método del gradiente descendente:** este algoritmo se basa en el cálculo del gradiente de la función a optimizar. Realiza movimientos hacia direcciones en las que el gradiente aumente progresivamente, ya que en el máximo de la función el gradiente será nulo. Este método tiene como principal inconveniente su poca eficiencia computacional, ya que el algoritmo progresará lentamente aunque se encuentre en la zona del mínimo.
- **Algoritmos genéticos:** los algoritmos genéticos tratan de emular el comportamiento de la evolución de las especies. Trabajan con una población de soluciones potenciales, donde se aplica el principio de supervivencia, de tal forma que se eligen las soluciones que mejor se aproximen al resultado requerido. Las soluciones posibles se codifican, usualmente en un código binario. El algoritmo comienza inicializando aleatoriamente la solución. Las nuevas soluciones se generarán a partir de las actuales mediante procesos de reproducción, mutación y recombinación, dando lugar a una nueva generación de soluciones. Estas se evalúan y se comprueba si alguna cumple el criterio de finalización. Este proceso lleva a que se consiga una población de soluciones que se va adaptando mejor a la solución

del problema, ya que se desechan aquellas soluciones intermedias que peor se ajustan.

- **Algoritmo del Temple Simulado (*Simulated Annealing*):** tiene su fundamento en el proceso físico de la búsqueda del estado de equilibrio. A una temperatura T elevada, los átomos de los metales están distribuidos aleatoriamente, y según ésta disminuye, los átomos tienden a situarse en una configuración que minimiza la energía del sistema. Esta configuración se alcanza aunque algunos estados intermedios impliquen variaciones positivas en la configuración de la energía. Si el sistema no alcanza el equilibrio puede quedar atrapado en un mínimo local que produzca irregularidades en la configuración final de los átomos.

En el algoritmo se parte de una temperatura inicial T_0 elegida empíricamente, y de un valor aleatorio de los parámetros que minimizan la función. A una temperatura dada se generan valores aleatorios de los parámetros de los que depende la función que se desea minimizar, cuya amplitud depende de la temperatura actual del sistema. La temperatura se reduce poco a poco, de tal forma que los valores generados están cada vez más localizados alrededor del punto que ha proporcionado un valor mejor de la función. Tras situarse en una nueva configuración de los vértices e_i , se calcula la diferencia de energía $\Delta\epsilon = \epsilon(e_i) - \epsilon(e_i)$. Si $\Delta\epsilon > 0$ la posición se acepta, ya que la energía del modelo decrece. Sino se calcula la probabilidad de aceptación P , que depende de forma exponencial de la diferencia de energía $\Delta\epsilon$ y de la temperatura actual T_k de la siguiente forma:

$$P(\Delta\epsilon, T_k) = \exp^{-\left(\frac{\Delta\epsilon}{T_k}\right)} \quad (1.7)$$

Este algoritmo presenta muy buen comportamiento con funciones que presentan muchos mínimos locales.

1.1.4. Medida de similitud

La elección de la medida de similitud está estrechamente relacionada con el espacio de propiedades, ya que ésta se va a obtener sobre las características de la imagen que

se empleen para realizar el alineamiento. Veremos aquí las principales medidas de semejanza del alineamiento que se pueden encontrar en la literatura.

- **Correlación cruzada:** esta es una medida clásica dentro del alineamiento de imágenes. Produce buenos resultados cuando las imágenes sufren pequeños desplazamientos y se usa para encontrar un patrón dentro de ellas. En este método interviene una plantilla $T(x, y)$ y una imagen $I(x, y)$, siendo el tamaño de la plantilla menor que el de la imagen. La correlación cruzada entre $I(x, y)$ y $T(x, y)$ viene dada por:

$$C(u, v) = \sum_x \sum_y T(x, y) I(x - u, y - v) \quad (1.8)$$

donde la suma se extiende sobre la zona de superposición de la plantilla con la imagen. Esta medida presenta el inconveniente de que es sensible a cambios de amplitud (intensidad) y por este motivo se utiliza el coeficiente de correlación cruzada normalizado, dado por:

$$\gamma(u, v) = \frac{\sum_x \sum_y [T(x, y) - \bar{T}] [I(x - u, y - v) - \bar{I}(x - u, y - v)]}{\sqrt{\sum_x \sum_y [T(x, y) - \bar{T}]^2 \sum_x \sum_y [I(x - u, y - v) - \bar{I}(x - u, y - v)]^2}} \quad (1.9)$$

donde \bar{T} es el valor medio de la plantilla e $\bar{I}(x - u, y - v)$ es el valor medio de la imagen en la zona de solape. La correlación cruzada normalizada tiene como inconveniente su elevado coste computacional. Utilizando las propiedades de la transformada de Fourier, se puede implementar una versión muy rápida de la misma, ya que en el dominio transformado el cálculo de la correlación se obtiene mediante el producto de la transformada de la imagen original con el conjugado de la imagen flotante. Existen además implementaciones mucho más eficientes, como la presentada por Lewis [Lewis, 1995].

- **Información mutua:** la información mutua tiene su base en la *Teoría de la información* [Cover y Thomas, 1991]. La información mutua (IM) entre dos imágenes I_1 e I_2 viene dada por:

$$IM(I_1, I_2) = \sum_{i=1}^m \sum_{j=1}^n p(g_i, g_j) \ln \frac{p(g_i, g_j)}{p(g_i)p(g_j)} \quad (1.10)$$

donde m y n son los niveles de gris de las imágenes I_1 e I_2 respectivamente, $p(g_i)$ es la probabilidad de ocurrencia del nivel de gris del píxel i g_i en la imagen I_1 , $p(g_j)$ es la probabilidad de ocurrencia del nivel de gris g_j en la imagen I_2 y $p(g_i, g_j)$ es la probabilidad de que los niveles g_i y g_j ocurran en una misma posición en las imágenes I_1 y I_2 .

Debido a que esta es una de las medidas que utilizaremos en nuestro método de alineamiento, se extenderá esta descripción en el capítulo 3, en el que se describe el proceso de alineamiento de imágenes utilizando la Información Mutua.

- **Medidas extraídas de la Matriz de Coocurrencia de los Niveles de Gris:** Bro-Nielsen [Bro-Nielsen, 1998] propone la utilización de una serie de medidas estadísticas extraídas de la *Matriz de Coocurrencia de los Niveles de Gris*. Estas medidas son la energía, varianza, entropía, información mutua, inercia, momento, o el estadístico de Woods. A partir de este estudio, se concluye que la mejor de estas medidas es la información mutua, motivo por el cual no nos extenderemos más.

1.2. Alineamiento de imágenes

Existen varias clasificaciones posibles de los métodos de alineamiento, según se tome como criterio el tipo de transformación que se busca recuperar, la dimensionalidad de las imágenes (2D/3D, con o sin dimensionalidad temporal), la modalidad de las mismas (monomodales o multimodales), el dominio de aplicación del método (imágenes médicas, topológicas, seguimiento de objetos, ...) o la naturaleza de la base del alineamiento (utilización o no de marcadores, que pueden ser intrínsecos o extrínsecos).

En este apartado analizaremos los métodos de alineamiento siguiendo este último criterio, la naturaleza de la base del alineamiento, centrándonos además en el dominio del ajuste de imágenes médicas.

Siguiendo esta clasificación, los métodos de alineamiento de imágenes pueden agruparse en dos grandes categorías: métodos basados en la extracción de características y métodos basados en la intensidad de las imágenes. Los métodos del primer grupo siguen dos procesos: en la primera etapa se extrae algún tipo de marcador de la imagen, como propiedades de la imagen o marcadores externos. En la segunda fase se calcula la transformación que mejor superpone las marcas extraídas. Dependiendo de las características de las imágenes, la primera etapa puede resultar de por sí un proceso complicado, con el agravante de que en estos métodos la precisión del ajuste depende en gran medida de la calidad de la segmentación inicial de las características. En el segundo grupo se encuentran aquellos algoritmos que proceden al alineamiento de las imágenes utilizando directamente la intensidad de las mismas o bien alguna propiedad calculada directamente sobre estas.

Enumeraremos aquí algunos de los trabajos más relevantes de ambos grupos que nos ayudarán a tener una perspectiva global de las técnicas de alineamiento de imágenes, así como de los problemas dentro del campo de la visión artificial que se pueden abordar mediante las mismas. Para una revisión más exhaustiva sobre técnicas de alineamiento pueden consultarse los trabajos de Brown [Brown, 1992], Maintz y Viergever [Maintz y Viergever, 1998] o Van den Elsen y col. [Van den Elsen y col., 1993].

1.2.1. Técnicas basadas en la extracción de características

Como ya hemos comentado, las técnicas de alineamiento basadas en la extracción de características basan su funcionamiento en la extracción previa al proceso de alineamiento de una serie de marcadores, bien sean externos (extrínsecos), como algún tipo de mecanismo externo al paciente que se introduce en la imagen, o bien internos (intrínsecos), como información inherente al propio paciente que se registra en la imagen, y que servirán de punto de referencia al proceso de ajuste.

La figura 1.7 representa el procedimiento general de cualquiera de estos algoritmos. En la primera fase del proceso (marcada con (1) en la figura) se extraen las características de las imágenes que se utilizarán como puntos de referencia en el alineamiento. Este proceso puede ser desde un simple filtro detector de bordes a complejos procedimientos de segmentación. En la etapa marcada como (2) en la figura 1.7 se lleva a

cabo el proceso iterativo en el cual se alinea la imagen flotante contra la de referencia utilizando las imágenes de características como entrada del proceso. El resultado será la imagen flotante transformada de forma que se maximiza la superposición de la misma con la imagen original (o bien la matriz de dicha transformación).

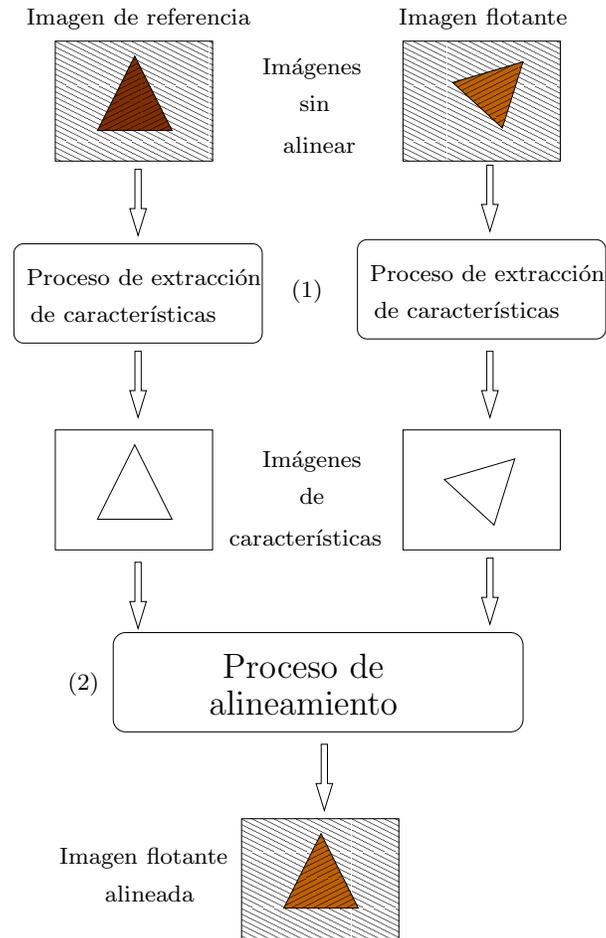


Figura 1.7: Esquema del proceso de alineamiento basado en la extracción de características. En el paso (1) se extraen las características relevantes de las imágenes de referencia y flotante. En el paso (2) se lleva a cabo el alineamiento, mediante un proceso iterativo en el cual se transforma la imagen flotante hasta que se obtiene un alto grado de solapamiento con la imagen de referencia.

Domingo y col. [Domingo y col., 1997] llevan a cabo el alineamiento de angiografías de retina mediante el ajuste robusto de los puntos de bifurcación detectados utilizando el operador de Förstner adaptado a los distintos niveles de saturación mediante una

ecualización del histograma. Para guiar el proceso de búsqueda cada punto detectado se asocia a un valor invariante a la rotación y traslación, calculado utilizando la correlación cruzada en una vecindad local. Un proceso de búsqueda estocástico asocia pares de puntos de una imagen con los de la otra, y un proceso de optimización iterativo final refina el resultado. Según los resultados mostrados en este trabajo, el proceso de alineamiento funciona correctamente, si bien los tiempos de procesamiento resultantes de ejecutar el método en un ordenador HP-735 son demasiado elevados (entre 114 y 152 segundos, con una media de 135 segundos) para los requerimientos de nuestro problema.

Pinz y col. [Pinz y col., 1995, Pinz y col., 1998] proponen un método de este tipo en el cual se utiliza un algoritmo basado en la intensidad para extraer las características relevantes de la imagen (venas, fovea, escotoma, ...) que se utilizan como base en el alineamiento. El problema de este método es que el tiempo de alineamiento de las imágenes es bastante elevado, por lo que no es adecuado para grandes conjuntos de datos.

Wade y Fitzke [Wade y Fitzke, 1998] desarrollaron un método rápido y robusto, que emplea una variante del algoritmo de detección de picos en la correlación cruzada normalizada, con una estrategia de búsqueda para acelerar la velocidad de operación. Sin embargo, en este método se asume que las transformaciones poseen sólo componente traslacional entre los fotogramas de referencia y flotante, descartando las componentes de rotación y de escalado.

Noack y Sutton [Noack y Sutton, 1994] presentan un método de registro muy rápido para secuencias de *SLO*, pero tampoco consideran las rotaciones, y los tiempos de cálculo reseñados se refieren a alineamientos de pequeñas zonas de 32×32 píxeles que se utilizan para calcular la correlación entre ambas imágenes, y con ello el nivel de ajuste de las mismas.

En la aproximación propuesta por Zana y Klein [Zana y Klein, 1999] se utilizan los vasos detectados tras una larga secuencia de operaciones, incluyendo operadores morfológicos y transformada Laplaciana. Después se extraen los puntos de bifurcación aplicando reconocimiento de patrones a un conjunto de configuraciones posibles. Para cada punto se calcula una medida invariante, en este caso relacionada con la

geometría local del vaso. Finalmente una transformada bayesiana de Hough selecciona los mejores candidatos para un proceso iterativo posterior, en el que los puntos de bifurcación se ajustan de forma precisa hasta que se alcanza la convergencia deseada. En este trabajo tan solo se muestran los resultados de la evaluación del método sobre 29 pares de imágenes, lo cual parece insuficiente para determinar la robustez de un método de alineamiento. Además en el trabajo descrito se considera un margen de error válido de 10 píxeles, margen que en nuestro caso es demasiado amplio, pues los vasos tienen en muchos casos un diámetro inferior a 10 píxeles, y es necesario, como veremos en el capítulo de resultados, que los puntos seleccionados por el experto clínico se mantengan en el interior de las venas a lo largo de toda la secuencia. Por otro lado, al igual que la gran mayoría de las técnicas analizadas, el método es demasiado lento para la evaluación de las secuencias de *SLO*.

Hsu y Loew [Hsu y Loew, 2001] proponen una técnica que realiza la segmentación utilizando una versión aproximada y eficiente del conocido método de detección de bordes de Canny [Canny, 1986] combinado con un algoritmo de etiquetado para el alineamiento multimodal de imágenes de tomografía computerizada y resonancia magnética del cráneo. Los bordes extraídos se utilizan en un método de ajuste basado en el cálculo de los mínimos cuadrados mediante un método que minimiza la suma de los cuadrados de las distancias de cada par de puntos.

Finalmente, Lloret y col. [Lloret y col., 2001] muestran los primeros pasos del método que aquí presentamos, basado en la extracción de las denominadas líneas cresta y líneas valle.

1.2.2. Técnicas basadas en la la intensidad de las imágenes

Las técnicas anteriores dan buenos resultados cuando es posible extraer las características que utilizan como referencia. Sin embargo, como veremos en el caso de las videoangiografías de *SLO*, existen muchos fotogramas en los que no es posible extraer dichas características, sobre todo debido a que el contraste es muy bajo. Por ello es necesario recurrir a técnicas que trabajen directamente sobre los valores de gris de la imagen o sobre alguna medida extraída de los mismos. Estos métodos presentan la ventaja adicional de que trabajan con toda la información disponible en la imagen, no

con partes de la misma, por lo que suelen ser más robustos ante oclusiones o desaparición de partes de la imagen, a costa, eso sí, de unos requerimientos computacionales más elevados. Estos métodos funcionan buscando, en un cierto espacio de transformaciones, aquella que maximiza un criterio que mide la similitud de intensidad entre píxeles.

Peli y col. [Peli y col., 1987] presentan un método de alineamiento de imágenes de la retina basado en el cálculo de una medida de similitud, la Suma de los Valores Absolutos de las Diferencias (SVAD) (ecuación 1.11) y la determinación del error acumulado E modificado mediante un factor de normalización, para corregir los cambios de intensidad (ecuación 1.12). La aproximación es similar a la del cálculo de la correlación: una plantilla calculada automáticamente mediante un filtro adaptativo se desplaza por la imagen buscando el máximo de la SVAD, definido como:

$$SAVD(m, n) = \sum_{j=1}^J \sum_{k=1}^K |F_1(j, k) - F_2(j - m, k - n)| \quad (1.11)$$

donde $F_1(j, k)$ es una ventana de tamaño $J \times K$ que se utiliza como plantilla, y $F_2(j, k)$ es la ventana de búsqueda de tamaño $M \times N$, mayor que el de F_1 . El error acumulado se define como:

$$E(m, n) = \sum_j \sum_k |(F_1(j, k) - TT) - (F_2(j - m, k - m) - SS(m, n))| \quad (1.12)$$

donde TT es la media del nivel de gris en la plantilla F_1 y SS la media local de la sección de la plantilla dentro de la ventana de búsqueda F_2 .

Según los autores, el tiempo de cálculo del método es de unos 20 segundos para cada par de fotogramas, lo cual nos llevaría a varias horas de cálculo para cada secuencia de *SLO*. Además, para cada secuencia, es necesario calcular la plantilla que se utilizará para buscar la transformación que ajuste las imágenes.

Ott y col. [Ott y col., 1990] proponen un método de alineamiento basado en el cálculo y minimización de la distancia especificada en la ecuación 1.13, que representa la diferencia entre una plantilla y la imagen.

$$D = \sum_{x,y} |T(x,y) - I(x,y)| \quad (1.13)$$

donde $T(x,y)$ son los niveles de gris de la plantilla, $I(x,y)$ son los niveles de gris de la imagen, y la suma se toma sobre el tamaño de la plantilla.

Si bien este método proporciona buenos resultados, el tiempo de cálculo es demasiado elevado, y además es muy dependiente del nivel de desajuste entre imágenes.

Cheng y col. [Cheng y col., 1994] presentan un método que trabaja en el dominio de la transformada de Fourier, ajustando las imágenes mediante el empleo de unos descriptores invariantes a rotaciones, y que además representan las rotaciones y escalados como traslaciones en el espacio paramétrico. El procedimiento utiliza la fase espectral de los descriptores calculados a partir de las imágenes de referencia y flotante como entrada a un filtro ajustado, cuya salida máxima proporcionará la transformación buscada. Sin embargo, este método sólo es evaluado por los autores con traslaciones y rotaciones pequeñas en varios dominios, entre ellos imágenes médicas. Además es un proceso computacionalmente muy costoso, pues al cálculo de los descriptores hay que unir el cálculo de los filtros ajustados en fase, que son los que permiten detectar las rotaciones y escalados.

Bro-Nielsen [Bro-Nielsen, 1998] presenta un estudio completo de varias medidas calculadas a partir de la matriz de coocurrencia de los niveles de gris que, aunque no se aplican al alineamiento de imágenes oftalmológicas, sí nos sirven para extraer alguna conclusión útil para nuestro trabajo. Para la estimación de las probabilidades, y debido a la complejidad computacional de los cálculos de la matriz de coocurrencias, Bro-Nielsen realiza un submuestreo con un factor de 2, y a partir de dicha matriz se construyen las distintas probabilidades. Como método de optimización se utiliza el algoritmo de Powell descrito anteriormente. Finalmente, de las distintas funciones de alineamiento estudiadas, la que mejor resultados produce es la información mutua.

La información mutua (IM) es una medida que se utilizó para el alineamiento multimodal de imágenes de tomografía computerizada y resonancia magnética del cráneo [Viola y Wells III, 1995], siendo después utilizada en multitud de trabajos [Wells y col., 1996, Studholme y col., 1996, Maes y col., 1997, Pluim y col., 2001] y

proporcionando muy buenos resultados. La diferencia más significativa entre todos estos trabajos es la forma de estimación de las probabilidades que se utilizan para su cálculo, y el método de maximización de la misma. Así, Studholme y col. [Studholme y col., 1996], Maes y col. [Maes y col., 1997] y Plum y col. [Plum y col., 2001] estiman las probabilidades simples y conjuntas a partir de las distribuciones de los niveles de gris de los pares de píxeles, construyendo el histograma de cada imagen y el histograma conjunto. Sin embargo, mientras que estos autores utilizan el algoritmo de Powell para su maximización, en los otros dos trabajos [Studholme y col., 1996, Maes y col., 1997] se utiliza un método en el que se combina la evaluación multiresolución y la optimización multiresolución. Dada una transformación inicial T_0 , evalúan la información mutua para un conjunto de 13 transformaciones, siendo las traslaciones y rotaciones utilizadas función a su vez de la resolución en la que se encuentran en la pirámide multiresolución. Una vez optimizada en el nivel actual, se pasa al nivel siguiente.

Por otro lado, Viola y Wells [Viola y Wells III, 1995] y Wells y col. [Wells y col., 1996] utilizan un método no paramétrico de estimación de las probabilidades, el denominado *método de la ventana de Parzen*. Para la maximización de la *información mutua* utilizan un método estocástico análogo al descenso de gradientes con un parámetro λ de aprendizaje, cuyo valor es proporcional a la aproximación de la derivada de la información mutua.

Likar y Pernus [Likar y Pernus, 2001] presentan un algoritmo de alineamiento basado en el cálculo de la información mutua normalizada siguiendo también una estructura jerárquica, para ajustar imágenes de músculo humano, obteniendo buenos resultados. Además utilizan para el cálculo de las probabilidades la aproximación propuesta por Maes y col. [Maes y col., 1997], descartando el método de las ventanas Parzen [Viola y Wells III, 1995]. En el trabajo desarrollado en esta tesis doctoral también hemos realizado la implementación de ambos métodos (como se verá en el capítulo 3) llegando a las mismas conclusiones, ya que el método de las ventanas Parzen implica una sobrecarga computacional que lo hace inviable para nuestra aplicación.

Los métodos de alineamiento basados en la información mutua [Ritter y col., 1999, Pardo y col., 2001] también han dado buenos resultados para el ajuste de imágenes

de retina. El trabajo de Ritter y col. [Ritter y col., 1999] utiliza el alineamiento basado en información mutua combinado con el método del Temple Simulado para su maximización. A pesar de que los resultados son muy buenos, los tiempos de cálculo son bastante elevados, debido sobre todo a la aproximación jerárquica que utilizan para mejorar los resultados. Además, las imágenes utilizadas en este trabajo representan tan sólo un pequeño ejemplo de las que aparecen en las secuencias de *SLO*. El método presentado en Pardo y col. [Pardo y col., 2001] es muy similar, aunque realiza un estudio más exhaustivo de las imágenes que pueden aparecer en las videoangiografías *SLO*. Por otro lado, tiene el mismo inconveniente: el elevado tiempo de cálculo necesario para cada par de imágenes, que si bien no es alto en sí mismo, sí es demasiado alto para que este método sirva para el alineamiento de secuencias completas.

En resumen, como acabamos de ver, existen gran variedad de aproximaciones al alineamiento de imágenes oftalmológicas. Sin embargo, el alineamiento de grandes secuencias de imágenes presenta varios problemas que hacen que el ajuste de todos los fotogramas de una secuencia sea un proceso más complicado:

1. Es necesario un tiempo reducido para el alineamiento de cada fotograma.
2. Existe gran cantidad de ruido en las imágenes.
3. Aparecen grandes diferencias entre fotogramas consecutivos, debidos a movimientos rápidos del ojo.
4. Dificultad en la selección del fotograma de referencia: debido a que el nivel del contraste presente en los fotogramas varía mucho (nivel del contraste muy bajo en los fotogramas del principio de la secuencia, nivel del contraste muy alto en el medio de la secuencia y otra vez nivel bajo del contraste en los fotogramas del final de la secuencia), alinear todos los fotogramas contra un único fotograma de referencia es prácticamente inviable, por lo que será necesario elegir varios fotogramas de referencia.

Es por ello que, si bien muchas de las técnicas revisadas podrían llevar a cabo el alineamiento de muchos fotogramas de la secuencia ofreciendo buenos resultados, ninguna de ellas por sí sola sería válida para realizar el alineamiento de una secuencia

completa, debido a la variabilidad de las imágenes o al tiempo que emplean en realizar el proceso de ajuste.

El trabajo que presentamos aquí soluciona todos estos problemas mediante la combinación de dos técnicas de alineamiento, y así donde una de ellas encuentra problemas, se podrá aplicar la otra para corregirlos. Sin embargo, esto presenta el inconveniente de que hay que encontrar la forma de fusionar los resultados de ambos métodos, además de la forma de determinar cuándo aplicar cada uno de ellos. Veremos en los capítulos siguientes, además de la presentación de cada una de las técnicas, cómo se han resuelto estas dos cuestiones, de forma que el resultado final sea la secuencia de *SLO* alineada en un tiempo adecuado para la utilización del sistema en un entorno hospitalario.

Capítulo 2

Alineamiento de imágenes de *SLO*

Las imágenes objeto de este trabajo se obtienen mediante el dispositivo denominado *Oftalmoscopio con Láser de Barrido (Scanning Laser Ophthalmoscope, SLO)*. El SLO se basa en una técnica en la que se ilumina el fondo del ojo con un rayo láser barriendo la superficie de la retina, capturando a continuación la luz reflejada mediante un detector (ver figura 2.1). La imagen del fondo se visualiza en tiempo real en un monitor que ayuda a localizar los vasos. La longitud de onda infrarroja que se utiliza puede penetrar opacidades del medio, como cataratas, opacidades de la córnea y hemorragias. La figura 2.2 muestra un equipo de SLO y un ejemplo de la imagen capturada.

Las principales ventajas del SLO sobre las cámaras tradicionales son:

1. No es necesario dilatar la pupila para la adquisición de imágenes ya que el SLO puede capturar imágenes incluso con una apertura muy pequeña.
2. El paciente tolera mejor el procedimiento ya que la intensidad de la luz que necesita el *SLO* para iluminar la retina es muy inferior a la requerida por la cámara de fondo de ojo, con lo cual se reduce significativamente la molestia para

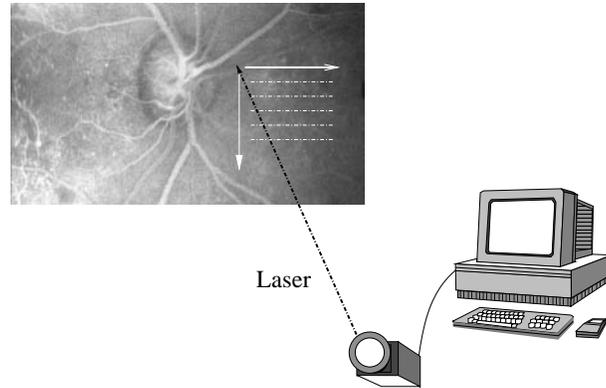


Figura 2.1: Fotografía de un fondo de ojo donde se representa la trayectoria del láser sobre la retina. Las flechas indican la dirección horizontal y vertical que sigue el láser.

el paciente así como el riesgo de dañar la retina si se llevan a cabo demasiadas exposiciones

3. La resolución proporcionada por el SLO es mucho mejor que la de las cámaras convencionales.
4. Las longitudes de onda infrarrojas del SLO pueden penetrar opacidades del medio, como cataratas, opacidades de la córnea y hemorragias, proporcionando así información no observable mediante otras técnicas tradicionales.
5. Puede capturarse un mayor número de imágenes con una velocidad muy superior (el *SLO* puede capturar imágenes a velocidades variables de hasta la frecuencia de un vídeo, 25Hz).
6. El número de imágenes que se capturan sólo está limitado por la capacidad de almacenamiento del disco duro del ordenador al que se encuentre conectado el dispositivo de adquisición.
7. Desde el punto de vista de la calidad de la imagen adquirida, la geometría de la óptica del *SLO* y lo reducido del haz de luz láser reduce la variación de la iluminación que se produce en la imagen, si se compara con las imágenes fotográficas que se empleaban usualmente.

El fondo del ojo se ilumina con un rayo láser estrecho (que puede ser de tres tipos: Argón (488 nm, azul; 514 nm, verde), Helio-Neón (HeNe) (632.8 nm rojo) o un

diodo láser infrarrojo (780 nm)), que barre la superficie de la retina (figura 2.1) y un detector que captura la luz reflejada. El láser barre el área del objeto píxel a píxel. El tamaño de cada punto de láser sobre la retina es de 10 micras y dura 100 ns. Cada línea está compuesta por 640 puntos y dura 64 microsegundos. La separación vertical entre líneas es de 7 micras. El número de líneas depende del sistema de TV utilizado y puede ser de 525 o 625. Los tamaños típicos de las imágenes obtenidas son 1024×1024 o 512×512 . La imagen del fondo del ojo se visualiza en tiempo real en un monitor, normalmente con una frecuencia de entre 20 y 30 fotogramas por segundo.

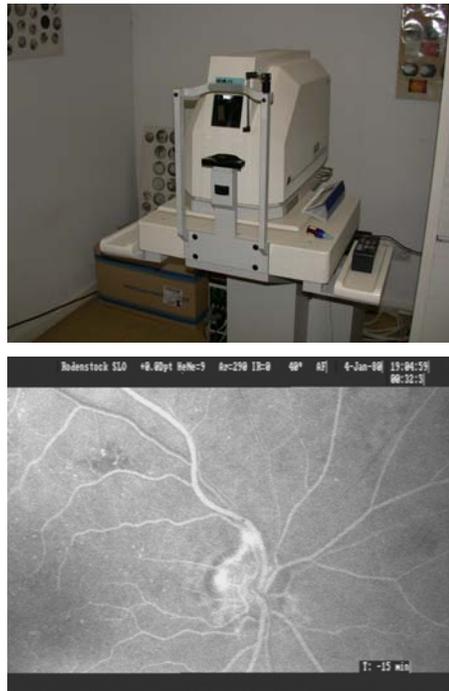


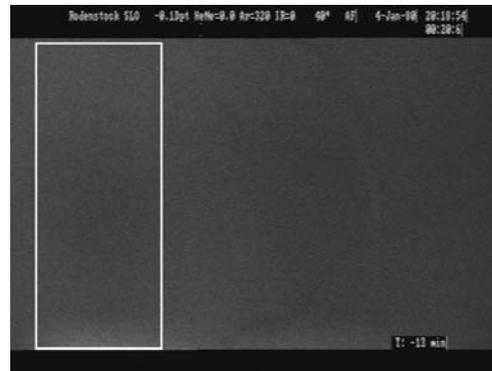
Figura 2.2: Fotografía de un equipo de *SLO* (arriba) y ejemplo de fotograma de una secuencia de *SLO* (abajo).

Sin embargo, hay que mencionar que el *SLO* también presenta algunas desventajas frente a otros sistemas. El equipo es muy costoso y de una gran complejidad técnica. Su resolución espacial sigue siendo menor que la de una retinografía realizada con película fotográfica.

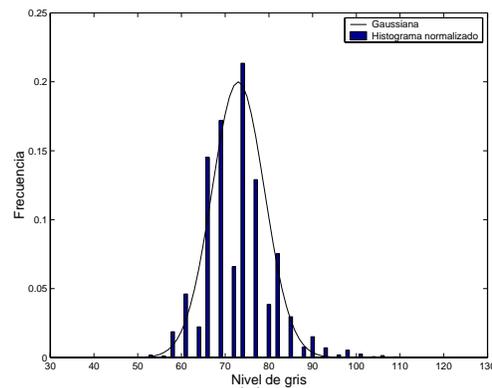
Una característica muy importante en el análisis de cualquier imagen es la relación señal/ruido, que en el caso de los fotogramas de *SLO* es un valor bajo. El ruido en

los fotogramas lo origina la naturaleza óptica de los dispositivos de adquisición: la luz que mide el dispositivo debe entrar y salir de la retina, que está compuesta por un conjunto de tejidos traslúcidos que dificultan el modelado. Además los dispositivos de transmisión y grabación añaden algo de ruido a la señal original.

Para minimizar el problema del ruido se ha realizado un estudio que analiza la naturaleza del mismo: en la figura 2.3(a) se tomó una ventana que contiene sólo información de fondo de imagen, y se calculó el histograma de los valores obtenidos. El histograma, como se puede apreciar en la figura 2.3(b), se aproxima bien mediante una curva Gaussiana de media $\mu = 73,2$ y desviación típica $\sigma = 9,12$.



(a)



(b)

Figura 2.3: Para analizar el ruido presente en los fotogramas de las secuencias de *SLO* se tomó un área del fondo de distintas imágenes (a) y se calculó su histograma normalizado (b). Se aprecia claramente cómo el fondo se puede modelar de forma aproximada mediante una Gaussiana, que en este ejemplo tendría media $\mu = 73,2$ y desviación típica $\sigma = 9,12$.

Los resultados obtenidos al calcular histogramas del fondo para otros fotogramas de esta secuencia así como para fotogramas de otras secuencias fueron muy similares. En [Lois y col., 1999] puede encontrarse un estudio similar aunque mucho más detallado. En este trabajo los autores evalúan la repetibilidad de las medidas del fondo en imágenes *SLO*, con un grupo de 10 pacientes y 10 voluntarios. Una vez que expertos oftalmólogos seleccionan manualmente los valores de fondo, se observan diferencias en los rangos (1,4 – 13,5) para los voluntarios y (1,5 – 9,3) para los pacientes. Por lo tanto, podemos concluir que antes de utilizar los valores de las imágenes originales, es conveniente convolucionarlas con un filtro Gaussiano para eliminar algo de ruido. Este proceso de suavizado es similar al que se utiliza en el cálculo de las derivadas de las imágenes.

2.1. Registro basado en extracción de características: crestas y valles

Cuando se captura una secuencia de *SLO*, como ya se ha comentado, el paciente ha de permanecer quieto y con la mirada fija en un punto durante un período de tiempo que comprende entre 30 y 40 segundos, lo cual es prácticamente imposible, por lo que se producirán movimientos del ojo, que pueden llegar a ser de hasta $600^\circ/s$ [Markov y col., 1993]. Es por esto que será necesario realizar un paso previo de alineamiento de los fotogramas para poder realizar análisis cuantitativos de la secuencia adquirida.

El método que describimos en este trabajo sigue el procedimiento de ajuste utilizado habitualmente por los oftalmólogos, en el sentido de que se utilizan como líneas guía aquellas características comunes a ambas imágenes. En contraste con que otros métodos existentes en la bibliografía, no restringimos la comparación a los puntos de bifurcación de los vasos, sino que se comparan las estructuras más significativas de la imagen, es decir, el árbol arterio-venoso. Al contrario que las aproximaciones basadas en el ajuste utilizando como referencia los puntos de bifurcación [Zana y Klein, 1999, Domingo y col., 1997], nuestro método no tiene una dependencia tan grande de la calidad de la segmentación. En este sentido, este método puede considerarse similar al propuesto por Pinz y col. [Pinz y col., 1998].

El algoritmo se divide en dos partes. La primera, que trabaja a escala local, es la encargada de extraer de forma adecuada las crestas y valles de la imagen a partir de las arterias y venas de la misma. La segunda toma la estructura de crestas y valles obtenida para llevar a cabo el proceso de alineamiento. Debido a que el proceso de segmentación es idéntico para todas las imágenes y las características de las mismas varían de forma impredecible, el alineamiento debe ser robusto frente a la existencia de estructuras que no aparecen en alguna de las dos imágenes, o bien ante estructuras que no se solapan al 100 %.

2.1.1. Método de extracción de crestas y valles

En las técnicas de alineamiento basadas en la extracción de características es necesario encontrar algún tipo de marcador cuya extracción sea posible en todas las imágenes, y que de alguna forma caracterice a esa imagen entre todas las posibles.

Los vasos son unos marcadores fiables en las imágenes de retina ya que se trata de estructuras casi rígidas que son visibles en todas las modalidades. Además pueden verse como picos y valles si se ve a las imágenes como relieves topográficos, tal y como muestra la figura 2.4

Entre las muchas definiciones de *pliege*, aquella basada en los conjuntos de niveles de curvatura extrínseca (*level set extrinsic curvature, LSEC*) posee propiedades de invarianza muy útiles. Dada una función $L : \mathbb{R}^d \rightarrow \mathbb{R}$, el conjunto de niveles para una constante l está formado por el conjunto de puntos $\{\mathbf{x} | L(\mathbf{x}) = l\}$ (Fig. 2.5). Para imágenes 2D L puede ser visto como un relieve topográfico, y los conjuntos de nivel son sus curvas de nivel. El mínimo negativo de la curvatura de las curvas de nivel κ , nivel a nivel, forma las denominadas *curvas valle*, mientras que los máximos positivos forman las *curvas cresta*. El LSEC (en 2D) puede expresarse en términos de las derivadas de L siguiendo métodos de cálculo tensorial [ter Haar Romeny y Florack, 1993] como indica la ecuación (2.1):

$$\kappa = (2L_x L_y L_{xy} - L_y^2 L_{xx} - L_x^2 L_{yy}) (L_x^2 + L_y^2)^{-\frac{3}{2}} \quad (2.1)$$

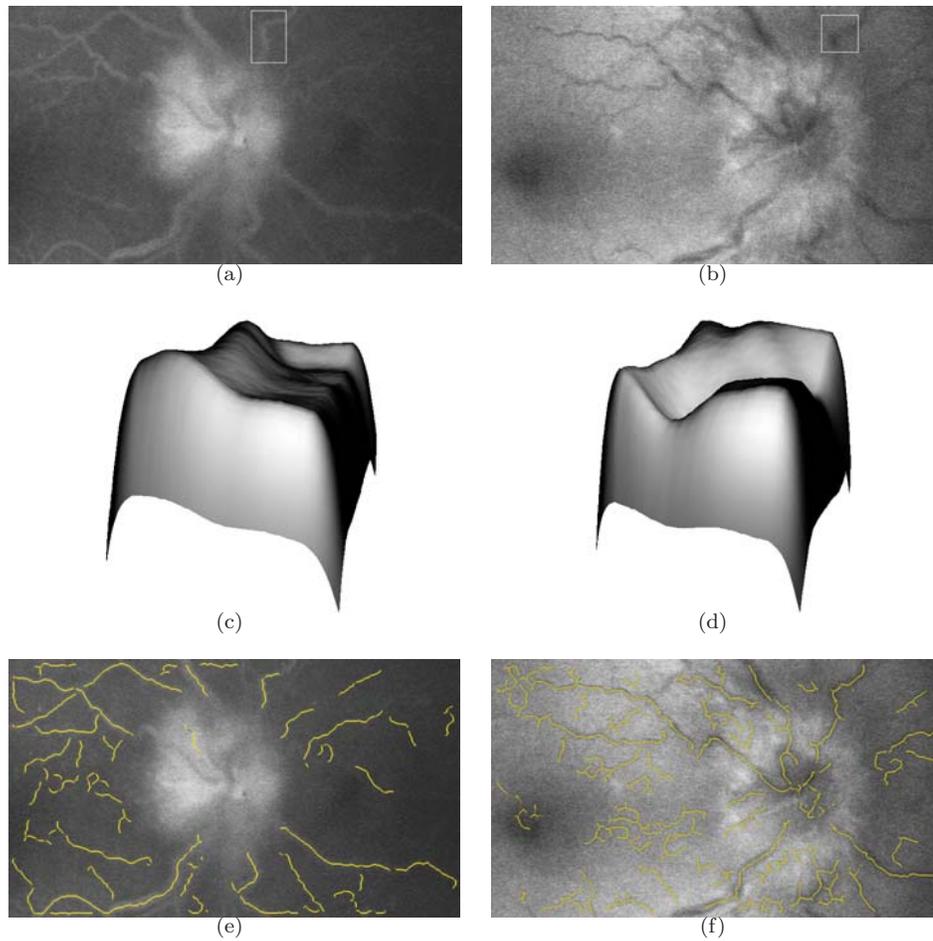


Figura 2.4: (a) y (b) muestran los fotogramas número 700 y 3100 de una secuencia de 3500 imágenes de *SLO*. (c) y (d) muestran la región de interés de las imágenes (a) y (b) representada como un relieve topográfico. Obsérvese como, a pesar de la variación en intensidad y contraste, las crestas y valles son claramente visibles. Finalmente, (e) y (f) muestran las imágenes (a) y (b) con las crestas y valles superpuestos en amarillo.

donde

$$L_\alpha = \frac{\partial L}{\partial \alpha}, L_{\alpha\beta} = \frac{\partial^2 L}{\partial \alpha \partial \beta}, \alpha, \beta \in \{x, y\}$$

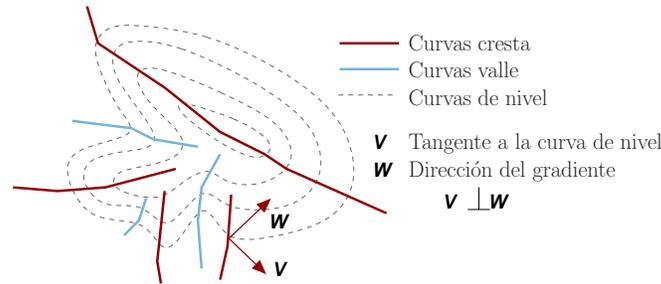


Figura 2.5: Interpretación gráfica de los términos relacionados con las curvas de nivel: Los mínimos negativos de la curvatura κ de las curvas de nivel, nivel a nivel, forman las *curvas valle*, y los máximos positivos forman las *curvas cresta*. Los vectores v y w representan la tangente y la dirección del gradiente, respectivamente, en cada punto de las curvas de nivel, siendo además perpendiculares el uno con respecto al otro.

Sin embargo, la discretización del LSEC empleada habitualmente está mal definida en algunos casos, dando lugar a discontinuidades inesperadas en el centro de objetos alargados. Estos problemas se deben a que la definición del LSEC es demasiado local, lo que no resulta apropiado en el dominio discreto.

Para evitar los problemas de las discontinuidades alrededor de puntos críticos, hemos empleado el operador (*Multilocal Level Set Extrinsic Curvature - Structure Tensor*) $\tilde{\kappa}_d$, un operador multi-local basado en el LSEC, tal y como se define en [López y col., 2000].

En $2D$ (ver figura 2.5) κ puede definirse a través de su relación con las líneas de pendiente, que son las líneas que integran el campo vectorial del gradiente \mathbf{w} y son, por lo tanto, ortogonales a las curvas de nivel. Debido a esta ortogonalidad, cuando las curvas de nivel son líneas rectas paralelas, las líneas de pendiente también son paralelas y rectas, y cuando las curvas de nivel se doblan, las líneas de pendiente convergen/divergen. En cálculo vectorial existe un operador *divergencia*, que mide este grado de paralelismo. La divergencia de un campo vectorial d -dimensional $\mathbf{u} : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\mathbf{u}(\mathbf{x}) = (\mathbf{u}^1(\mathbf{x}), \dots, \mathbf{u}^d(\mathbf{x}))^t$ se define como [Schey, 1973]:

$$\operatorname{div}(\mathbf{u}) = \sum_{i=1}^d \frac{\partial u^i}{\partial x^i} \quad (2.2)$$

si denotamos por $\mathbf{0}_d$ el vector cero d -dimensional, el campo vectorial normalizado del gradiente de $L : \mathbb{R}^d \rightarrow \mathbb{R}$ $\bar{\mathbf{w}}$ se define como:

$$\bar{\mathbf{w}} = \begin{cases} \mathbf{w}/\|\mathbf{w}\| & \text{if } \|\mathbf{w}\| > 0 \\ \mathbf{0}_d & \text{if } \|\mathbf{w}\| = 0 \end{cases} \quad (2.3)$$

y entonces se puede demostrar ([López y col., 2000]) que:

$$\kappa_d = -\operatorname{div}(\bar{\mathbf{w}}) \quad (2.4)$$

En $2D$ tomemos el punto \mathbf{x} donde se desea calcular la divergencia de un campo vectorial $2D$. Sea \mathcal{C} una curva cerrada simple en \mathbb{R}^2 parametrizada por l , y que encierra al punto \mathbf{x} . Sea \mathbf{n} su vector normal unitario y ω el área encerrada por \mathcal{C} (figura 2.6). Entonces la divergencia de \mathbf{u} en \mathbf{x} también puede definirse como [Schey, 1973]:

$$\operatorname{div}(\mathbf{u}) = \lim_{\omega \rightarrow 0} \frac{1}{\omega} \int_{\mathcal{C}} \mathbf{u}^t \cdot \mathbf{n} \partial l \quad (2.5)$$

En [López y col., 2000] se muestra cómo la sustitución de la definición local de κ_d por una definición multilocal basada en una versión discretizada de la ecuación 2.5, donde la multilocalidad se obtiene asumiendo que el vecindario \mathcal{W} alrededor de un punto \mathbf{x} , o de forma análoga, su límite cerrado \mathcal{C} , es un parámetro seleccionable. Es decir, para calcular $\operatorname{div}(\bar{\mathbf{w}})$ en \mathbf{x} se tienen en cuenta los vectores gradiente a lo largo de la ruta \mathcal{C} alrededor de \mathbf{x} .

Siguiendo este razonamiento, para una dimensión d dada, denotaremos por $\bar{\kappa}_d$ el operador MLSEC (*Multilocal Level-Set Extrinsic Curvature*) basado en las ecuaciones 2.3, 2.4 y 2.5, dado un \mathcal{C} seleccionado. El operador MLSEC para un dominio discreto se define como [López y col., 2000]:

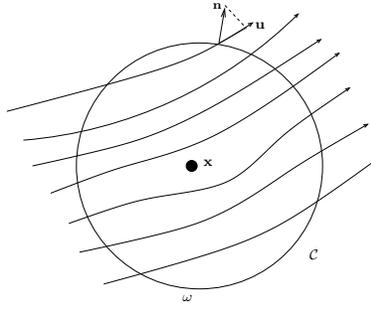


Figura 2.6: Geometría implicada en la definición de la divergencia del campo vectorial \mathbf{u} en el punto \mathbf{x} . \mathcal{C} es una curva cerrada simple en \mathbb{R}^2 parametrizada por l que encierra al punto \mathbf{x} , con vector normal unitario es \mathbf{n} . ω representa el área encerrada por \mathcal{C} .

$$\bar{\kappa}_d = -\text{div}(\bar{\mathbf{w}}) = -\frac{d}{r} \sum_{k=1}^r \bar{\mathbf{w}}_k^t \cdot \mathbf{n}_k \quad (2.6)$$

donde r es la adyacencia seleccionada que vendrá dada por el \mathcal{C} específico (por ejemplo, en $2D$ podemos elegir entre adyacencia 4 o adyacencia 8), y d es la dimensionalidad del espacio.

Una vez que se ha establecido $\bar{\kappa}_d$ como una buena medida de las crestas y valles, todavía puede ser mejorada prefiltrando el campo vectorial del gradiente de la imagen para incrementar el grado de atracción/repulsión en las curvas-cresta y curvas-valle, que es lo que mide κ . Esto puede hacerse utilizando un tensor estructural, que es una herramienta bien conocida para el análisis de texturas orientadas [Bigun y col., 1991, Jähne, 1993, Jähne, 2002].

En el espacio d -dimensional, dada una vecindad simétrica de tamaño σ_I centrada en un punto dado \mathbf{x} , lo que denotaremos por $\mathcal{N}(\mathbf{x}; \sigma_I)$, el tensor estructural se define como la matriz simétrica semidefinida positiva de tamaño $d \times d$:

$$\mathbf{M}(\mathbf{x}; \sigma_I) = \mathcal{N}(\mathbf{x}; \sigma_I) * (\mathbf{w}(\mathbf{x}) \cdot \mathbf{w}^t(\mathbf{x})) \quad (2.7)$$

donde “ $*$ ” representa la operación de convolución.

El autovector que se corresponde con el menor autovalor de $\mathbf{M}(\mathbf{x}; \sigma_I)$, por ejemplo $\mathbf{w}'(\mathbf{x}; \sigma_I)$, proporciona la *orientación predominante* en \mathbf{x} , que es perpendicular a la *orientación predominante del gradiente*. Nótese que el gradiente de una función apunta en la dirección de máximo cambio, y la orientación predominante es perpendicular a esta dirección debido a que la anisotropía aparece en forma de valores de gris similares a lo largo de una orientación y grandes variaciones perpendiculares a ella.

Este análisis asume que en cada vecindad hay una sola orientación. Para verificar este supuesto, se introduce una *medida de confianza* normalizada: se asocia a cada orientación un valor real $C \in [0, 1]$, que puede obtenerse a partir de los autovalores del tensor estructural. La similaridad de los autovalores del tensor estructural implica isotropía y, como resultado, C debería tener un valor próximo a 0. Por lo tanto, denotando por $\lambda_1, \dots, \lambda_d$ los autovalores de \mathbf{M} , una opción lógica consiste en comprobar si la suma de las diferencias cuadráticas de los mismos:

$$\lambda_{\Delta}(\mathbf{x}; \sigma_I) = \sum_{i=1}^d \sum_{j=i+1}^d (\lambda_i(\mathbf{x}; \sigma_I) - \lambda_j(\mathbf{x}; \sigma_I))^2 \quad (2.8)$$

supera un umbral predefinido c característico para λ_{Δ} en la estructura que deseamos realizar. Una función adecuada es [Niessen y col., 1997]:

$$C(\mathbf{x}; \sigma_I; c) = 1 - e^{-\lambda_{\Delta}(\mathbf{x}; \sigma_I)^2 / 2c^2} \quad (2.9)$$

Como conclusión, para calcular las líneas cresta y las líneas valle de una imagen se siguen los siguientes pasos:

1. Cálculo del campo vectorial del gradiente \mathbf{w} y del campo del tensor estructural \mathbf{M} . Utilizaremos una vecindad que seguirá una distribución Gaussiana para calcular \mathbf{M} :

$$\mathbf{M}(\mathbf{x}; \sigma_I) = \begin{pmatrix} s_{11}(\mathbf{x}; \sigma_I; \sigma_D) & s_{12}(\mathbf{x}; \sigma_I; \sigma_D) \\ s_{12}(\mathbf{x}; \sigma_I; \sigma_D) & s_{22}(\mathbf{x}; \sigma_I; \sigma_D) \end{pmatrix} \quad (2.10)$$

$$\begin{aligned}
s_{11}(\mathbf{x}; \sigma_I; \sigma_D) &= G(\mathbf{x}; \sigma_I) * (L_x(\mathbf{x}; \sigma_D)L_x(\mathbf{x}; \sigma_D)) \\
s_{12}(\mathbf{x}; \sigma_I; \sigma_D) &= G(\mathbf{x}; \sigma_I) * (L_x(\mathbf{x}; \sigma_D)L_y(\mathbf{x}; \sigma_D)) \\
s_{22}(\mathbf{x}; \sigma_I; \sigma_D) &= G(\mathbf{x}; \sigma_I) * (L_y(\mathbf{x}; \sigma_D)L_y(\mathbf{x}; \sigma_D))
\end{aligned}$$

El nuevo parámetro σ_D denota la desviación estándar del núcleo Gaussiano implicado en el proceso de diferenciación necesario para calcular \mathbf{w} . El parámetro σ_D se denomina *escala de diferenciación*, frente a σ_I , que es la *escala de integración*. La escala de diferenciación se ajusta al tamaño de las estructuras cuya orientación se está buscando, mientras que la escala de integración se ajusta al tamaño de la vecindad en la que una orientación es dominante.

2. Análisis de los autovalores de \mathbf{M} . El autovector normalizado \mathbf{w}' correspondiente al mayor autovalor nos da la orientación predominante del gradiente. En el análisis con el tensor estructural, direcciones opuestas son tratadas por igual. Así, para recuperar la dirección se pone \mathbf{w}' en el mismo cuadrante 2D que \mathbf{w} . De esta forma obtenemos el nuevo campo vectorial $\tilde{\mathbf{w}}$:

$$\tilde{\mathbf{w}} = \text{signo}(\mathbf{w}'^t \cdot \mathbf{w})\mathbf{w}' \quad (2.11)$$

donde

$$\text{signo}(x) = \begin{cases} +1 & \text{si } x > 0 \\ -1 & \text{si } x < 0 \\ 0 & \text{si } x = 0 \end{cases} \quad (2.12)$$

De esta forma los vectores de atracción/repulsión se refuerzan.

3. Cálculo de la nueva medida de las crestas o valles utilizando el operador MLSEC-ST definido como:

$$\tilde{\kappa}_d = -\text{div}(\tilde{\mathbf{w}}) \quad (2.13)$$

4. Cálculo de la medida de confianza C que permite descartar crestas y valles que aparecen en zonas isotrópicas. De esta forma, $\tilde{\kappa}_d C$ tiene una menor respuesta que $\tilde{\kappa}_d$ en las regiones isotrópicas, y será lo que tomaremos como medida final de la cresta o valle.

La figura 2.7 ilustra el proceso de obtención de las crestas y valles sobre dos imágenes distintas, extraídas de dos secuencias pertenecientes a dos pacientes.

2.1.2. Alineamiento de las imágenes

Una vez que se han extraído las crestas y valles de los fotogramas que van a alinearse, el siguiente paso es llevar a cabo su alineamiento. La aproximación más directa para conseguirlo consiste en ejecutar un proceso iterativo de optimización de alguna función que mida el grado de superposición de ambas imágenes: se toma una imagen como referencia mientras que la otra se transforma iterativamente hasta que la función alcanza un máximo global. Una función adecuada para medir la calidad del alineamiento es la función de correlación, definida en su versión discreta como:

$$\text{Corr}_{\mathcal{T}} = f(x, y) \circ g(x, y) = \sum_{x, y \in f} f(x, y) \cdot g(\mathcal{T}(x, y))$$

donde \circ representa la operación de correlación, f y g son las imágenes con las crestas y valles de los fotogramas, (x, y) es un punto en ambas imágenes y \mathcal{T} representa la transformación cuyos parámetros estamos evaluando. Un punto clave consiste en no transformar todos los píxeles de la imagen, sino sólo aquellos cuyo valor supera un umbral fijo. Este procedimiento puede llegar a reducir hasta en un 95 % la carga computacional del algoritmo.

Este tiempo todavía puede reducirse más si se realizan los cálculos en el dominio transformado, especialmente si se utiliza la Transformada Rápida de Fourier, dominio en el que para calcular la correlación se puede emplear el teorema de la correlación [González y Woods, 1992, Oppenheim y col., 1996]. Dadas dos imágenes $f(x, y)$ y $g(x, y)$ con transformadas de Fourier $F(u, v)$ y $G(u, v)$ respectivamente, tenemos que:

$$f(x, y) \circ g(x, y) \Leftrightarrow F^*(\mathbf{u}, \mathbf{v})G(\mathbf{u}, \mathbf{v}) \quad (2.14)$$

donde $F^*(\mathbf{u}, \mathbf{v})$ representa el complejo conjugado de $F(\mathbf{u}, \mathbf{v})$, la transformada de $f(x, y)$. Esta implementación es mucho más manejable computacionalmente que la versión en el dominio espacial, y si se utiliza la transformada rápida de Fourier (que

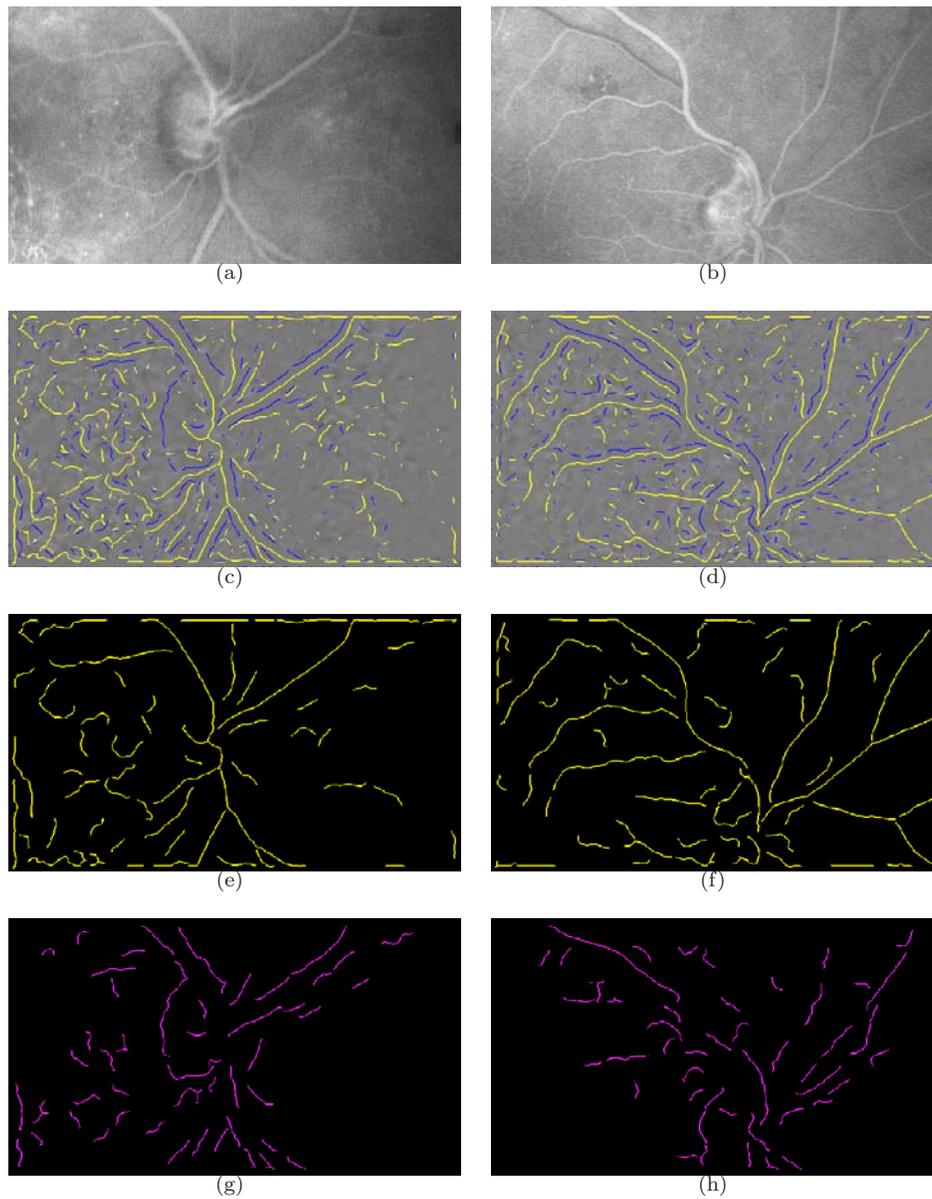


Figura 2.7: Proceso de obtención de las crestas y valles sobre dos imágenes extraídas de dos secuencias *SLO*. (a) y (b) muestran los fotogramas originales. En (c) y (d) se muestra el resultado obtenido utilizando el operador LSEC. A continuación se muestran las crestas ((e), (f)) y valles ((g) y (h)) obtenidos mediante el operador MLSEC-ST.

reduce la complejidad de $\mathcal{O}(M^2)$ a $\mathcal{O}(M \log_2 M)$, el tiempo de computación es realmente bajo, factor imprescindible en el problema que estamos tratando.

La función $\text{Corr}_{\mathcal{T}}$ junto con los cinco parámetros de la transformación (traslación en x e y , rotación y escalado en x e y) definen un espacio de búsqueda que es difícil de optimizar debido a que:

1. la función es no-monótona, es decir, posee muchos máximos locales,
2. el cálculo de la medida de similitud $\text{Corr}_{\mathcal{T}}$ acarrea una cierta carga computacional ya que implica la transformación de imágenes 2D (aunque este factor, como se acaba de comentar, se atenúa con la implementación antes descrita), y
3. los parámetros de traslación y rotación no pueden ser desacoplados para reducir la dimensionalidad del espacio de búsqueda.

Una aproximación que resuelve los dos primeros problemas consiste en realizar la búsqueda dentro del espacio de parámetros con múltiples resoluciones. La multiplicidad de resoluciones se consigue construyendo dos pirámides, como se propone en [Van den Elsen y col., 1995], donde las imágenes de crestas y valles de los fotogramas *SLO* están en la base de las pirámides, y cada nivel es una versión calculada a partir de la anterior pero con la mitad de resolución, continuando hacia la cima hasta que las imágenes tienen un tamaño de aproximadamente 64 píxeles (aunque este parámetro del modelo, el tamaño mínimo de las imágenes, es ajustable). La figura 2.8 esquematiza el proceso de alineamiento de dos imágenes utilizando este método.

Tomando como referencia la figura 2.8 vamos a analizar cada uno de los pasos que se siguen en el algoritmo de alineamiento:

Paso 1

El primer paso del algoritmo de alineamiento consiste en extraer las crestas y valles de los dos fotogramas que se va a alinear mediante el método anteriormente descrito.

Paso 2

En el paso 2 se construyen dos pirámides de imágenes a partir de las imágenes de características obtenidas en el paso 1. En estas pirámides, cada nivel se construye

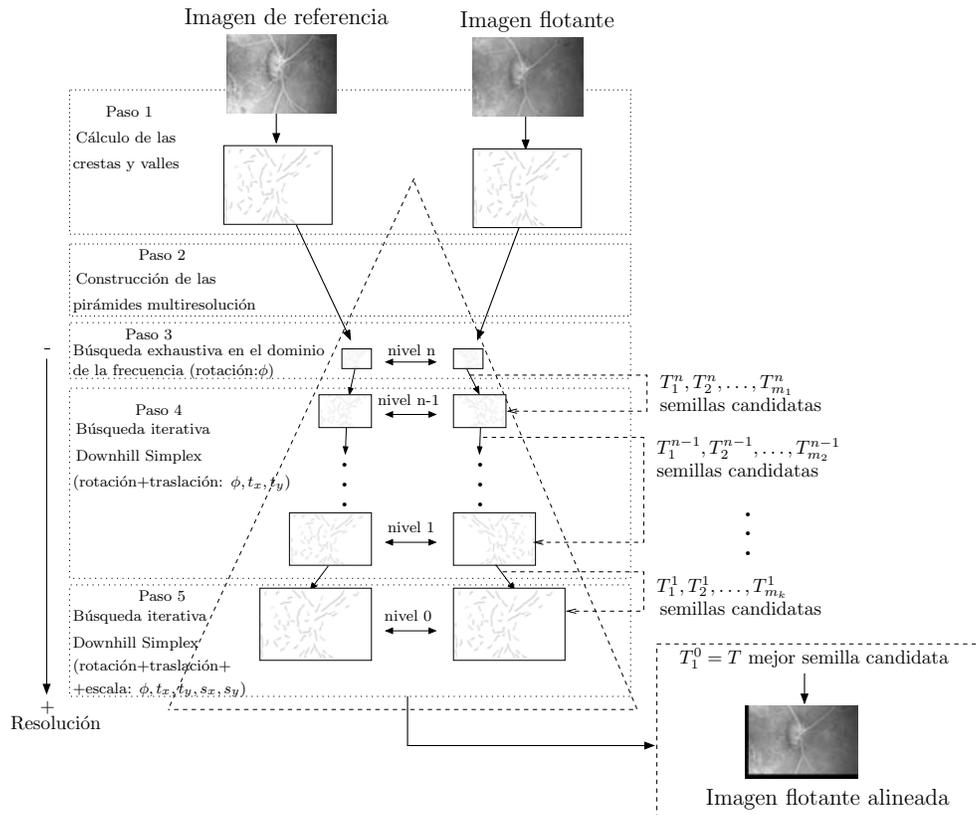


Figura 2.8: Alineamiento utilizando el método basado en la extracción de crestas. En un primer paso se extraen las crestas y valles de las imágenes de referencia y flotante. A continuación (paso 2) se construyen las dos pirámides multiresolución que se utilizarán en el alineamiento. Cada nivel de la pirámide contiene una imagen con la mitad de resolución que el nivel anterior, por lo que en la cima de la pirámide las imágenes tienen la mínima resolución, y así es factible realizar una búsqueda más exhaustiva (paso 3), que proporcionará el primer conjunto de transformaciones semilla, que servirán en el siguiente nivel para iniciar la búsqueda. En los niveles $n-1$ a 0 la búsqueda se realiza partiendo de las semillas del nivel superior y utilizando el algoritmo Downhill Simplex (paso 4) con traslación y rotación como parámetros, que proporcionará un nuevo conjunto de semillas que a su vez se irán pasando a los niveles inferiores, reduciendo el número de las mismas hasta que, en el nivel 0 se obtenga una sola semilla (paso 5), pero con el escalado incorporado, dando un resultado más preciso.

tomando como base la imagen del nivel anterior (el nivel 0 está compuesto por las imágenes de características originales) y realizando un submuestreo de la misma, de forma que se reduce el tamaño a la mitad.

A la hora de realizar el submuestreo, se analizaron varias posibilidades: tomar el píxel central de una vecindad, tomar el valor medio o tomar el valor mayor. Si se tiene en cuenta que estamos trabajando sobre las imágenes de crestas y valles y que estas tienen 1 píxel de ancho, lo más importante es no perder esta información, por lo que la opción elegida finalmente fue la de tomar el mayor valor de la ventana de píxeles correspondiente (figura 2.9), para así no eliminar ninguna cresta o valle existente.

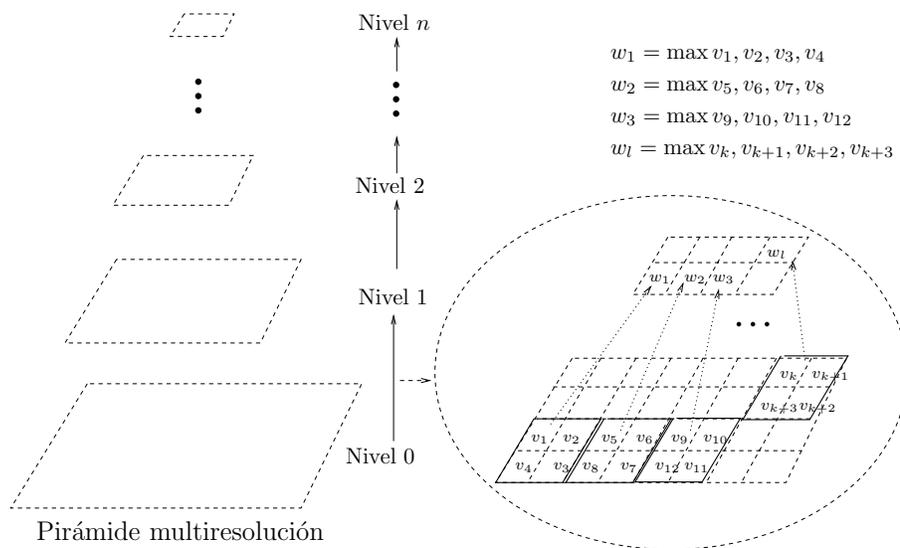


Figura 2.9: Representación del proceso de generación de las imágenes que constituyen los distintos niveles de la pirámide multiresolución. Cada nivel de la pirámide está formado por una versión muestreada del nivel inferior, en el que se toma el mayor valor de cada 4 píxeles, preservando así las posibles crestas y valles que existieran en esa vecindad.

Paso 3

El proceso de búsqueda comienza en la parte superior de las pirámides, donde el pequeño tamaño de las imágenes permite llevar a cabo una búsqueda más exhaustiva. En este nivel no buscamos una solución precisa, sino un conjunto de soluciones de partida o soluciones semilla que nos lleven a la solución final.

La aproximación más trivial consiste en iterar sobre todo el conjunto de posibles parámetros seleccionando los que proporcionen mejores valores para la función de ajuste. Sin embargo, esta solución implica un coste computacional excesivo por el alto número de intentos que hay que realizar. Tal y como hemos comentado antes, una aproximación mucho más eficiente consiste en operar en el dominio de la Transformada de Fourier. El teorema de correlación permite calcular n^3 traslaciones con el coste computacional de una única multiplicación. Utilizando esta solución, sólo es necesario transformar las imágenes con el ángulo de rotación.

$$f(x, y) \circ g(x, y) \Leftrightarrow F^*(\mathbf{u}, \mathbf{v})G(\mathbf{u}, \mathbf{v}) \quad (2.15)$$

donde $F^*(\mathbf{u}, \mathbf{v})$ representa el complejo conjugado de $F(u, v)$, la transformada de $f(x, y)$.

A la hora de implementar esta búsqueda en el dominio de Fourier, surgieron dos problemas. El primero es que, para poder calcular su transformada rápida de Fourier, las imágenes han de tener un tamaño potencia de dos. Además, si las imágenes no tenían un tamaño cuadrado múltiplo de 2 tanto en alto como en ancho, aparecían valores inesperados que llevaban a falsos positivos en la búsqueda. Para resolverlo, antes de llevar a cabo la búsqueda, es necesario ampliar las imágenes de referencia y flotante hasta un tamaño potencia de 2. El esquema de la figura 2.10 ilustra el método empleado para la búsqueda del primer conjunto de valores de transformación (valores semilla) que se propagarán a los niveles inferiores de la pirámide.

Al final de este proceso obtendremos la transformación que optimiza el valor de la correlación en este nivel de la pirámide. En la tabla 2.1 se puede apreciar cómo varía el tiempo de alineamiento empleado al variar el número de niveles de la pirámide, y con ello el tamaño final de las imágenes en el nivel superior de la misma. Estos tiempos fueron obtenidos para un par de fotogramas de tamaño 692×394 píxeles. La primera columna de dicha tabla indica el número de niveles que componían cada una de las pirámides multiresolución, mientras que la segunda columna indica el tamaño en píxeles (ancho o alto, pues como hemos indicado anteriormente las imágenes han de ser cuadradas para poder operar con la transformada rápida de Fourier), de las imágenes en el nivel superior de la pirámide. Las tres siguientes columnas indican los

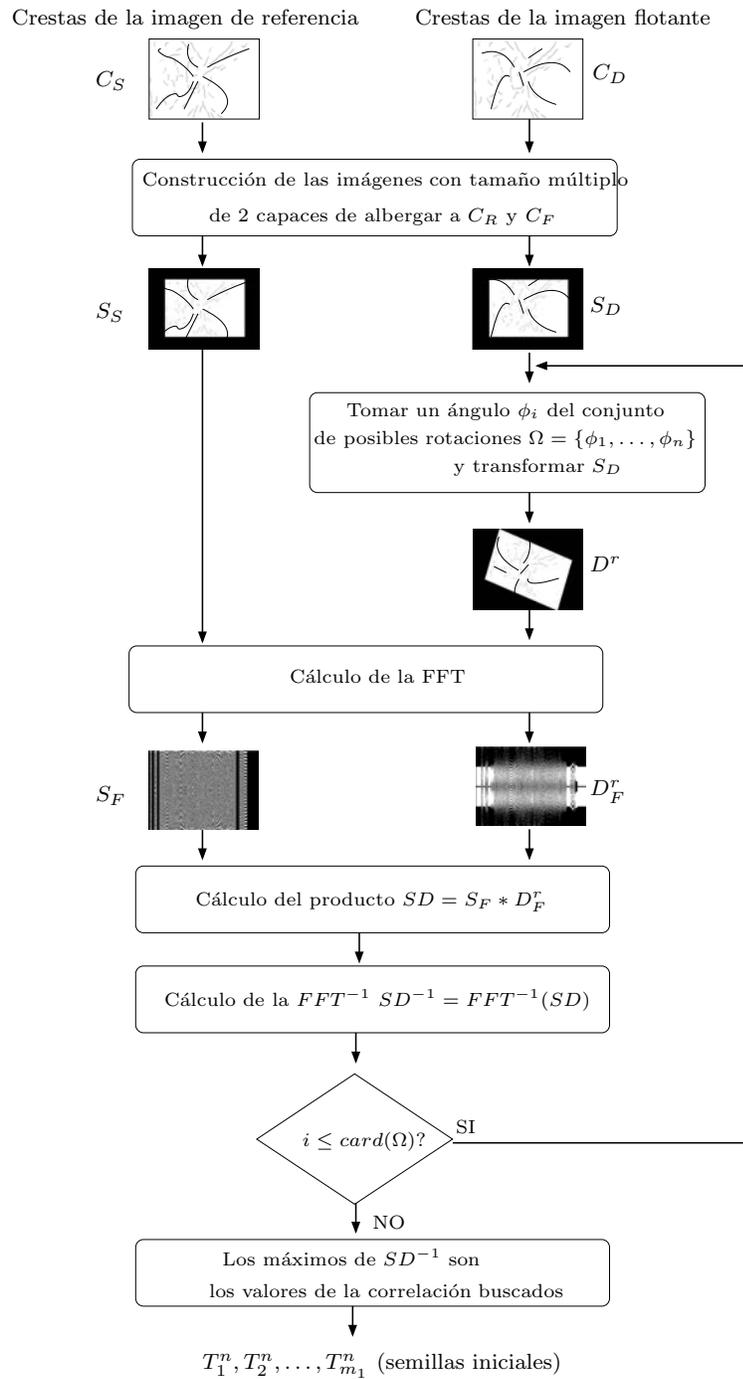


Figura 2.10: Esquema que muestra la técnica empleada para la búsqueda de semillas iniciales en el nivel superior de la pirámide multiresolución

tiempos necesarios, en segundos, para calcular la transformada rápida de Fourier, la imagen flotante transformada y el tiempo total de alineamiento de ambas imágenes, respectivamente.

Niveles	Tam.imágenes último nivel	t.FFT (s)	t.Transformación (s)	t.Alineamto. (s)
1	1024	0,0686	6,8137	61,9556
2	512	0,0168	1,3518	15,9669
3	256	0,0042	0,2083	4,4794
4	128	0,0010	0,0313	2,7393
5	64	0,0001	0,0013	2,5706

Tabla 2.1: Tiempos obtenidos al variar el número de niveles de la pirámide multiresolución, variando con ello el tamaño de las imágenes en el nivel más alto, en el que se lleva a cabo una búsqueda más exhaustiva en el dominio de la Transformada de Fourier. La primera columna indica el número de niveles de la pirámide. La segunda indica el tamaño de las imágenes del nivel superior, y las tres últimas indican los tiempos (en segundos) necesarios para calcular la transformada de Fourier, la imagen flotante transformada y el tiempo total de alineamiento, respectivamente.

Sin embargo, la mejor transformación en cada nivel no es necesariamente la que lleva al mejor resultado final, debido a que la aproximación jerárquica introduce falsos máximos. Es por ello que, para que el método sea más robusto, se toman los mejores valores de cada nivel que denominaremos semillas (denotadas en la figura 2.8 por $\{T_1^n, T_2^n, \dots, T_{m_1}^n\}$ en el nivel n , $\{T_1^{n-1}, T_2^{n-1}, \dots, T_{m_2}^{n-1}\}$ en el nivel $n - 1$ hasta $\{T_1^1, T_2^1, \dots, T_{m_k}^1\}$ en el nivel 1 y $\{T\}$ en el nivel 0), como puntos de inicialización del proceso de búsqueda en el nivel inmediatamente inferior.

Hemos seleccionado los siguientes valores para la búsqueda en este nivel: se muestra la pirámide hasta el nivel 5, para el proceso de búsqueda se considera como variable la rotación de las imágenes tomando 10 ángulos de rotación como conjunto de valores a probar, con una separación entre ellos de 0,1 radianes.

Si se muestreara un nivel más la pirámide hasta llegar al nivel 6 las imágenes tendrían un tamaño de 21×24 píxeles de ancho y alto respectivamente, y como

muestra la figura 2.11 con esta resolución el nivel de detalle es demasiado bajo y se pierden demasiadas crestas y valles, lo que unido a que aumentarían los tiempos de cálculo, nos lleva a descartar el aumento a 6 niveles.

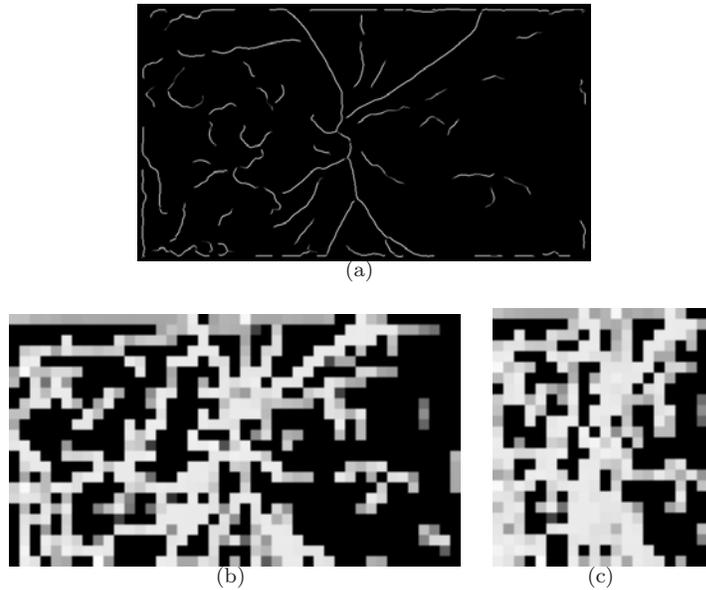


Figura 2.11: Muestra de las imágenes en los niveles superiores de la pirámide con 6 niveles. En la fila superior se muestra la imagen original de crestas y valles, que constituirá el nivel más bajo de la pirámide, mientras que en la fila inferior se muestran las imágenes de los niveles 5 (b) y 6 (c), tras haberlas aumentado de tamaño hasta cuadruplicar el original. Se aprecia claramente cómo la imagen del nivel 6 (c) pierde gran parte de las características necesarias para que la búsqueda en el nivel superior resulte útil. Sin embargo, no ocurre esto con la imagen del nivel 5 (b), que será la del nivel superior de la pirámide si se consideran sólo 5 niveles.

La tabla 2.2 muestra la evolución del proceso de búsqueda en el nivel superior de la pirámide para cada una de los 10 ángulos que se toman como punto de partida, y las 8 semillas que se toman en el nivel superior (ver tabla 2.3, incluyendo el número de semillas por nivel).

Puede apreciarse cómo en la lista de mejores soluciones se almacenan los ángulos para los que la correlación proporciona unos valores más altos. En el ejemplo, a partir de los 0° no mejora el resultado, por lo que los parámetros de la solución se mantienen.

Rotación aplicada (radianes)	Conjunto mejores soluciones (t_x, t_y, α)	Correlación
-0.4363	(6.0,2.0,-0.4363)	8017662.5
	(-1.0,-10.0,-0.4363)	8038701.0
	(-1.0,13.0,-0.4363)	8073998.0
	(-1.0,0.0,-0.4363)	8080737.5
	(0.0,1.0,-0.4363)	8132846.0
	(-1.0,-11.0,-0.4363)	8134517.0
	(-12.0,5.0,-0.4363)	8301839.0
	(6.0,1.0,-0.4363)	8474340.0
-0.3490	(2.0,-1.0,-0.3490)	8391516.0
	(2.0,12.0,-0.3490)	8455127.0
	(6.0,1.0,-0.4363)	8474340.0
	(2.0,11.0,-0.3490)	8508669.0
	(2.0,0.0,-0.3490)	8703538.0
	(2.0,3.0,-0.3490)	8783170.0
	(2.0,2.0,-0.3490)	8815636.0
	(2.0,1.0,-0.3490)	8893733.0
-0.2617	(6.0,1.0,-0.2617)	8829371.0
	(1.0,4.0,-0.2617)	8840798.0
	(-1.0,3.0,-0.2617)	8841495.0
	(2.0,1.0,-0.3490)	8893733.0
	(6.0,2.0,-0.2617)	8921727.0
	(0.0,4.0,-0.2617)	8923463.0
	(-1.0,5.0,-0.2617)	8936563.0
	(-1.0,4.0,-0.2617)	9194510.0
-0.1745	(4.0,2.0,-0.1745)	9312256.0
	(2.0,-10.0,-0.1745)	9365308.0
	(3.0,5.0,-0.1745)	9503752.0
	(0.0,0.0,-0.1745)	9550701.0
	(3.0,4.0,-0.1745)	9633020.0
	(2.0,-8.0,-0.1745)	9807047.0
	(0.0,1.0,-0.1745)	9869814.0
	(2.0,-9.0,-0.1745)	9916861.0

Tabla 2.2: sigue en la página siguiente

Rotación aplicada (radianes)	Conjunto mejores soluciones (t_x, t_y, α)	Correlación
-0.0872	(-1.0,-5.0,-0.0872)	10378206.0
	(1.0,-8.0,-0.0872)	10451900.00
	(1.0,-7.0,-0.0872)	10568260.00
	(-1.0,-6.0,-0.08726)	10657272.0
	(3.0,-1.0,-0.0872)	10659476.0
	(0.0,-6.0,-0.0872)	10968344.0
	(0.0,-8.0,-0.0872)	11493528.0
	(0.0,-7.0,-0.0872)	11512715.0
0.0	(3.0,-8.0,0.0)	13674249.0
	(2.0,-4.0,0.0)	14279848.0
	(2.0,-7.0,0.0)	15451454.0
	(2.0,-5.0,0.0)	18360112.0
	(3.0,-5.0,0.0)	19558148.0
	(2.0,-6.0,0.0)	19680204.0
	(3.0,-7.0,0.0)	22253336.0
	(3.0,-6.0,0.0)	26864596.0
0.0872	(3.0,-8.0,0.0)	13674249.0
	(2.0,-4.0,0.0)	14279848.0
	(2.0,-7.0,0.0)	15451454.0
	(2.0,-5.0,0.0)	18360112.0
	(3.0,-5.0,0.0)	19558148.0
	(2.0,-6.0,0.0)	19680204.0
	(3.0,-7.0,0.0)	22253336.0
	(3.0,-6.0,0.0)	26864596.0

Tabla 2.2: sigue en la página siguiente

Rotación aplicada (radianes)	Conjunto mejores soluciones (t_x, t_y, α)	Correlación
0.1745	(3.0,-8.0,0.0)	13674249.0
	(2.0,-4.0,0.0)	14279848.0
	(2.0,-7.0,0.0)	15451454.0
	(2.0,-5.0,0.0)	18360112.0
	(3.0,-5.0,0.0)	19558148.0
	(2.0,-6.0,0.0)	19680204.0
	(3.0,-7.0,0.0)	22253336.0
	(3.0,-6.0,0.0)	26864596.0
0.2617	(3.0,-8.0,0.0)	13674249.0
	(2.0,-4.0,0.0)	14279848.0
	(2.0,-7.0,0.0)	15451454.0
	(2.0,-5.0,0.0)	18360112.0
	(3.0,-5.0,0.0)	19558148.0
	(2.0,-6.0,0.0)	19680204.0
	(3.0,-7.0,0.0)	22253336.0
	(3.0,-6.0,0.0)	26864596.0
0.3490	(3.0,-8.0,0.0)	13674249.0
	(2.0,-4.0,0.0)	14279848.0
	(2.0,-7.0,0.0)	15451454.0
	(2.0,-5.0,0.0)	18360112.0
	(3.0,-5.0,0.0)	19558148.0
	(2.0,-6.0,0.0)	19680204.0
	(3.0,-7.0,0.0)	22253336.0
	(3.0,-6.0,0.0)	26864596.0

Tabla 2.2: Valor de las transformaciones a través del proceso de búsqueda en el nivel superior de la pirámide. La primera columna indica el ángulo de rotación que se aplica a la imagen dinámica. La segunda columna indica el conjunto de los 8 mejores valores semilla encontrados hasta ese momento, mientras que la tercera indica el valor de la correlación de estas soluciones.

Paso 4

Una vez obtenidas las semillas del nivel superior a partir de la búsqueda más exhaustiva en el dominio de Fourier, se continúa descendiendo en la pirámide para refinar estos resultados iniciales, ejecutando en cada nivel una búsqueda hasta llegar al nivel de máxima resolución, cuya convergencia nos proporcionará el valor máximo de la correlación. La figura 2.12 representa los procesos que se llevan a cabo en este paso.

Para llevar a cabo la optimización de los valores de la correlación se han analizado diferentes métodos de maximización de funciones, como son el método iterativo denominado *Downhill Simplex* [Press y col., 1992], el método estocástico denominado *Temple simulado* [Press y col., 1992] o el algoritmo de Powell en su variante multi-dimensional [Press y col., 1992], aunque los resultados obtenidos con el primero de ellos fueron los mejores a la hora de llevar a cabo la maximización de la correlación (sobre todo a nivel computacional, ya que el segundo es un algoritmo no determinista y el tercero, además de ser también más lento, presenta un peor comportamiento ante máximos locales). En cada nivel el proceso se repite partiendo de las semillas obtenidas en el nivel superior, y termina cuando el algoritmo alcanza una zona en la que la diferencia entre los valores máximo y mínimo en un entorno no superan un umbral dado. El valor de tolerancia (umbral en la diferencia entre los valores máximo y mínimo) y el número de semillas determinan en gran medida la robustez del método así como el tiempo de cálculo final.

Paso 5

Finalmente, en el nivel 0 de la pirámide, además de la búsqueda se realiza un proceso posterior de refinamiento de la solución (figura 2.13), tratando de ajustar además de las traslaciones y rotación, la escala de ambas imágenes.

La tabla 2.3 muestra un ejemplo de la traza de las semillas obtenidas en los niveles de la pirámide para dos fotogramas de una de las secuencias mostrados en la figura 2.14, con una pirámide de 5 niveles (el nivel superior en el que se ejecuta la búsqueda más exhaustiva se muestra en la tabla 2.2). Además, la segunda fila de la figura 2.14 muestra dos perspectivas distintas del mapa de la correlación que se obtiene desplazando la imagen 2.14(b) en el rango ± 316 píxeles y posteriormente calculando

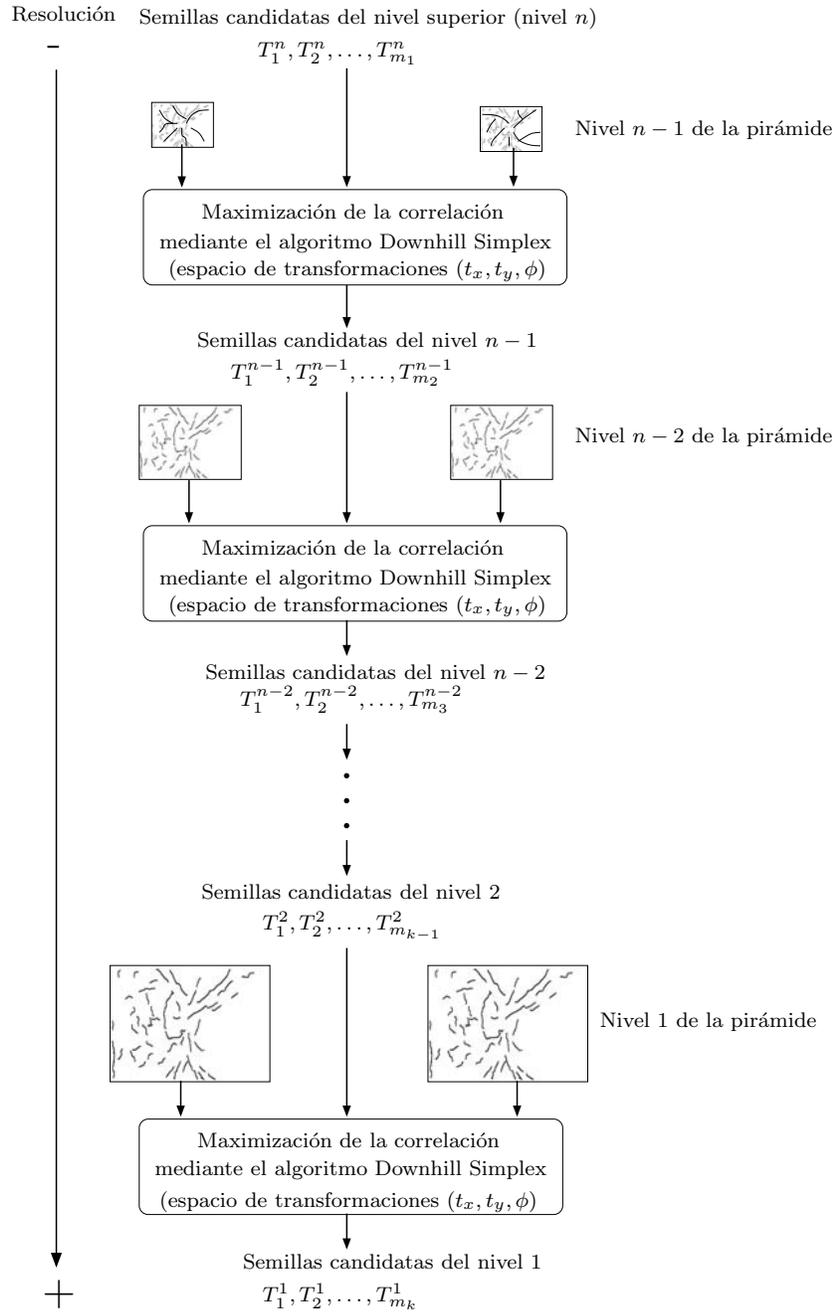


Figura 2.12: Esquema que ilustra el proceso de obtención de los valores de las transformaciones (sólo con rotación y traslación en estos niveles) que optimizan el valor de la correlación en los niveles intermedios de la pirámide, mostrando asimismo la propagación de dichos valores semilla de un nivel al inmediatamente inferior en la pirámide multiresolución, en el que se utilizarán como puntos de partida del proceso de ajuste.

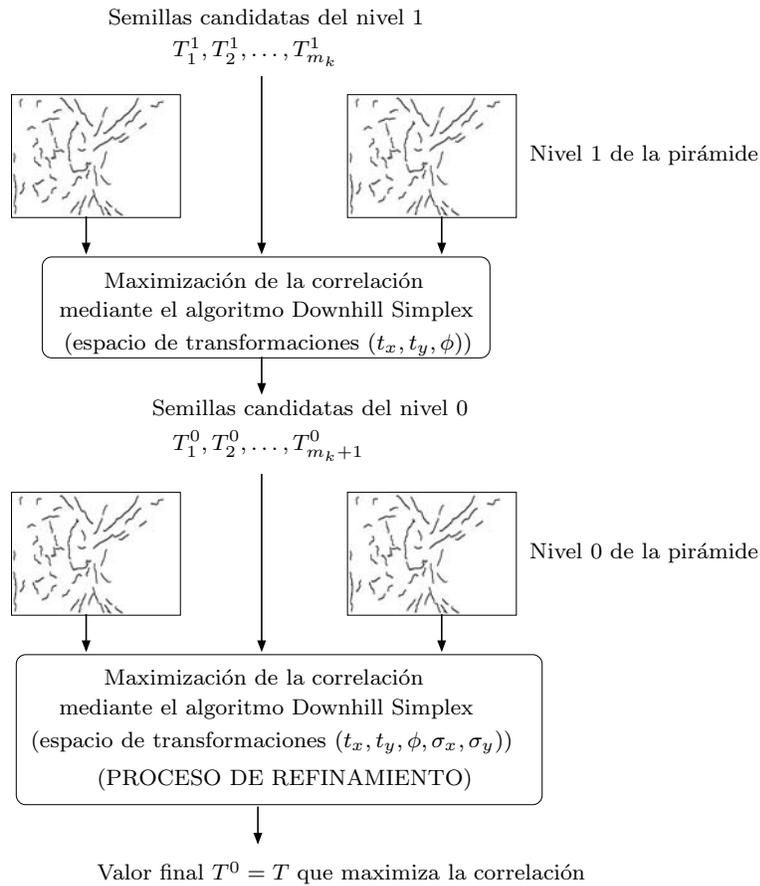


Figura 2.13: Esquema que ilustra el proceso de obtención de los valores de las transformaciones que optimizan el valor de la correlación en el nivel 0 de la pirámide, incluyendo además el proceso de refinamiento que tiene lugar al final, y en el que se incluye la escala (en x e y) en el espacio de búsqueda.

la correlación cruzada entre la imagen y la flotante transformada.

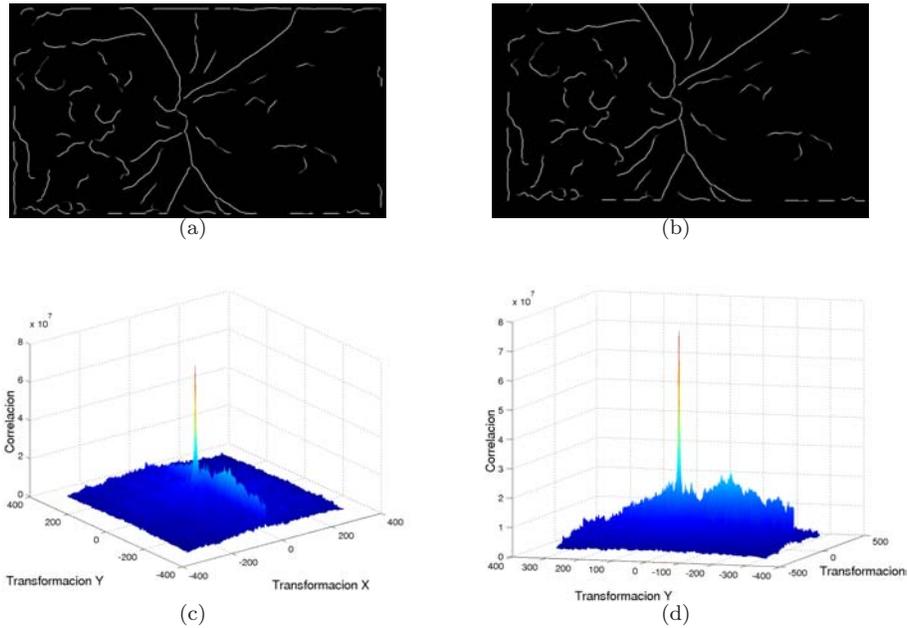


Figura 2.14: Ejemplo de imágenes de crestas y los valores que se obtienen al correlacionar una de ellas con la otra transformada. (a) y (b) muestran las imágenes de crestas que se emplearán para ilustrar el funcionamiento del método Downhill Simplex, y entre las cuales existe, como muestra la tabla 2.3, una diferencia de 22 píxeles en el eje X y 24 píxeles en el eje Y. La rotación que las ajusta perfectamente es prácticamente nula (0.000776 grados). En (c) y (d) se muestran dos perspectivas del mapa con los valores de la correlación obtenidos al transformar la imagen (b) en el rango ± 316 píxeles para posteriormente calcular la correlación entre el resultado de dicha transformación y la imagen (a).

Para llevar a cabo el proceso de maximización de la función de correlación se ha seleccionado el método iterativo Downhill Simplex, propuesto por Nelder y Mead [Nelder y Mead, 1965], e implementado como se propone en [Press y col., 1992]. Esta maximización iterativa se lleva a cabo en todos los niveles de la pirámide excepto en el superior, en el que, como ya se ha comentado, se lleva a cabo una búsqueda más exhaustiva en el dominio de la frecuencia. La búsqueda en cada nivel comienza en los puntos dados por las semillas que se obtienen en el nivel inmediatamente superior, y termina cuando el algoritmo alcanza una zona en la que la diferencia entre los valores máximo y mínimo en un entorno de cada semilla son inferiores a un umbral.

Nivel de la pirámide	Tamaño de las imágenes	# semillas generadas	Valores de las semillas ($t_x, t_y, \alpha, s_x, s_y$)	Correlación ($\times 10^3$)
Nivel 3	86×59	6	(2.9912,-6.0027,0.0001)	27,311.550
			(2.9935,-5.9999,0.0000)	27,336.866
			(2.9970,-6.0048,0.0001)	27,339.314
			(2.9962,-6.0019,0.0000)	27,350.308
			(2.9998,-6.0015,0.0001)	27,365.276
			(2.9986,-5.9985,0.0000)	27,367.860
Nivel 2	173×98	4	(5.0428,-5.9717,0.0012)	32,728.598
			(5.0250,-5.9899,0.0007)	32,777.808
			(5.0078,-6.0028,0.0002)	32,800.640
			(5.0073,-5.9911,0.0002)	32,812.182
Nivel 1	346×197	2	(11.0160,-12.0584,0.0028)	53,384.816
			(10.9981,-12.0460,0.0022)	53,392.008
Nivel 0	692×394	1	(22.2278,-24.4210,0.0049)	85,935.328
Nivel 0 (refinamiento)	692×394	1	(20.6620,-24.3768,0.0035, 0.9955,1.0004) (solución final)	90,191.584

Tabla 2.3: Traza obtenida por la ejecución del proceso de alineamiento de un par de imágenes (figura 2.14), mostrando las semillas extraídas en cada uno de los niveles de la pirámide multiresolución, excepto el nivel superior, en el que se ejecuta una búsqueda más exhaustiva, que se mostró en la tabla 2.2. Nótese que el parámetro de escalado no se considera hasta el paso de refinamiento en el nivel 0.

El apéndice C describe con más detalle este método de optimización.

2.1.3. Evaluación del método de alineamiento basado en crestas

El procedimiento habitual en el Servicio de Oftalmología del *Complejo Hospitalario Universitario de Santiago de Compostela (CHUS)* para adquirir secuencias de *SLO* comienza con la inyección de 3 ml de fluoresceína sódica en la vena antecubital del brazo derecho del paciente. A continuación, el oftalmoscopio (*SLO-101*, Rodenstock con láser de Argon, 488 nm) en modo de fluoresceína, comienza la adquisición de las imágenes. Estas se digitalizan mediante una tarjeta digitalizadora Matrox RT5000 en un PC convencional a una velocidad de 25 fotogramas por segundo durante unos 40 segundos, con una resolución de 720×576 píxeles y 256 niveles de gris. De aquí se puede deducir que el número de fotogramas que componen una secuencia ronda los 10,000.

Tal y como se mostró en la figura 1.3, los fotogramas de las secuencias de *SLO* presentan características muy diferentes, dependiendo del momento en que se capture la imagen tomando como referencia el instante en el que se inyecta el contraste: los fotogramas iniciales apenas poseen contraste. Según avanza el tiempo el nivel de contraste aumenta, y con él la iluminación en el fondo del ojo, permitiendo observar muchos más detalles, hasta que se satura el nivel de contraste. En base a esto, hemos dividido la secuencia en tres secciones tal y como muestra la figura 2.15: una primera sección en la que los fotogramas tienen poco contraste, una sección intermedia en la que aumenta la cantidad de fluoresceína y se aprecian más claramente las venas, arterias y demás elementos del fondo del ojo, y una etapa final en la que se satura el nivel de contraste. Si el nivel de fluoresceína es alto, será posible extraer las crestas correspondientes a los vasos, y podrá utilizarse con éxito un proceso basado en la extracción de características. La tabla 2.4 muestra los resultados obtenidos al aplicar el método de registro basado en crestas y valles a cinco secuencias de *SLO* diferentes, cada una de ellas dividida en las 3 secciones que hemos comentado. Los fotogramas localizados en la mitad de cada área se seleccionan como fotogramas de referencia, ya que son los más representativos de ese entorno, y es necesario que el fotograma de

referencia, contra el que se alinearán todos los demás fotogramas, sea representativo de la sección a la que pertenece, para que el alineamiento de los fotogramas sea lo más preciso posible.

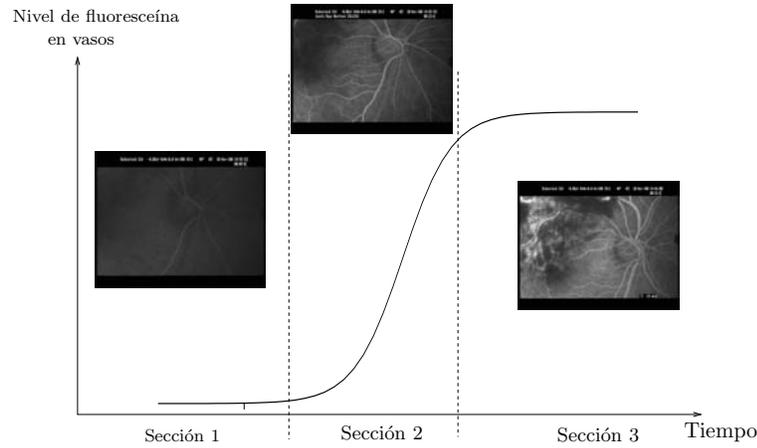


Figura 2.15: Las secuencias *SLO* se dividen en tres secciones, caracterizadas por el nivel de fluoresceína presente en el ojo en cada una de ellas

A partir de los datos que se muestran en la tabla 2.4 se deduce que el método basado en crestas y valles no obtiene buenos resultados en la sección 1, en la que el nivel de fluoresceína es todavía bajo, por lo que es difícil obtener unas crestas y valles que permitan una alineación buena, y así el Error Cuadrático Medio (ECM) es alto y el porcentaje de éxitos en el alineamiento es bajo. Esto puede apreciarse perfectamente en las crestas y valles obtenidos en tres imágenes (figura 2.16). Este hecho nos lleva a considerar que el método basado en crestas no va a ser suficiente para ajustar las imágenes en las que el contraste sea bajo, por lo que será necesaria una alternativa, representada por el método de alineamiento basado en la intensidad que se describe en el siguiente capítulo, para así poder alinear todos los fotogramas de una secuencia, con independencia de la posición del fotograma en la secuencia (y consecuentemente con independencia del nivel de fluoresceína presente en el árbol arteriovenoso en ese instante).

Para la evaluación del método de alineamiento basado en crestas, se ha diseñado un banco de pruebas que se ejecutó para 5 pacientes distintos. El banco de pruebas consiste en ejecutar el proceso de ajuste con 200 fotogramas de cada una de las seccio-

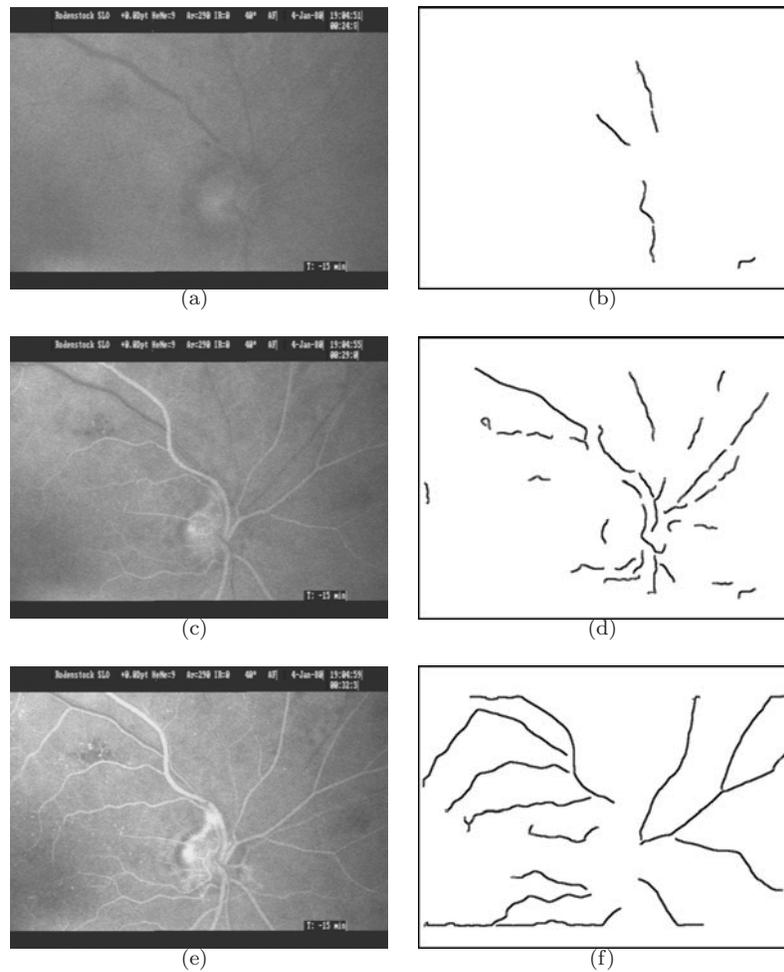


Figura 2.16: Muestras de los fotogramas representativos de cada una de las secciones de la figura 2.15 (columna de la izquierda) y sus correspondientes crestas extraídas (a la derecha): (a) representa a los fotogramas de la sección 1, (c) es un fotograma representativo de la sección 2 y (e) es un fotograma representativo de la sección 3. La figura muestra cómo la información de las crestas y valles extraídos es suficiente para conseguir una buena alineación en las secciones 2 y 3 (reflejado en las imágenes de crestas obtenidas (d) y (f)) pero no para la sección 1 (imagen (b)).

	Sección 1			Sección 2			Sección 3		
	Transf. éxito (%)	ECM	\bar{t} (s)	Transf. éxito (%)	ECM	\bar{t} (s)	Transf. éxito (%)	ECM	\bar{t} (s)
A	16	17.6	4.4	86	6.2	2,7	98	0.4360	2,5
B	10	12.9	4.4	80	5.3	2,1	91	0.1123	2,5
C	8	18.9	4.5	82	7.3	2,7	95	0.5432	2,5
D	11	10.3	4.6	79	6.5	2,8	95	0.1643	2,4
E	15	9.2	4.5	75	7.1	2,7	98	0.9523	2,5

Tabla 2.4: Estadísticas del experimento realizado para la evaluación de la alineación basada en crestas y valles para cinco secuencias de *SLO* pertenecientes a cinco pacientes distintos (*A – E*): la primera columna de cada sección representa el porcentaje de transformaciones recuperadas correctamente, la segunda representa el error cuadrático medio de las transformaciones obtenidas y finalmente la tercera indica el tiempo medio necesario (en segundos) para realizar el alineamiento de las imágenes.

nes en las que se dividen las secuencias *SLO* (ver figura 2.15) dependiendo del nivel de fluoresceína presente en las venas. A cada uno de estos fotogramas seleccionados se le aplicaron 50 transformaciones aleatorias, con valores de ± 100 píxeles de traslación y $\pm 5^\circ$ de rotación. Estas cotas para las transformaciones aleatorias representan los valores máximos presentes en los fotogramas de las distintas secuencias de *SLO* analizadas.

Los resultados obtenidos con este experimento se encuentran representados en la tabla 2.4, que muestra en su primera columna el porcentaje de transformaciones recuperadas con éxito, en la segunda columna incluye el error cuadrático medio (ECM) de las transformaciones obtenidas y en la tercera se muestra el tiempo medio necesario para realizar el alineamiento de cada par de imágenes (en segundos). Para medir el desalineamiento de las imágenes después del proceso de ajuste se seleccionan manualmente cuatro puntos en la imagen dinámica, de forma que estos se utilizan como puntos de referencia.

El error de alineamiento para cada fotograma se define como:

$$RE = \frac{1}{4} \sum_{j=1}^4 D(T(x_j), x_j) \quad (2.16)$$

donde D representa la distancia Euclídea y $T(x_j)$ es el punto x_j después de aplicarle la transformación obtenida T . Así, la primera columna de la tabla 2.4 muestra el porcentaje de alineamientos recuperados cuya distancia a la transformación aplicada es igual o inferior a ± 3 píxeles. El error considerado de ± 3 píxeles se considera aceptable debido a que no afecta al objetivo final del proceso, que es el cálculo de las curvas de dilución, pues este proceso conlleva un ajuste robusto a una curva, y por tanto no se verá alterado por esta pequeña variación.

La primera de las secciones presenta unos niveles de error muy elevados, de lo cual se deduce que el método basado en crestas no permite resolver el problema del alineamiento de las imágenes en esta sección. Esto viene motivado por el hecho de que las imágenes en esta etapa tienen muy poco contraste, lo que dificulta en gran medida la extracción de las características necesarias para el alineamiento. Sin embargo, en las otras dos secciones se puede apreciar cómo se incrementa notablemente el porcentaje de imágenes alineadas correctamente, sobre todo en la sección 3, en la que los niveles de fluoresceína son más altos. El método de alineamiento basado en crestas demuestra así ser muy adecuado para el alineamiento de los fotogramas de las secciones 2 y 3, no sólo por los altos porcentajes de éxito y bajo Error Cuadrático Medio (ECM), sino también por el reducido tiempo necesario para el ajuste de cada par de imágenes, como muestran las columnas con los tiempos medios por secuencia en la tabla 2.4. Pero tal y como veremos en el capítulo 4, es necesario tener también alineados los fotogramas de la sección 1, ya que para estimar las curvas de dilución se utiliza un conjunto de puntos extraídos de los fotogramas de las tres secciones. Si este conjunto de puntos no proviene de una secuencia bien alineada, el ajuste de la curva no será lo suficientemente preciso, y con ello cualquier medida que extraigamos de ella tendrá un error elevado. Veremos en el siguiente capítulo cómo resolver el problema del alineamiento de los fotogramas de la sección 1 de las secuencias de *SLO*.

Además, tal y como se mostrará en el capítulo 4 en el que analizaremos la metodología del proceso de alineamiento de las secuencias completas, incluido, como veremos en el capítulo siguiente el de los fotogramas de la sección 1, el reducido número de

fotogramas que quedan mal alineados en las secciones 2 y 3 no influirán en el resultado final. Esto se conseguirá empleando una técnica robusta de ajuste que eliminará estos fotogramas de los cálculos de las curvas de dilución, objetivo final del proceso de alineamiento de las secuencias de *SLO*.

Capítulo 3

Método de alineamiento basado en Información Mutua

En el capítulo anterior hemos visto cómo ante la ausencia de información relevante en las imágenes, con los métodos de alineamiento basados en la extracción de puntos característicos se obtiene un error elevado. Con objeto de solucionar este problema hemos decidido incorporar una nueva técnica que utiliza como medida del alineamiento una medida de información basada en la reducción de la incertidumbre de una variable debido al conocimiento de la otra. Esta medida, denominada *Información Mutua* es un concepto clásico de la teoría de la información que mide la cantidad de información que una variable contiene sobre la otra o la reducción de la incertidumbre de una variable debido al conocimiento de la otra. Este estadístico se derivó inicialmente en los estudios de teoría de la comunicación [Shannon, 1948, Shannon y Warren, 1963] como una medida de información entre el emisor y el receptor en cualquier punto del canal de comunicación, aunque aquí veremos cómo se puede utilizar para realizar el alineamiento de imágenes. En el apéndice A se describen con mayor detalle los fundamentos de la *teoría de la información* necesarios para comprender el funcionamiento del método de alineamiento de imágenes basado en *información mutua*.

Para abordar el alineamiento de dos imágenes es necesario calcular previamente cuál es el valor de información mutua para ambas imágenes, y luego resolver el pro-

blema del ajuste de las mismas. El cálculo de la *información mutua* se basa en la estimación de las distribuciones de probabilidad asociadas a la información existente en ambas imágenes, que en nuestro caso será el nivel de gris de los píxeles que las componen.

En primer lugar veremos cómo se calcula la *información mutua* en las imágenes que pretendemos alinear, y posteriormente cómo se utiliza ese valor para realizar el ajuste de las mismas (la figura 3.1 muestra un esquema del proceso de alineamiento utilizando la *información mutua*).

Para calcular las probabilidades necesarias para la determinación de la *información mutua* se han estudiado las dos técnicas más utilizadas en la literatura, la basada en el método no paramétrico denominado *Ventana de Parzen* [Duda y Hart, 1973, Viola y Wells III, 1995], y el método basado en el cálculo de las probabilidades a partir de las frecuencias de los niveles de gris [Maes y col., 1997].

3.1. Estimación basada en el método *de la ventana de Parzen*

La entropía es una medida de la incertidumbre o aleatoriedad de una variable aleatoria. En el caso que nos ocupa, las variables aleatorias con las que vamos a tratar son los niveles de gris de las imágenes que pretendemos alinear, por lo que las funciones de distribución se refieren a la probabilidad de que un valor tome un determinado nivel de gris.

Las entropías descritas en el apéndice A se definen en términos de integrales sobre las densidades de probabilidad asociadas con las variables aleatorias (imágenes) X e Y . En algunas ocasiones (por ejemplo cuando se alinean imágenes médicas) no se dispone de dichas densidades. Los métodos no paramétricos de estimación de la función de densidad trabajan directamente con muestras del conjunto, sin la utilización de estimadores.

El primer paso a la hora de estimar la entropía de una muestra consiste en aproximar la función de densidad de probabilidad subyacente de los niveles de gris $P(g)$

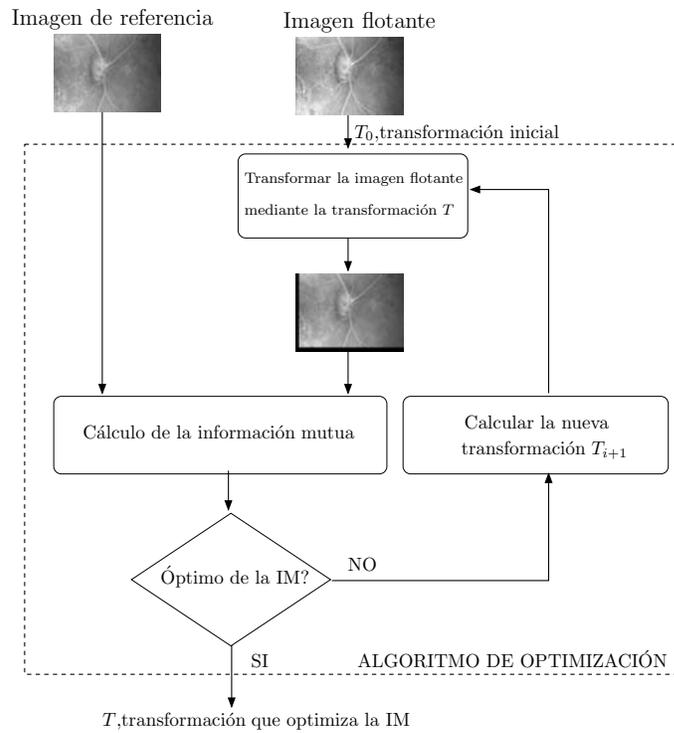


Figura 3.1: Esquema del alineamiento empleando *información mutua* (IM) como medida del ajuste. La *información mutua* representa el grado de ajuste de las imágenes, por lo que se buscará la transformación que maximiza dicha medida en un proceso iterativo.

mediante la superposición de funciones centradas en los elementos de una muestra A tomada de la imagen I :

$$P(g) \approx P^*(g) = P^*(g, A) \equiv \frac{1}{N_A} \sum_{g_a \in A} R(g - g_a) = E_A[R(g - g_a)] \quad (3.1)$$

donde N_A es el número de elementos de la muestra A , R es una función de densidad auxiliar (también llamada *función ventana*) cuya integral es 1 y E representa la esperanza matemática (ecuación A.2). A $P^*(x)$ se la conoce como la *Estimación de densidades mediante la Ventana de Parzen* [Duda y Hart, 1973, Viola y Wells III, 1995]. La calidad de la aproximación depende de la forma de la función ventana R y de su anchura. Diferentes funciones ventana conducirán a estimaciones de la densidad muy diferentes. En este desarrollo asumiremos que la función ventana tiene la forma de una función de densidad gaussiana (figura 3.2). Esto simplifica en gran medida el análisis de la técnica, así como los requisitos computacionales (la entropía de una gaussiana puede calcularse directamente de su varianza), aunque no es necesario que R tenga esta forma. Puede utilizarse como soporte cualquier función diferenciable, como por ejemplo la función de densidad de Cauchy, que también sería una buena elección.

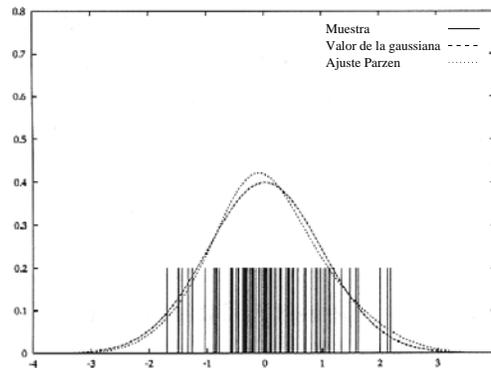


Figura 3.2: Estimación de una gaussiana de media 0 y varianza 1 a partir de 100 muestras (líneas verticales negras) mediante una función de densidad ventana de Parzen.

La función de densidad gaussiana es:

$$G_\psi(x) \equiv \frac{1}{2\pi\sqrt{|\psi|}} \exp\left(-\frac{1}{2}x^t\psi^{-1}x\right) \quad (3.2)$$

donde ψ es la varianza de la gaussiana, $|\psi|$ es su determinante y t representa la operación de trasposición de matrices. Sustituyendo R por G en la ecuación 3.1:

$$P^*(g, A) = \frac{1}{N_A} \sum_{g_a \in A} G_\psi(g - g_a) \quad (3.3)$$

La estimación de densidad Parzen $P^*(g)$ y la función ventana de Parzen R se pueden definir sobre datos escalares o vectoriales. Cuando I es un vector, ψ es la matriz de covarianzas de una gaussiana multidimensional.

Intuitivamente, el método de la ventana de Parzen calcula una media local (en la ventana definida) de la muestra. La función ventana R define una región centrada en g en la que los puntos de la muestra contribuyen a la estimación de la densidad. Los puntos que se encuentran fuera de la ventana no contribuyen. La estimación de la densidad en g es el cociente entre el número de puntos muestreados dentro de la ventana y el número total de puntos de la muestra, N_A .

Desafortunadamente, evaluar la integral de la entropía:

$$H(g) \approx -E[\ln P^*(g, A)] = - \int_{-\infty}^{\infty} \ln P^*(g, A) \partial g \quad (3.4)$$

es extremadamente difícil (ver [Viola, 1995]). Sin embargo, puede aproximarse esta integral como la media muestral

$$H(g) \approx H^*(g) \equiv -\frac{1}{N_B} \sum_{g_b \in B} \ln P^*(g_b) \quad (3.5)$$

donde N_B es el número de elementos de una segunda muestra B . La media muestral converge hacia la media real con una velocidad proporcional a $\frac{1}{\sqrt{N_B}}$.

Ahora ya podemos expresar una aproximación para la entropía de una variable aleatoria I como sigue:

$$H(I) \equiv -\frac{1}{N_B} \sum_{g_b \in B} \ln \frac{1}{N_A} \sum_{g_a \in A} G_{\psi}(g_b - g_a) \quad (3.6)$$

De esta forma, utilizando dos muestras A y B de las imágenes podemos calcular la entropía: la primera se utiliza para estimar la densidad mientras que la segunda se utiliza para estimar la entropía.

3.2. Estimación de las probabilidades utilizando las frecuencias de los niveles de gris

En la segunda técnica analizada para la obtención de las funciones de probabilidad nos basamos en la propuesta de Maes y col. [Maes y col., 1997], que estiman las probabilidades utilizando la distribución de los niveles de gris de los pares de píxeles y los valores de los píxeles en la parte común de las dos imágenes. La figura 3.3 ilustra este método: el histograma conjunto de las imágenes de referencia y flotante contendrá en la posición (g_1, g_2) el número de veces que los valores de gris g_1 y g_2 han ocupado la misma posición en ambas imágenes.

Así, dadas dos imágenes I y J con m y n niveles de gris respectivamente, la *información mutua* entre las dos imágenes vendrá dada por (ecuación A.6):

$$IM(I, J) = H(I) - H(J|I) = \sum_{i=1}^m \sum_{j=1}^n P(g_i, g_j) \ln \frac{P(g_i, g_j)}{P(g_i)P(g_j)} \quad (3.7)$$

donde $P(g_i)$ es la probabilidad de ocurrencia del nivel de gris g_i en la imagen I , $P(g_j)$ es la probabilidad de ocurrencia del nivel de gris g_j en la imagen J y $P(g_i, g_j)$ es la probabilidad de que los niveles g_i y g_j ocurran en una misma posición en las imágenes I y J , respectivamente, una vez aplicada la transformación pertinente a la imagen flotante J .

La probabilidad $P(g_i, g_j)$ se estima usando la distribución de pares de píxeles de la parte común de las dos imágenes cuando sobre la imagen flotante se ha realizado una transformación rígida. Así:

$$\begin{aligned}
 P(g_i) &= \frac{h(i)}{N} \\
 P(g_j) &= \frac{h(j)}{N} \\
 P(g_i, g_j) &= \frac{h(i, j)}{N}
 \end{aligned}
 \tag{3.8}$$

donde $h(i)$ representa el valor del histograma para el nivel de gris i , $h(i, j)$ representa el histograma conjunto, y N es el número de puntos en la parte común del espacio de muestras.

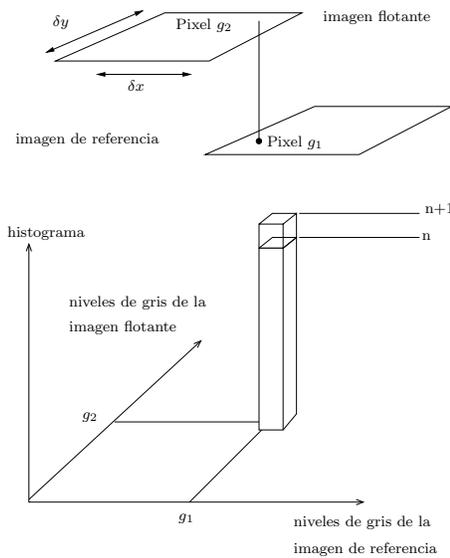


Figura 3.3: Cálculo del histograma conjunto de dos imágenes. El histograma conjunto de las imágenes de referencia y flotante contendrá en la posición (g_1, g_2) el número de veces que los valores de gris g_1 y g_2 han ocupado la misma posición en ambas imágenes.

El histograma conjunto de las dos imágenes se obtiene tomando un píxel de la imagen de referencia y el correspondiente píxel, una vez aplicada la transformación, de la imagen flotante e incrementando en una unidad la entrada del histograma de dos dimensiones correspondiente a los dos valores obtenidos, tal y como muestra la figura 3.3.

Las distribuciones de probabilidad marginales se obtienen calculando los histogramas de la zona de superposición de cada imagen por separado. Para ello se suma el

histograma conjunto sobre las filas y sobre las columnas:

$$P(g_i) = \frac{\sum_{j=1}^m h(g_i, g_j)}{N} \quad (3.9)$$

$$P(g_j) = \frac{\sum_{i=1}^n h(g_i, g_j)}{N}$$

La figura 3.4 muestra las áreas que se utilizan para calcular la probabilidad conjunta. Una vez transformada la imagen flotante (figura 3.4(d)), algunas áreas de la misma dejarán de contener información relevante, por lo que aportan datos erróneos al histograma conjunto, y no se considerarán en el cálculo del mismo. En la imagen original se elimina la misma zona no considerada en la imagen transformada (zona negra en la figura 3.4(c)). Así obtendremos la probabilidad conjunta empleando sólo aquellos píxeles que realmente contienen información común sobre el fondo del ojo.

Desde el punto de vista computacional, el comportamiento del método de la ventana de Parzen es diferente con respecto a los esquemas paramétricos. Construir un modelo con parámetros implica lentas búsquedas en el espacio paramétrico, y el tamaño de la muestra influye proporcionalmente en el tiempo de búsqueda, mientras que en un modelo Parzen no se requieren estos cálculos. Sin embargo, en el modelo paramétrico, una vez estimados los parámetros del mismo, el tiempo requerido para los cálculos es pequeño y no crece al aumentar el número de muestras. Evaluar $P^*(x)$ es más costoso computacionalmente y requiere un tiempo proporcional al tamaño de la muestra. El coste computacional final de los dos métodos depende del modo en el que sean usados.

Para ilustrar todo este proceso de obtención de la información mutua, en la figura 3.5 se muestra el mapa de los valores de *información mutua* obtenidos al desplazar una imagen sobre si misma ± 25 píxeles. En la figura 3.5(b) se muestra el mapa de la información mutua calculado mediante el método de estimación de probabilidades que acabamos de describir, mientras que la figura 3.5(c) muestra el mapa de la *información mutua* obtenido utilizando el método de la ventana de Parzen, tomando una muestra de 810 elementos.

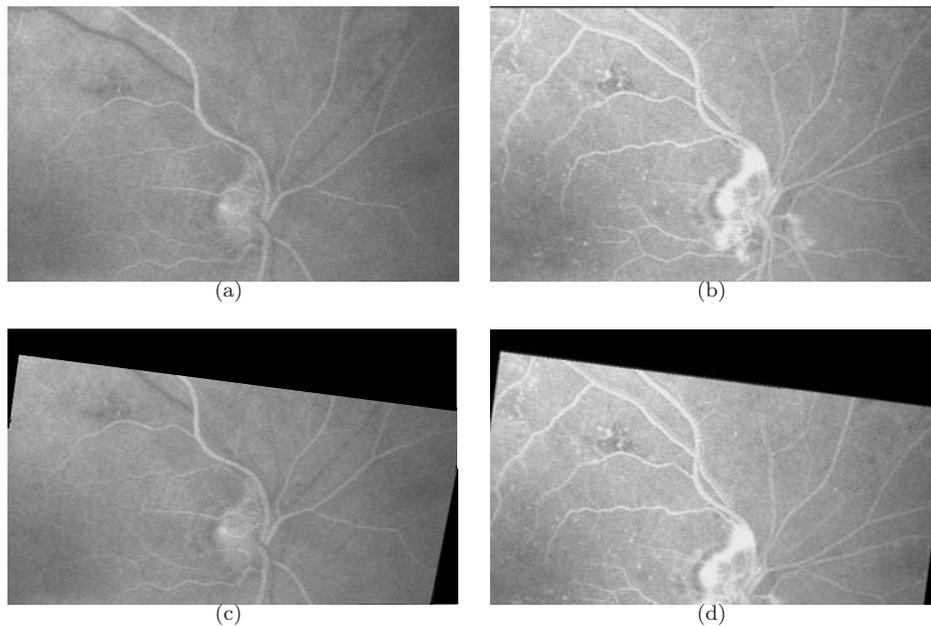


Figura 3.4: Área de superposición de las imágenes, que se utiliza para calcular el histograma conjunto de las imágenes de referencia y flotante. (a) y (b) muestran las imágenes original y flotante, respectivamente. Una vez aplicada la transformación que se desea evaluar para saber si es la que maximiza el área de superposición, en la imagen flotante se eliminan ciertas áreas del cálculo del histograma, marcadas en negro en la imagen (d). Estas mismas zonas serán eliminadas en la imagen original (zonas marcadas en negro en la figura (c)), de forma que sólo se empleará en el cálculo del histograma la información común a ambas imágenes transformadas.

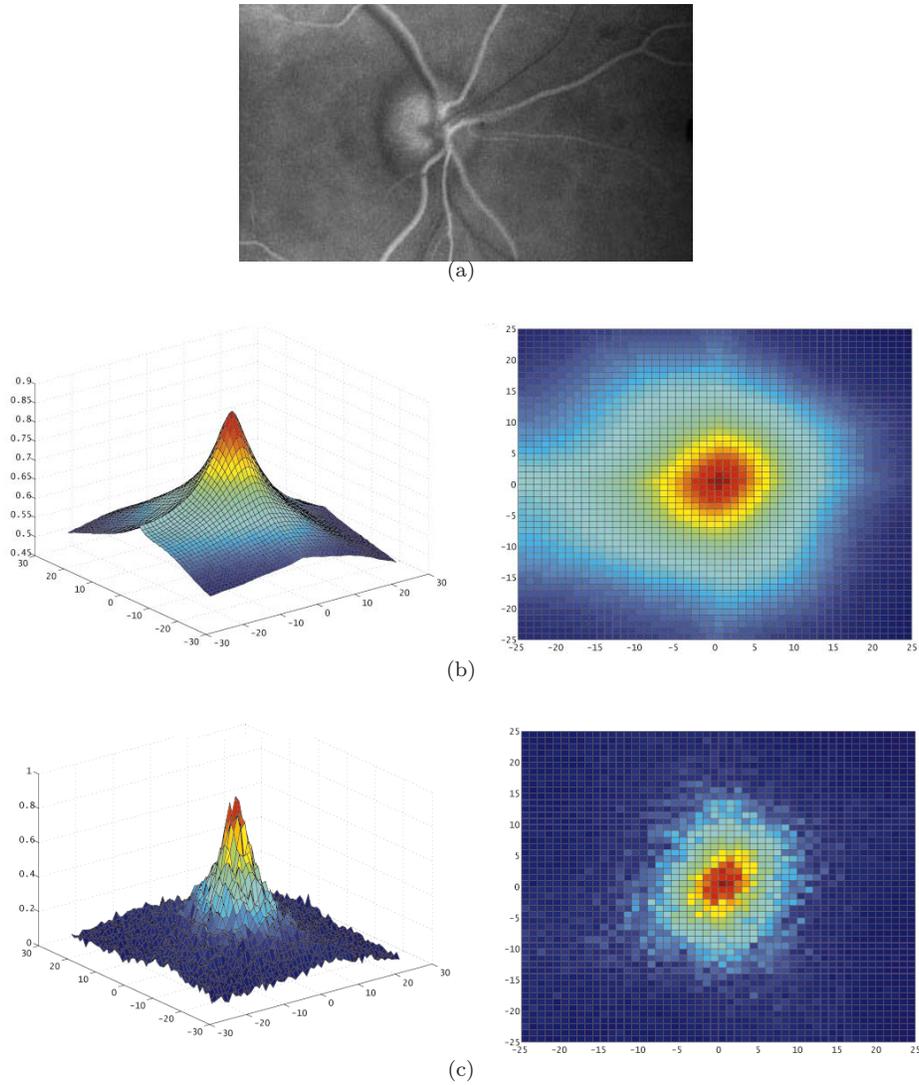


Figura 3.5: Mapas de *información mutua* obtenidos mediante los dos métodos de estimación de las funciones de probabilidad estudiados, sobre la imagen (a). (b) resultado obtenido empleando la estimación basada en frecuencias. (c) resultado obtenido mediante el método de la ventana de Parzen tomando 810 muestras.

Además de todo esto, para reducir los tiempos de cálculo, se utiliza la imagen reducida a la mitad de resolución, es decir, se realiza un submuestreo tomando uno de cada 4 píxeles. Como muestra la figura 3.6, el mapa de la información mutua obtenido utilizando todos los píxeles resultantes de este submuestreo contiene la información suficiente para que el resultado del alineamiento no varíe al disminuir la precisión. En dicha figura se puede apreciar cómo el punto donde se localiza el máximo de la *información mutua*, que deberá estar en las coordenadas $(-20,667, -24,293)$, se mantiene, eso sí, trasladado al $(-10,300, -12,140)$, la mitad del original, debido al submuestreo que se aplica a la imagen.

Para comprobar cuantitativamente que la reducción en la resolución no afectaba a los resultados obtenidos se llevó a cabo el siguiente experimento: se tomaron aleatoriamente 200 imágenes de las distintas secuencias de *SLO* y se construyó una versión reducida de cada una de ellas con la mitad de resolución, tal y como acabamos de comentar. Después se aplicó una transformación también aleatoria a cada una de las imágenes, tanto en su versión original como en la reducida (la misma a ambas), y se procedió a ejecutar el proceso de alineamiento. El error obtenido al utilizar las versiones reducidas fue, en promedio, 1,309 píxeles, considerando como error la diferencia entre la transformación recuperada sobre la imagen original y la recuperada con la imagen con la mitad de resolución. Mediante esta técnica obtenemos unos buenos resultados mejorando de forma significativa los tiempos de cálculo de la función de densidad, y por extensión los tiempos de estimación de la *información mutua*.

3.3. Algoritmo de Maximización de la *información mutua*

Una vez calculada la *información mutua*, hemos de determinar el método que se utilizará para optimizar dicho valor, ya que, como se ha comentado anteriormente, el alineamiento es un proceso iterativo en el que se busca la optimización de la medida de similitud (la *información mutua* en este caso). Las imágenes de videoangiografía con las que trabajamos tienen un tamaño de 720×576 píxeles, y pueden sufrir desplazamientos de hasta 100 píxeles tanto en la dirección vertical como en la horizontal,

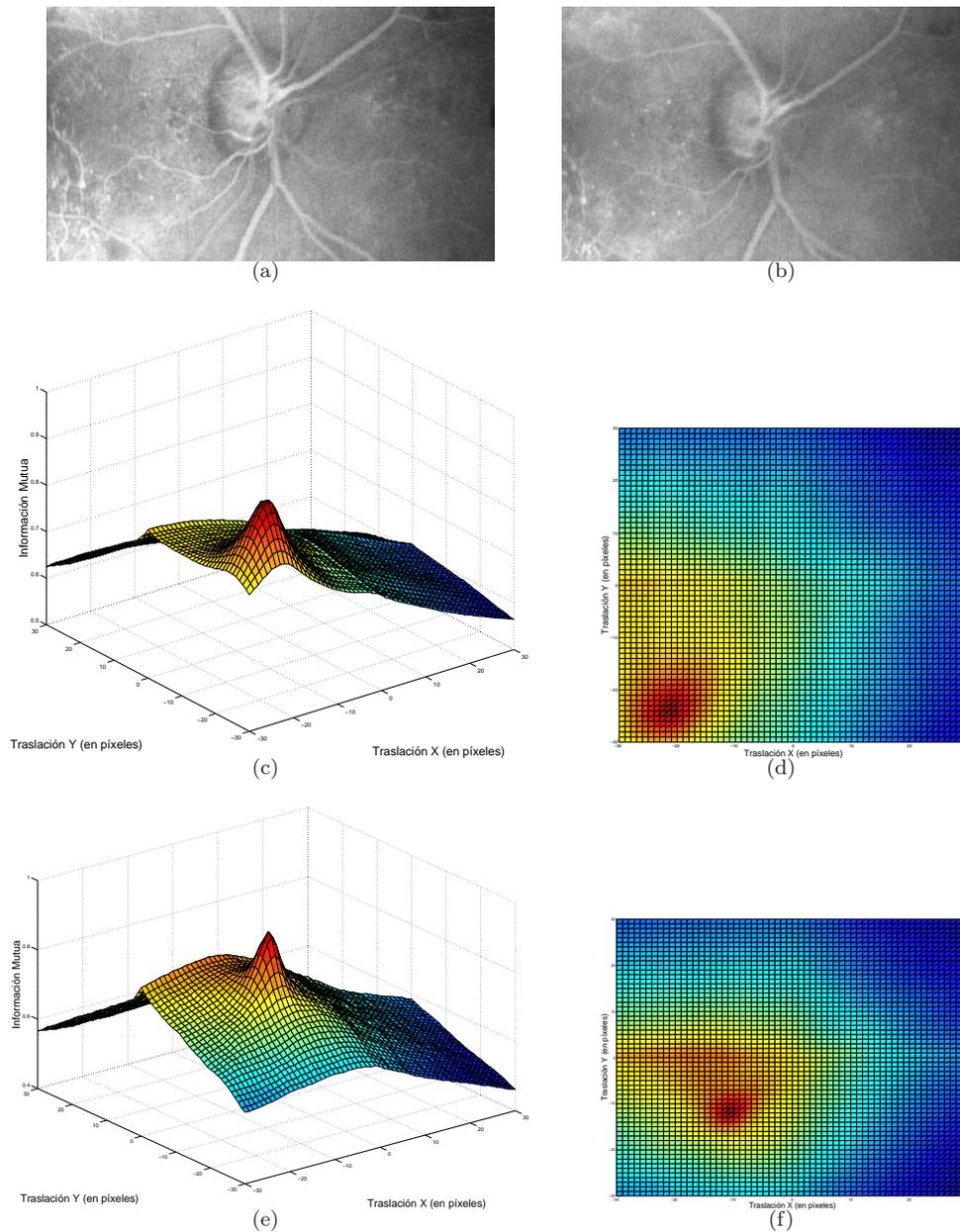


Figura 3.6: El resultado de alinear las imágenes (a) y (b) proporciona como resultado los mapas de *información mutua* de las figuras (c) y (e). En (c) y (d) se muestra el mapa de *información mutua* obtenido al alinear las imágenes (a) y (b) con la resolución original, mientras que (e) y (f) muestran el mapa de *información mutua* cuando se ha reducido a la mitad la resolución de las imágenes para reducir el tiempo de proceso del alineamiento.

lo que implica un espacio de búsqueda de 10^5 muestras, demasiado grande para que una exploración exhaustiva sea un procedimiento viable.

El algoritmo utilizado en este trabajo para maximizar la *información mutua* es el *Temple Simulado*. Esta técnica pertenece a la clase de los algoritmos probabilísticos e iterativos. Pretende simular el proceso de enfriamiento que se utiliza para los metales, consistente en elevar la temperatura hasta que el material se funde y, a continuación, disminuir la temperatura hasta que se alcanza el punto de congelación empezando en un estado inicial determinado. Además permite solventar el problema de los máximos locales que aparecen al utilizar la información mutua como medida del ajuste, ante los cuales el método downhill-simplex presenta problemas. Para una descripción más completa véase el apéndice C.2.

Se han analizado también los resultados obtenidos con el método *Downhill Simplex* desarrollado por Nelder y Mead [Nelder y Mead, 1965], que utiliza el concepto de un simplex, que es un polígono de $N + 1$ vértices en N dimensiones. El método busca una solución óptima local a un problema con N variables cuando la función objetivo a optimizar varía de forma suave. El método sigue una serie de pasos, la mayoría de los cuales consisten en mover el punto del simplex en el que la función a optimizar tiene el valor máximo (para una descripción más detallada consúltese el apéndice C).

Tan sólo estudiaremos aquí los resultados obtenidos con cada uno de ellos, que nos han llevado finalmente a elegir el Temple Simulado como estrategia de búsqueda. Pueden encontrarse descripciones más detalladas sobre los métodos de optimización comentados en [Press y col., 1992, Ingber, 1993].

3.3.1. Comparativa de los métodos de cálculo de la *información mutua*

En los apartados anteriores hemos descrito diferentes métodos para el cálculo de la *información mutua*, así como para llevar a cabo su maximización. Queda por determinar cuál es el procedimiento que, utilizando la *información mutua* calculada mediante una de estas técnicas y mediante su optimización con uno de los dos algoritmos descritos, proporciona mejores resultados. Para realizar dicha evaluación se han

empleado fotogramas de la sección 1 de la figura 2.15, en la que el método basado en el cálculo de crestas y valles no consigue ajustar correctamente los fotogramas que no poseen todavía fluoresceína. Además debemos recordar que en el caso de la *información mutua* se ajustan los parámetros de traslación en X e Y, no la rotación ni el escalado.

Siguiendo esta idea, se han realizado dos experimentos: con el primero se trata de analizar la efectividad en cuanto a cómo influye el número de muestras tomadas de la imagen tanto en el tiempo empleado en el proceso de alineamiento de los fotogramas como en el porcentaje de éxito y precisión obtenida. En el segundo experimento se analizará la eficiencia de cada uno de los algoritmos de maximización de la *información mutua* que se han descrito.

Existen en la literatura algunos trabajos en los que se analiza el funcionamiento de los distintos métodos de obtención de la función de densidad de probabilidad, como el presentado por Glavinović [Glavinović, 1996], en el que se compara el método de obtención de la función de densidad de la ventana de Parzen con el basado en la frecuencia del histograma, los métodos aquí descritos. En él se llega a la conclusión de que el método basado en la ventana de Parzen es claramente superior ya que necesita menos muestras para obtener buenos resultados. Como veremos en los resultados de nuestros experimentos, esto no sucederá en nuestro caso.

En primer lugar, veremos cómo influye el número de muestras tomadas de la imagen tanto en el tiempo empleado en el proceso de alineamiento de los fotogramas como en el porcentaje de éxito y precisión obtenida. Para ello se diseñó el siguiente experimento, que se ejecutó sobre 200 imágenes de cada una de las 5 secuencias estudiadas:

1. Seleccionar la imagen de referencia (para evitar imágenes con errores causados por el proceso de adquisición).
2. Seleccionar una transformación aleatoria en el rango ± 100 píxeles de traslación, que aplicada nos dará la imagen flotante.
3. Seleccionar un número de muestras (inicialmente 800).

4. Calcular la estimación de probabilidad mediante los valores obtenidos de las muestras tomadas.
5. Alinear las imágenes utilizando el método de la ventana de Parzen.
6. Almacenar tanto el tiempo empleado como el error obtenido.
7. Incrementar el número de muestras.

La figura 3.7 muestra los tiempos empleados en el alineamiento de los 200 fotogramas de cada una de las 5 secuencias evaluadas, en función del número de muestras empleado en la estimación de probabilidades del método de la ventana de Parzen. Recordemos que para la estimación de la función de probabilidad son necesarias dos muestras A y B (cuyos tamaños conforman los ejes X e Y de dicha gráfica), una para estimar la densidad y otra para estimar la entropía.

Observando la figura 3.7(a) se puede concluir que sería necesario tomar muestras de 4000 elementos para que el error cuadrático medio bajase de los 4 píxeles, pero si observamos la figura 3.7(b), vemos que el tiempo necesario para alinear cada fotograma con este número de muestras oscila entre los 26 y los 35 segundos, tiempos demasiado elevados teniendo en cuenta el gran número de fotogramas que hay que alinear en cada secuencia.

Dado que en el caso de la utilización del método de estimación de las probabilidades el número de muestras es constante e igual a la mitad de los puntos existentes en la imagen, el tiempo de cálculo de las probabilidades es constante. La tabla 3.1 muestra, para 5 secuencias de SLO , el tiempo medio \bar{t} necesario para estimar la función de densidad de probabilidad sobre 200 fotogramas tomados aleatoriamente, así como la dispersión de estos tiempos, medida mediante la desviación típica σ de las muestras.

Se puede apreciar con claridad cómo este método es computacionalmente muy superior al de la ventana de Parzen, por lo que será el que utilizemos a la hora de estimar las probabilidades que se emplearán para calcular el valor de la información mutua.

Una vez analizado el primer experimento y seleccionada la técnica de obtención de las probabilidades que se utilizan para estimar la *información mutua*, el segundo

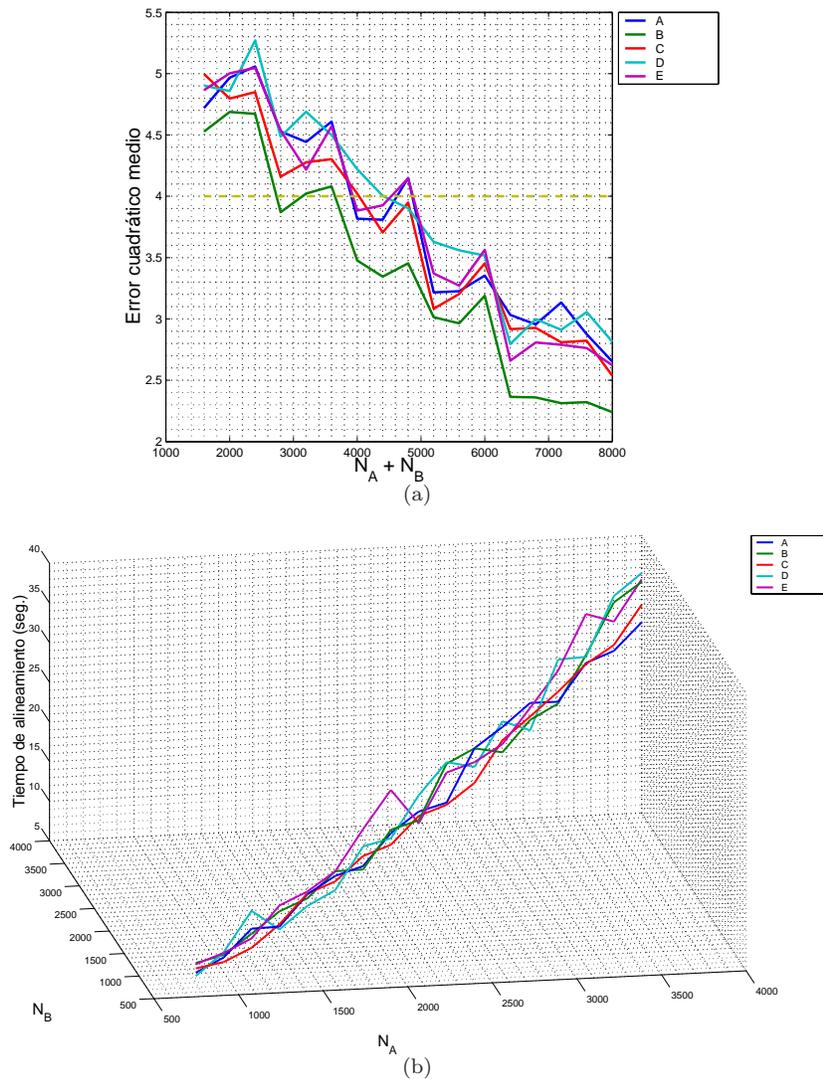


Figura 3.7: Influencia del número de muestras en el tiempo de cálculo de la información mutua. (a) evolución del error cuadrático medio del alineamiento en función del número de elementos de las muestras (N_A , N_B son el número de muestras de A y B respectivamente). (b) tiempos medios empleados en el alineamiento de cada par de imágenes para distintos tamaños de las dos muestras que son necesarias para la estimación de probabilidades utilizando el método de la ventana de Parzen.

	\bar{t} (s)	σ
A	6.322	0.084
B	6.347	0.024
C	6.185	0.055
D	6.027	0.051
E	6.404	0.052

Tabla 3.1: Análisis de los tiempos medios de cálculo de la función de densidad de probabilidad para cinco pacientes distintos empleando el método de estimación basado en el histograma de la imagen. La segunda columna muestra el tiempo medio (en segundos) necesario para obtener la estimación de las probabilidades, mientras que la tercera columna muestra la desviación típica de dichos tiempos.

experimento realizado determinará cuál de los métodos de optimización (Temple Simulado o Downhill Simplex) proporciona mejores resultados. El procedimiento que hemos seguido para obtener estos datos ha sido el siguiente (para 200 imágenes de cada una de las 5 secuencias estudiadas):

1. Seleccionar una imagen de la sección 1.
2. Seleccionar una transformación aleatoria para el eje X y otra para el eje Y, limitando el valor máximo de esta transformación al rango ± 100 .
3. Aplicar a la imagen la transformación seleccionada.
4. Calcular el alineamiento de las imágenes utilizando el método basado en *información mutua*, minimizando este valor con cada uno de los dos métodos de optimización ya comentados.
5. Almacenar tanto el tiempo como el error cometido en el alineamiento.

La Tabla 3.2 muestra los resultados obtenidos con este segundo experimento, resultados que aparecen de forma gráfica en la figura 3.8. En la tabla se pueden apreciar los tiempos de cálculo empleados en el alineamiento de los fotogramas de cada una de las secuencias analizadas, así como el error cuadrático medio (ECM) del error del ajuste, definido como:

$$ECM = \sqrt{\frac{\sum_{i=1}^{200} (\mathcal{T}_{i_a} - \mathcal{T}_{i_r})^2}{200}} \quad (3.10)$$

donde \mathcal{T}_{i_a} representa la transformación aplicada a la imagen y \mathcal{T}_{i_r} representa la transformación recuperada utilizando el método de alineamiento basado en la *información mutua*.

	Downhill Simplex			Temple Simulado		
	Tiempo (seg)	Tiempo por fotograma (seg)	ECM (píxeles)	Tiempo (seg)	Tiempo por fotograma (seg)	ECM (píxeles)
A	1285.922	6.429	4.103	1289.823	6.449	3.916
B	1270.339	6.351	3.980	1259.494	6.297	3.403
C	1365.833	6.829	8.559	1218.988	6.094	5.430
D	1379.860	6.899	4.553	1302.574	6.512	3.741
E	1356.424	6.782	4.950	1360.902	6.804	3.744

Tabla 3.2: Resultados del segundo experimento realizado sobre cinco secuencias de *SLO* pertenecientes a cinco pacientes distintos (A-E), realizado para determinar cuál de los dos métodos de maximización de la información mutua estudiados proporciona mejores resultados. La primera columna muestra, para cada uno de los dos métodos, el tiempo total empleado en el alineamiento de los 200 fotogramas, mientras que la segunda muestra el tiempo medio por fotograma. La tercera columna contiene el error cuadrático medio (en píxeles) de las transformaciones obtenidas para cada uno de los pacientes.

De estos datos se puede extraer la conclusión de que el método del Temple Simulado es en promedio ligeramente más rápido que el DownHill Simplex. Por otro lado, ambos métodos ofrecen un comportamiento muy similar en lo que al error en el ajuste se refiere aunque el *Temple Simulado* siempre obtiene un error por debajo del obtenido por el algoritmo Downhill Simplex. Por todo esto parece lógica la elección del primero como método de maximización de la función de medida de la bondad del alineamiento, la información mutua.

De las gráficas de la figura 3.8, queremos destacar la secuencia del paciente C, pues obtiene un error muy superior al de las otras secuencias estudiadas. Este error puede comprenderse fácilmente si observamos algunos de los fotogramas que componen

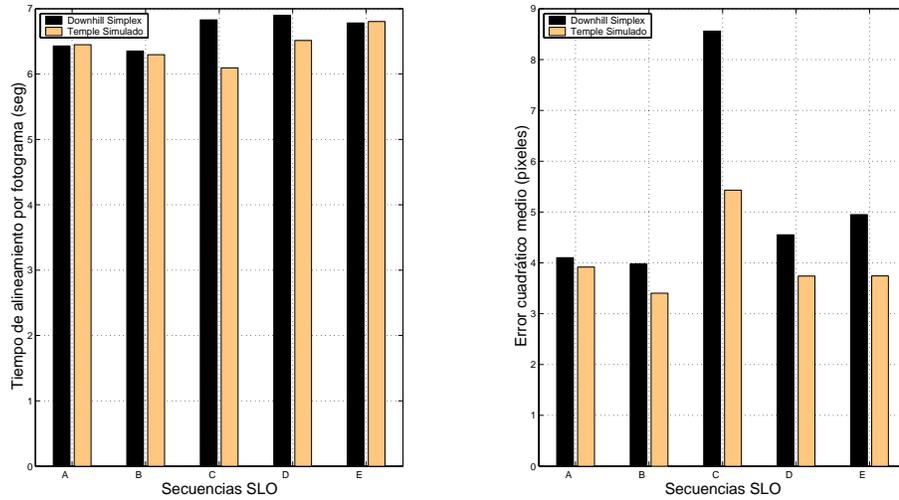


Figura 3.8: Análisis de los métodos de maximización de la *información mutua*. La gráfica de la izquierda muestra el tiempo empleado en el alineamiento de cada par de imágenes para cada una de las secuencias estudiadas, mientras que la de la derecha muestra el ECM cometido para cada una de ellas.

dicha secuencia (figura 3.9). En esta secuencia aparecen algunas oclusiones parciales (marcada con un rectángulo blanco en la figura 3.9) así como la fovea, señalada en 3.9 con un círculo gris, que resulta especialmente engañosa para un método como el basado en *información mutua*, que utiliza la información de los niveles de gris, y no la de estructuras como los vasos, y así puede tomar (erróneamente en este caso) la estructura de la fovea como referencia del alineamiento. Es necesario tener en cuenta que estas estructuras no representan ningún problema en las restantes secciones de la secuencia en las que se utiliza el alineamiento basado en crestas, por lo que, en vista de que el incremento en el error no es significativo, no se diseñará ninguna estrategia específica para tratar de evitar estas fuentes de error.

Una vez que se han determinado tanto el método de obtención de las probabilidades que se utilizan para calcular la *información mutua* (estimación utilizando las frecuencias de los niveles de gris) como la estrategia que se empleará para guiar el proceso de búsqueda (*Temple Simulado*), se procede a analizar el comportamiento de la técnica de alineamiento basado en *información mutua* sobre la sección 1 de la figura 2.15, en la cual el método de alineamiento basado en la extracción de crestas

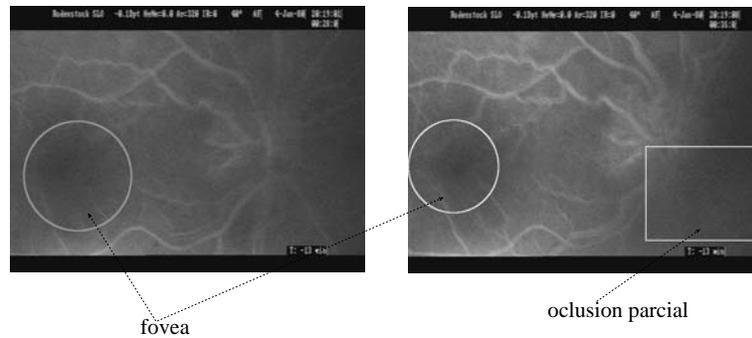


Figura 3.9: Fotogramas número 639 (izquierda) y número 834 (derecha) del paciente C. Se puede apreciar la baja calidad de la captura, con frecuentes oclusiones parciales (señaladas con un rectángulo blanco) que hacen que el alineamiento de esta secuencia presente más dificultades que en otras (véase por ejemplo los fotogramas 2.7(a) y 2.7(b), de los pacientes B y D). Además en esta secuencia aparece claramente reflejada la fóvea (zona marcada con un círculo blanco), que puede ser tomada por el algoritmo basado en *información mutua* como referencia para el alineamiento, aumentando así el número de ajustes erróneos.

no proporciona buenos resultados, con un número mayor de experimentos. Para ello se diseñó un nuevo experimento, en el que se procedió a alinear 500 fotogramas seleccionados aleatoriamente dentro de esta sección, para cada una de las 5 secuencias *SLO* estudiadas.

La Tabla 3.3 muestra los resultados obtenidos en este experimento. Las columnas de dicha tabla tienen el mismo significado que en la Tabla 3.2, es decir, la primera columna muestra el número de alineamientos obtenidos con éxito y la segunda indica el error cuadrático medio de dichos alineamientos.

De esta tabla se deduce que el método de alineamiento basado en *información mutua* permite alinear por término medio el 94.4% de las imágenes de la zona de las secuencias de *SLO* en las que la técnica basada en la extracción de características no proporcionaba buenos resultados, con lo que se dispone de dos métodos que se complementan a la hora de alinear los fotogramas de las secuencias, para así poder llevar a cabo el procesamiento de las secuencias completas. Por lo tanto, se empleará el método basado en *información mutua* para alinear los fotogramas pertenecientes a la sección 1, mientras que para las otras secciones se utilizará el método basado en

	Transformaciones con éxito (%)	ECM
A	96	3.114
B	98	2.216
C	88	4.341
D	96	4.618
E	94	4.914

Tabla 3.3: Estadísticas del experimento realizado para analizar el comportamiento del método de alineamiento basado en el cálculo de la *información mutua* sobre fotogramas de la sección 1 de la figura 2.15: la primera columna indica el porcentaje de alineamientos recuperados con éxito, mientras que la segunda indica el error cuadrático medio de las transformaciones recuperadas.

la extracción de crestas, debido a que proporciona una mayor precisión y mejores tiempos de cálculo.

Capítulo 4

Metodología de alineamiento de secuencias

Una vez establecida la necesidad de utilizar de forma combinada los métodos descritos en los dos capítulos precedentes, mostraremos en este capítulo la metodología para el alineamiento de una secuencia completa de imágenes de *SLO*. Para ello deberemos combinar ambos métodos de forma que cada uno de ellos se encargue de alinear aquellos fotogramas en los que el otro encuentra problemas debido a las características de luminosidad. Utilizaremos como método principal el basado en crestas por su velocidad y precisión y, donde este no funcione correctamente debido a que no es posible detectar marcadores anatómicos (sección 1 de la secuencia) usaremos el método basado en la *información mutua*.

4.1. Selección de imágenes de referencia

El procedimiento de alineamiento empleado por los expertos clínicos a la hora de abordar el análisis de las secuencias consiste en alinear uno a uno los fotogramas de la secuencia tras seleccionar los fotogramas de referencia, tomando puntos característicos de las imágenes (normalmente bifurcaciones de los vasos o los mismos vasos) como base para el alineamiento.

En el procedimiento que proponemos, el primer paso consiste en que el experto clínico seleccione manualmente dos fotogramas de la secuencia (F_1 y F_2 en la figura 4.1), que serán los fotogramas de referencia para la totalidad de los fotogramas de la secuencia. Hemos de comentar que, como es obvio, el primer fotograma F_1 deberá ser representativo de las imágenes pertenecientes a la primera parte de la secuencia (en la que el nivel de fluoresceína es bajo), mientras que el segundo fotograma F_2 deberá representar a las imágenes de la parte más brillante de la secuencia (en la que la fluoresceína ya ha invadido el árbol arterio-venoso de la retina). La selección de dos fotogramas de referencia es necesaria debido a la gran diferencia existente entre los pertenecientes a la zona clara y los de la zona oscura de la secuencia, lo que hace extremadamente difícil la obtención de buenos resultados utilizando únicamente un fotograma de referencia.

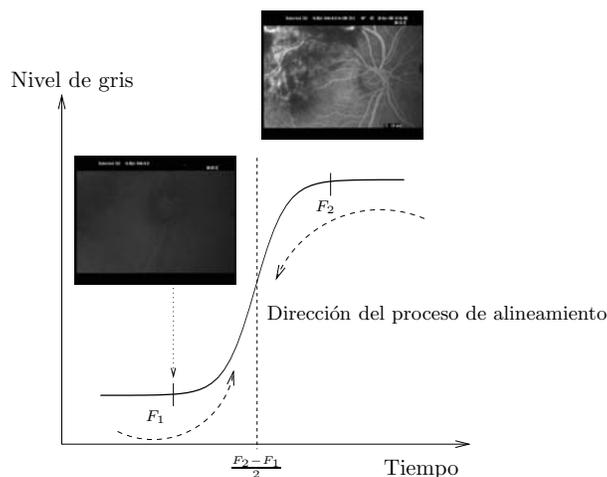


Figura 4.1: Fotogramas de referencia F_1 y F_2 seleccionados por el oftalmólogo que se usarán para alinear la secuencia de *SLO*.

Esta selección representa un paso fundamental en el proceso de alineamiento, ya que la calidad de los fotogramas seleccionados influye decisivamente en el comportamiento de los métodos de alineamiento. Además, se deben elegir F_1 y F_2 sin errores de adquisición, consecuencia de parpadeos involuntarios o de errores de sincronismo entre los dispositivos de adquisición, como se puede ver en la figura 4.2. En la imagen izquierda se muestra un fotograma en el que el paciente parpadeó, provocando la oclusión del fondo del ojo y la consiguiente eliminación de toda la información relevante,

por lo que resultaría imposible ya no sólo alinear cualquier otro fotograma con él (si se seleccionara como referencia), sino alinearlo a él mismo contra el fotograma de referencia que se seleccione. En la imagen de la figura 4.2(b) se muestra un fotograma en el que se produce un error de sincronización entre el dispositivo de barrido del fondo del ojo y el dispositivo de captura de vídeo, que da lugar a que los campos par e impar de la imagen aparezcan sin entrelazar, y con ello que se dupliquen los elementos que aparecen en la imagen. Veremos en el capítulo 5 cómo estos fotogramas, en los que el resultado del alineamiento será con toda probabilidad errónea, se eliminarán del resultado final.

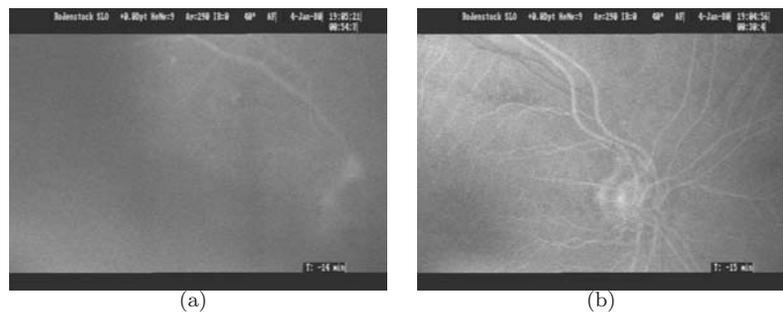


Figura 4.2: Ejemplos de fotogramas que no sirven como referencia. (a) error debido al parpadeo del paciente, dando lugar a que los vasos aparezcan ocluidas por el párpado. (b) error debido a problemas en el sincronismo del sistema de adquisición del vídeo con el sistema de captura de la imagen de fondo del ojo, dando lugar a que los campos par e impar de la imagen de barrido no se superpongan, duplicando las estructuras presentes en la imagen.

Una vez seleccionados los fotogramas de referencia, comienza el proceso de alineamiento de la secuencia, que se muestra de forma esquemática en la figura 4.3.

Tal y como muestran las figuras 4.1 y 4.3, el rango de fotogramas desde el número 1 hasta el número $\frac{F_2 - F_1}{2}$ serán alineados utilizando como fotograma el primero de referencia seleccionado por el oftalmólogo (F_1), mientras que los fotogramas desde el último de la secuencia hasta el número $\frac{F_2 - F_1}{2} + 1$ serán alineados utilizando como referencia el segundo seleccionado (F_2).

Una vez seleccionados los fotogramas de referencia, sólo necesitamos decidir qué método debemos emplear para alinear cada uno de los fotogramas. Esta decisión deberá to-

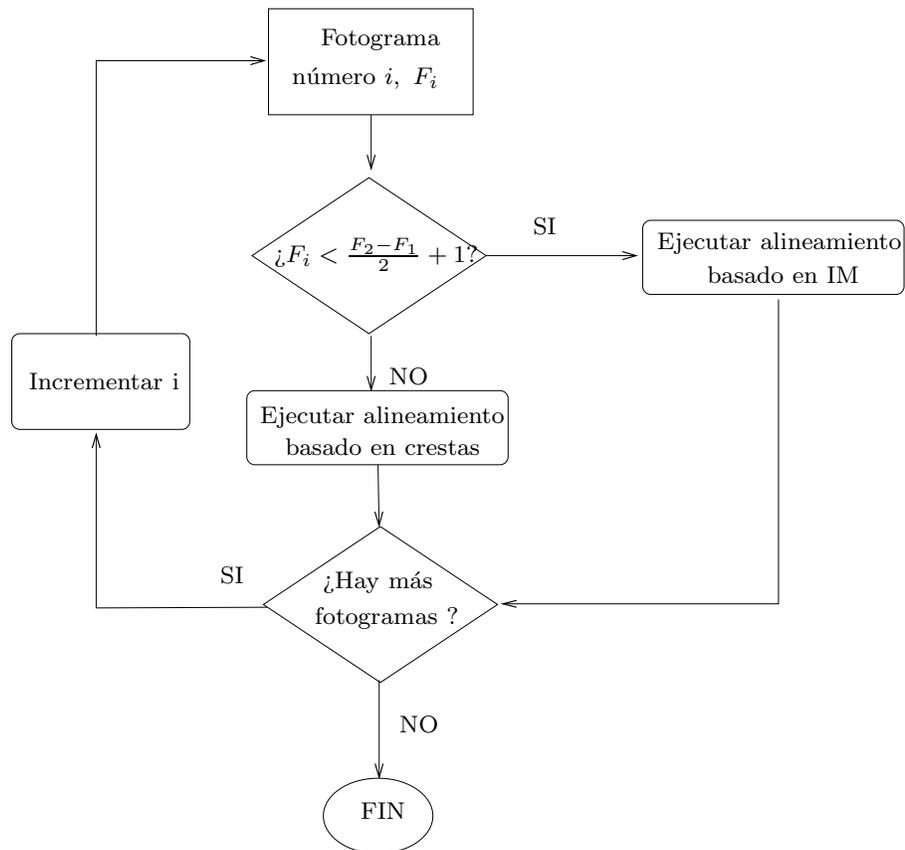


Figura 4.3: Esquema del proceso de ajuste de los fotogramas de una secuencia *SLO*. Hasta el fotograma $\frac{F_2-F_1}{2}$, se utiliza el método basado en *información mutua* para realizar el alineamiento de las imágenes, mientras que a partir de dicho fotograma se emplea el método basado en crestas, que ya tendrá suficiente información en la imagen para extraer las crestas que nos permitirán alinear correctamente el fotograma con la imagen de referencia.

marse de forma automática, sin que sea necesaria la intervención del operador. Para ello se probaron distintos métodos que describimos a continuación.

Debido a que el método de alineamiento basado en crestas es mucho más rápido y preciso, se trató de que el método de ajuste basado en información mutua sólo actuase en caso de que el valor de la correlación obtenido fuera inferior a un umbral dado. Sin embargo, la obtención de dicho umbral no es un cálculo inmediato, ya que la correlación es un parámetro sensible a las variaciones de intensidad de la imagen, por lo que aunque dos imágenes estén bien alineadas pueden tener un valor de correlación muy inferior a otras, e incluso inferior a otras que no estén bien alineadas pero que presenten niveles de gris altos, como las que aparecen cuando la fluorescencia está en los niveles más altos de saturación.

Para ilustrar este proceso, la tabla 4.1 muestra el valor de la correlación obtenido al ajustar las imágenes de la figura 4.4 usando el método de alineamiento basado en crestas. Dichas imágenes pertenecen a dos secuencias distintas (4.4(a),4.4(c) y 4.4(e) pertenecen a una secuencia y 4.4(g) a otra) con el objetivo de demostrar que el resultado de la correlación no determina de forma unívoca un buen alineamiento. Al lado de cada imagen se muestra la imagen de crestas obtenida, que es la que se correlaciona en el proceso de ajuste. Las tres primeras filas de la tabla 4.1 muestran el resultado de alinear las imágenes de una misma secuencia, obteniéndose valores para la correlación entre 11,6 y 25,8, rango muy amplio a pesar de que el ajuste es casi perfecto. Las tres últimas filas muestran el resultado de alinear imágenes de dos secuencias distintas. A pesar de esto, vemos que hay un valor de correlación muy alto, 23,0, que podría llevarnos a pensar que el ajuste ha sido bueno, cuando no ha sido así.

Otra posibilidad que se consideró fue la utilización de una medida normalizada, la *correlación cruzada normalizada* [Lewis, 1995] de dos imágenes F y G , definida por la ecuación 4.1.

$$\gamma = \frac{\sum_{x,y}[F(x,y) - \bar{F}][G(x,y) - \bar{G}]}{\left\{ \sum_{x,y}[F(x,y) - \bar{F}]^2 \sum_{x,y}[G(x,y) - \bar{G}]^2 \right\}^{0,5}} \quad (4.1)$$

Imagen original	Imagen flotante	Transformación obtenida (tx, ty, θ)	Correlación ($\times 10^6$)	Correlación Cruzada Normalizada
4.4(b)	4.4(d)	(-24,181; 0,100; 0,000694)	25,8	0,332
4.4(b)	4.4(f)	(17,592; 5,964; 1,267000)	15,4	0,076
4.4(d)	4.4(f)	(-39,044; 0,000; 0,000042)	11,6	0,061
4.4(b)	4.4(h)	(26,816; -0,350; 0,124039)	16,7	0,093
4.4(d)	4.4(h)	(117,257; -1,336; 0,002983)	12,1	0,067
4.4(f)	4.4(h)	(61,440; -0,338; 0,109800)	23,0	0,039

Tabla 4.1: Resultados obtenidos al alinear imágenes de distintas secuencias (las mostradas en la figura 4.4), mostrando la variabilidad de la correlación tanto en imágenes bien alineadas como en resultados erróneos del alineamiento. Las tres primeras filas muestran el resultado de alinear las imágenes de una misma secuencia, obteniéndose valores para la correlación entre $11,6 \times 10^6$ y $25,8 \times 10^6$, rango muy amplio a pesar de que el ajuste es casi perfecto. Las tres últimas filas muestran el resultado de alinear imágenes de dos secuencias distintas. A pesar de esto, vemos que hay un valor de correlación muy alto, $23,0 \times 10^6$, que podría llevarnos a pensar que el ajuste ha sido bueno, cuando no ha sido así.

donde \overline{G} representa el valor medio de la imagen dinámica y \overline{F} representa el valor medio de la imagen estática.

En el cálculo de la *correlación cruzada normalizada* no influyen las fluctuaciones en la intensidad, por lo que se podrían comparar los resultados obtenidos al alinear imágenes muy diferentes. Sin embargo, tal y como muestra la última columna de la tabla 4.1, otra vez nos encontramos con que esta medida no determina de forma unívoca el grado de bondad de un ajuste, ya que si bien el primer resultado de 0,332 se corresponde con el mejor de los alineamientos, también observamos que el alineamiento de los pares de imágenes 4.4(c)-4.4(e) y 4.4(c)-4.4(g) da un valor alrededor de 0,06, siendo el primero un buen ajuste mientras que el segundo se corresponde con un ajuste erróneo.

Debido a la dificultad en la determinación de una medida que indicara si un ajuste utilizando el método basado en crestas funcionó correctamente o no, se decidió que fuese el experto clínico el que, a la hora de elegir el primer fotograma de referencia, F_1 , lo hiciera de forma que dicho fotograma marcara de alguna forma el inicio de la zona en la que el nivel de fluoresceína presente en las arterias del ojo fuera suficiente para que el método basado en crestas tuviera suficiente información para la extracción de los marcadores, y así hasta el fotograma $\frac{F_2-F_1}{2}$ se utilizara el método basado en *información mutua*, mientras que a partir del fotograma $\frac{F_2-F_1}{2} + 1$ se utilizara el método basado en crestas.

Una vez seleccionados los fotogramas de referencia, comienza el proceso de alineamiento. Los fotogramas del principio de la secuencia se alinean utilizando el fotograma F_1 como referencia, y el proceso de alineamiento va desde el fotograma 1 hacia adelante. En el caso de los fotogramas de la parte media y final de la secuencia (desde el último hasta el $\frac{F_2-F_1}{2} + 1$) se alinean empleando como referencia el fotograma F_2 , y el proceso recorre desde el último hacia atrás (ver figura 4.1).

Una vez realizado el alineamiento, debido a las diferencias existentes entre los fotogramas de referencia F_1 y F_2 , es necesario llevar a cabo un proceso de reajuste de las transformaciones obtenidas para aquellos fotogramas situados entre los de referencia, es decir, el conjunto de fotogramas $\{x/F_1 < F_x < F_2\}$. Este hecho puede apreciarse en la figura 4.5, en la que se muestra la evolución del nivel de gris en un píxel se-

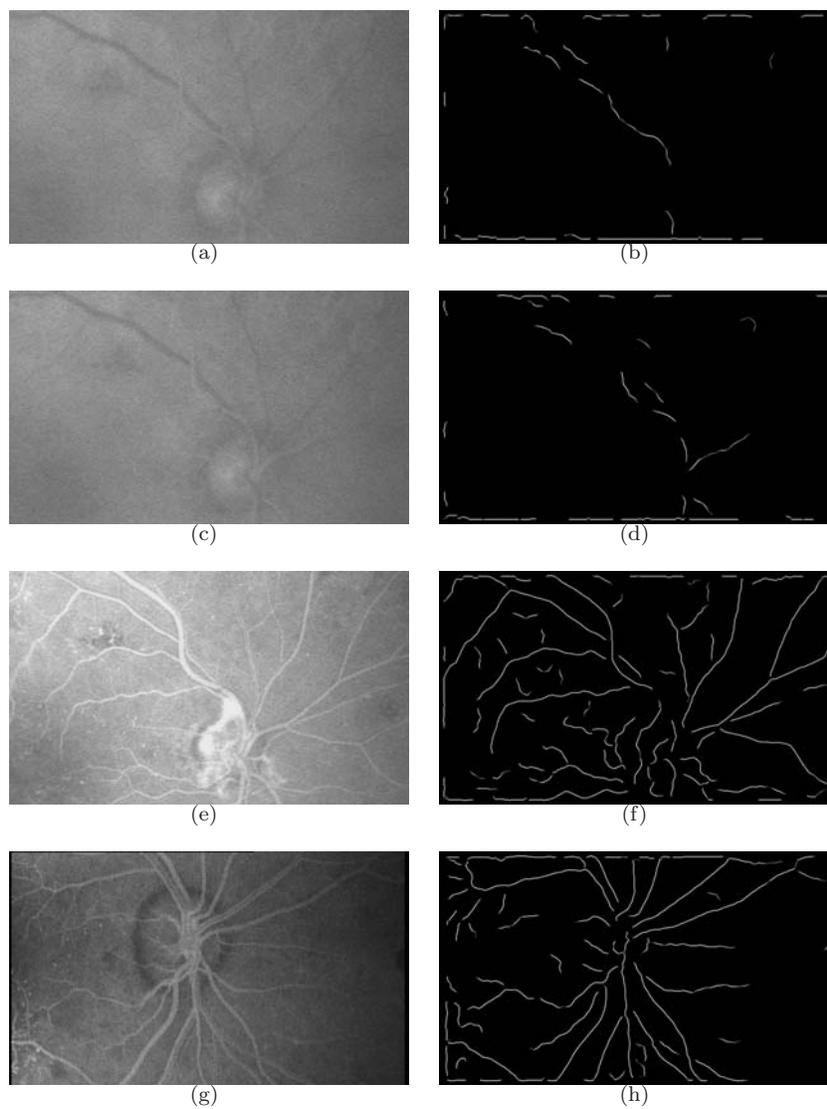


Figura 4.4: Fotogramas y sus correspondientes imágenes de crestas empleadas para evaluar el valor de la correlación como medida de evaluación de bondad del ajuste entre distintas ejecuciones. Las imágenes de las tres primeras filas pertenecen al mismo paciente, mientras que la última pertenece a otro paciente.

leccionado, y se ve claramente el salto que se produce en la zona en la que se unen los fotogramas alineados contra F_1 con los alineados contra F_2 , aproximadamente alrededor del fotograma $\frac{F_2 - F_1}{2}$.

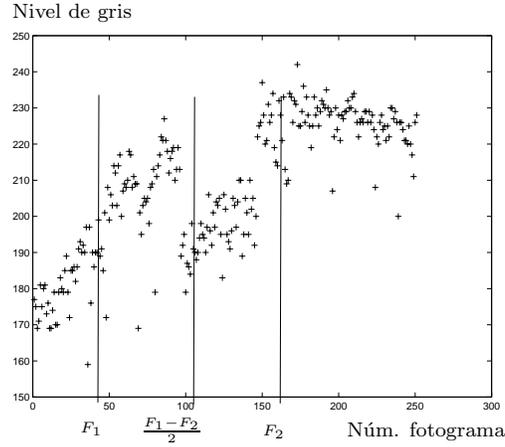


Figura 4.5: Evolución del nivel de gris en un punto seleccionado tras ejecutar el proceso de alineamiento de la secuencia. Se observa que aparece un salto entre las dos zonas definidas por los distintos fotogramas de referencia contra los que se ajustan las imágenes.

El proceso de reajuste consiste en alinear los fotogramas de referencia F_1 y F_2 utilizando cualquiera de ellos como referencia (esto es posible debido a que nuestro proceso de alineamiento cumple la propiedad de *consistencia* [Christensen y Johnson, 2001]). Denominemos por \mathcal{T}_i a la matriz de coordenadas homogéneas que representa la transformación obtenida en el alineamiento del fotograma F_i . El resultado del proceso de alineamiento será la matriz de transformación con coordenadas homogéneas \mathcal{T}_t que ajusta ambos fotogramas de referencia. Si se compone dicha matriz \mathcal{T}_t con cada una de las matrices de transformación \mathcal{T}_i , $F_1 < F_i < F_2$ pertenecientes a los fotogramas de la zona del salto, nos permitirá corregir el desalineamiento entre ambos fotogramas de referencia, de forma que las matrices de alineamiento serán ahora \mathcal{T}_i^* :

$$\mathcal{T}_i^* = \mathcal{T}_i \circ \mathcal{T}_t \quad , F_1 < F_i < F_2 \quad (4.2)$$

Una vez aplicada esta corrección, el salto que aparecía en la figura 4.5 se elimina y aparece una distribución de puntos uniforme, como puede apreciarse en la figura 4.6.

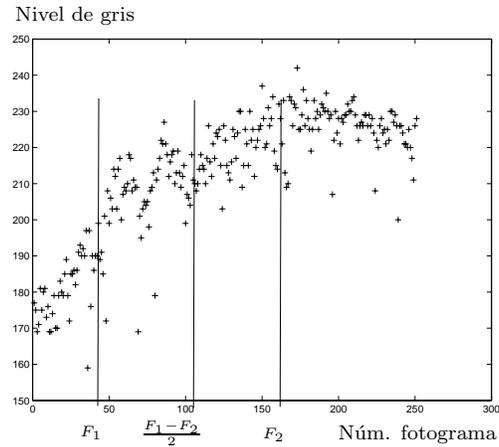


Figura 4.6: Una vez que se ha ajustado la zona entre los dos fotogramas de referencia el salto en la distribución de puntos desaparece, y la curva obtenida es similar a la que obtienen los oftalmólogos de forma manual.

Utilizando una imagen en la que se superponen los fotogramas de una secuencia, se puede llevar a cabo una evaluación visual del proceso de ajuste. La figura 4.7 muestra un esquema del proceso de construcción de estas imágenes, en las que el eje temporal constituye una tercera dimensión. Cada fotograma se apila sobre su precedente y en la parte superior de la imagen se muestra el resultado de aplicar un corte con un plano horizontal que pasa por un punto seleccionado de la imagen, mientras que en la parte izquierda se muestra el resultado de aplicar un corte con un plano vertical que pasa por el punto seleccionado.

En la figura 4.8 se muestra un ejemplo de dichas imágenes apiladas. Si el proceso de alineamiento tiene éxito, los vasos aparecerán como líneas rectas en las imágenes superior e izquierda, como puede verse en los recuadros señalados en las figuras 4.8(b) y 4.8(d). Sin embargo, si las imágenes no están correctamente alineadas, como en las figuras 4.8(a) y 4.8(c), aparecerán desplazamientos entre fotogramas vecinos, y los vasos se verán como líneas irregulares.

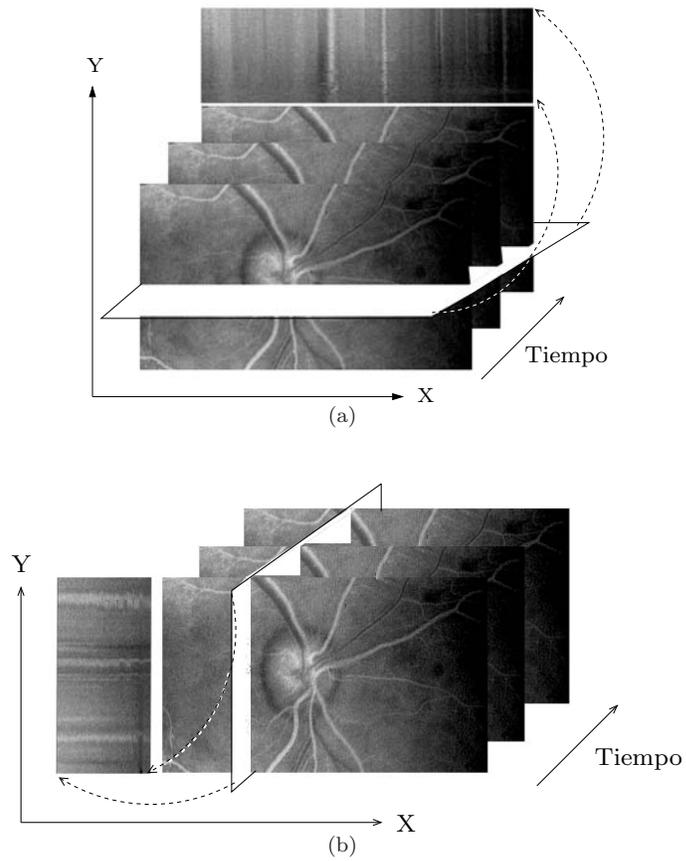


Figura 4.7: Esquema de construcción de las imágenes para la evaluación visual del resultado del alineamiento. Las zonas superior e izquierda son el resultado de cortar con un plano horizontal (a) y otro vertical (b), el cubo formado por el apilamiento de los fotogramas de la secuencia.

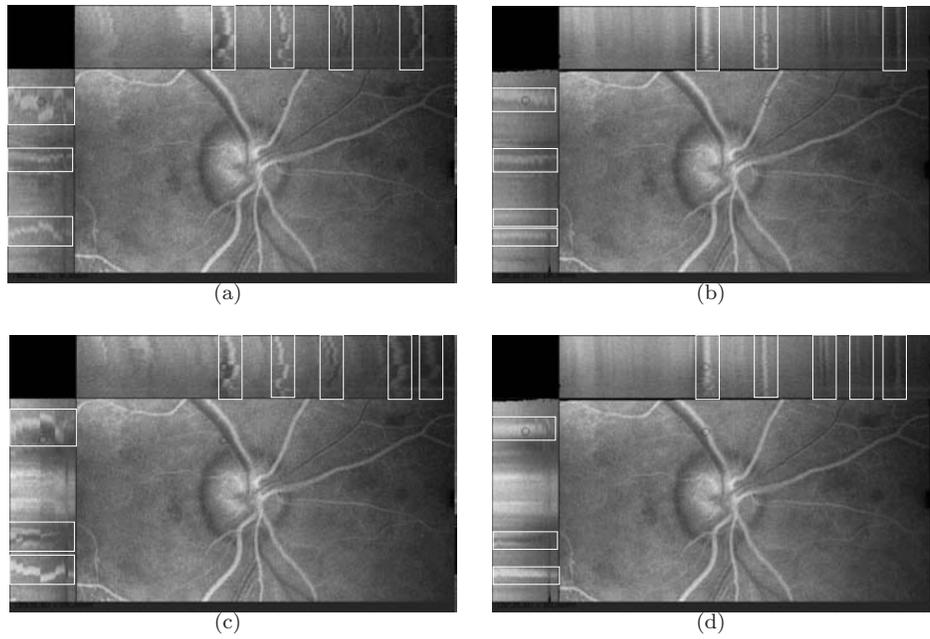


Figura 4.8: Imagen construida apilando los fotogramas de la secuencia. En (a) y (c) los fotogramas no están alineados, dando lugar a discontinuidades en los vasos (bandas zigzagueantes dentro de los rectángulos blancos señalados en las áreas superior e izquierda). (b) y (d) son los fotogramas una vez que han sido alineados, mostrando los vasos como líneas rectas (dentro de los rectángulos de las áreas superior e izquierda).

Capítulo 5

Análisis y evaluación de resultados

En este capítulo veremos los resultados obtenidos mediante el método de alineamiento de las secuencias de *SLO* descrito en esta memoria. En primer lugar presentaremos los resultados que justifican la utilización de los dos métodos elegidos para el alineamiento, para a continuación analizar los resultados obtenidos para varias secuencias completas. Además compararemos estos resultados con los obtenidos manualmente por varios expertos clínicos sobre las mismas secuencias de *SLO*, lo cual nos servirá de validación final, demostrando que nuestro método representa una herramienta enormemente útil en el entorno hospitalario.

El análisis de las secuencias capturadas nos lleva a acotar el rango de posibles transformaciones que sufre el ojo a lo largo de una secuencia. Dichos valores están incluidos en la Tabla 5.1. En los experimentos que presentamos en este capítulo restringiremos los valores de las transformaciones aplicadas a estos rangos.

Por último hemos de comentar que, debido al reducido valor que toma la variación del factor de escalado tanto en X como en Y , se han eliminado dichos factores de los experimentos. Así, se considerará que la escala no varía. Esto no afectará a la precisión requerida para el objetivo final, que es la obtención de las curvas de dilución

Parámetro	Valor mínimo	Valor máximo
Traslación en x	-100	+100
Traslación en y	-100	+100
Rotación (en grados)	-5	+5
Factor de escala en x	0,98	1,02
Factor de escala en y	0,98	1,02

Tabla 5.1: Rango de valores que puede alcanzar cada parámetro de una transformación afín para las imágenes de las secuencias *SLO*.

que permitirán llevar a cabo el cálculo de variables hemodinámicas como el tiempo arteria-vena.

5.1. Validación de las técnicas alineamiento

Para validar la robustez de los métodos de alineamiento (tanto el basado en la extracción de características como el basado en el cálculo de la *información mutua*), se han diseñado experimentos de forma que los fotogramas presenten condiciones muy diferentes. Cada experimento ha sido aplicado a 200 fotogramas elegidos aleatoriamente de cada una de las 5 secuencias de *SLO* que hemos venido mostrando a lo largo de este trabajo. Lo que no ha sido elegido aleatoriamente son los dos fotogramas de referencia de cada secuencia. Como se ha comentado anteriormente, dichos fotogramas han de ser seleccionados para que, por un lado, no presenten ningún problema de adquisición, y por otro sean representativos de la sección de la secuencia *SLO* en la que servirán como referencia para alinear los demás fotogramas. Cada fotograma se alinea tomando como referencia una u otra imagen de referencia en función de la sección a la que pertenezca.

Los experimentos cuyos resultados analizaremos son los siguientes:

1. **Experimento 1:** Aplicación de una transformación con traslación como única componente (sin rotación).

2. **Experimento 2:** Aplicación de una transformación con rotación como única componente (sin traslación en la transformación aplicada, pero posiblemente con traslación en la transformación recuperada).
3. **Experimento 3:** Aplicación de una transformación con rotación y traslación como componentes.

Además deberemos tener en cuenta que la componente de rotación no se aplica a los fotogramas en la sección 1 de la figura 4.1, es decir, aquellos en los que se utiliza la *información mutua* como técnica de alineamiento, ya que como hemos comentado en estos fotogramas sólo se tienen en cuenta los parámetros de traslación.

Para evaluar los resultados se utilizó la misma técnica que en la sección 2.1.3, que consiste en ejecutar, para 5 pacientes distintos, el proceso de ajuste con 200 fotogramas de cada una de las secciones en las que se dividen las secuencias *SLO* (ver Figura 2.15) dependiendo del nivel de fluoresceína presente en los vasos. A cada uno de estos fotogramas se le aplicaron 50 transformaciones aleatorias, con valores de ± 100 píxeles de traslación y $\pm 5^\circ$ de rotación. Estas cotas para las transformaciones aleatorias representan los valores máximos presentes en las secuencias *SLO* analizadas.

5.1.1. Experimento 1

En el experimento número 1 se aplica una transformación aleatoria con traslación (entre -100 y 100 píxeles) como única componente. Este proceso se aplica a 200 fotogramas seleccionados también aleatoriamente en cada una de las 5 secuencias *SLO* analizadas. La Tabla 5.2 muestra los resultados obtenidos. La primera columna de dicha tabla indica el número de transformaciones que han sido recuperadas con éxito, la segunda indica el error cuadrático medio de los alineamientos, la tercera el tiempo medio de cálculo necesario para cada alineamiento, mientras que las dos últimas indican la media μ y desviación típica σ de las diferencias entre las traslaciones aplicadas ($T_a = \sqrt{t_x^2 + t_y^2}$) y las traslaciones recuperadas ($T_r = \sqrt{t'_x{}^2 + t'_y{}^2}$), donde (t_x, t'_x) representan las traslaciones en el eje *X* y (t_y, t'_y) representan las traslaciones en el eje *Y*.

	Transformaciones con éxito (%)	ECM $\times 10^{-3}$	\bar{t} (seg.)	μ $\times 10^{-4}$	σ $\times 10^{-3}$
A	98	1,6	2,868	1,1	3,2
B	94	5,7	3,324	-4,0	3,8
C	90	36,8	3,877	26,0	566,0
D	96	1,3	3,302	-9,0	7,6
E	94	0,2	3,415	1,7	3,4

Tabla 5.2: Estadísticas del experimento realizado para analizar el resultado del alineamiento de fotografías a los que se les aplica una transformación aleatoria con componente única de traslación: la primera columna indica el porcentaje de alineamientos recuperados con éxito, la segunda indica el Error Cuadrático Medio de las transformaciones recuperadas y la tercera el tiempo necesario para el alineamiento (en segundos) de cada par de fotografías. Las columnas cuarta y quinta muestran la media (μ) y la desviación típica (σ) de la diferencia entre las transformaciones recuperadas y las aplicadas.

Podemos concluir a partir de los datos de la Tabla 5.2 que la componente de traslación se puede recuperar con éxito en la gran mayoría de los casos, y que la precisión del método de alineamiento cuando no existe componente rotacional en la transformación a recuperar es muy alta, igual o superior al 90% en todos los casos presentados.

La Figura 5.1 muestra gráficamente el resultado del experimento 1. En ella aparecen las diferencias entre las traslaciones recuperadas (T_r) y las aplicadas (T_a) para las distintas secuencias analizadas. Se aprecia claramente cómo la precisión del método permite recuperar con gran precisión las transformaciones aplicadas a los fotografías.

Hay que comentar que la elevada variabilidad (σ) de la tercera fila de la Tabla 5.2 se debe a que esta secuencia presenta un número alto de fotografías con errores en la adquisición. Esta dispersión se aprecia perfectamente en la gráfica de la Figura 5.2, donde se ve claramente cómo los puntos con un error elevado vienen dados por errores de alineamiento de fotografías de la secuencia C.

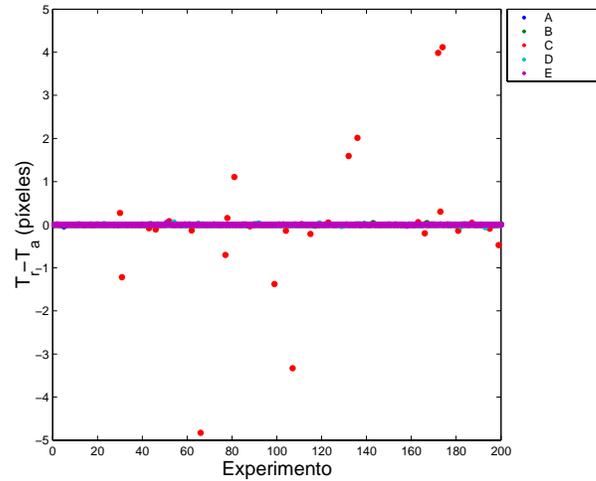


Figura 5.1: Resultados obtenidos del primer experimento de alineamiento de fotografías. El eje X señala los 200 fotografías alineados en cada secuencia, mientras que el eje Y muestra la diferencia entre la transformación recuperada T_r y la aplicada T_a , que idealmente debería ser 0, y que en la realidad se aproxima con gran precisión a este valor en la mayoría de los casos.

5.1.2. Experimento 2

En el experimento número 2 se aplica una transformación aleatoria con la rotación ($\phi \in [-5, 5]$) como única componente. Este proceso se aplica a 200 fotografías seleccionados también aleatoriamente en cada una de las 5 secuencias *SLO* analizadas. La Tabla 5.3 muestra los resultados obtenidos en este experimento. La primera columna indica el porcentaje de alineamientos recuperados con éxito, la segunda indica el error cuadrático medio de las transformaciones recuperadas, la tercera el tiempo medio necesario para el alineamiento (en segundos) de cada par de fotografías. Las columnas tercera y cuarta indican la media (μ) y desviación típica (σ) de la diferencia entre las rotaciones aplicadas (ϕ_a) y las rotaciones recuperadas (ϕ_r).

De la Tabla 5.3 se puede extraer que a pesar de que se ha reducido el porcentaje de alineamientos con éxito, éste sigue siendo elevado para todas las secuencias. La explicación de esta reducción en la precisión del alineamiento hay que buscarla en la mayor dificultad que presenta alinear imágenes que difieren en una rotación que ante diferencias debidas a traslaciones.

	Transformaciones con éxito (%)	ECM	\bar{t} (seg.) $\times 10^{-3}$	μ $\times 10^{-3}$	σ $\times 10^{-3}$
A	90	131,0	3,868	-93,0	55,0
B	84	28,0	3,432	20,0	57,0
C	91	8,0	3,977	56,0	55,0
D	86	13,0	3,802	-95,0	58,0
E	84	15,0	3,415	10,0	58,0

Tabla 5.3: Estadísticas del experimento realizado para analizar el resultado del alineamiento de fotografías a los que se les aplica una transformación aleatoria con componente única de rotación: la primera columna indica el porcentaje de alineamientos recuperados con éxito, la segunda indica el error cuadrático medio de las transformaciones recuperadas, la tercera el tiempo necesario para el alineamiento (en segundos) de cada par de fotografías. Las columnas tercera y cuarta indican la media (μ) y desviación típica (σ) de la diferencia entre las rotaciones aplicadas (ϕ_a) y las rotaciones recuperadas (ϕ_r).

La Figura 5.2 muestra gráficamente el resultado del experimento 2. En ella aparecen las diferencias entre las rotaciones recuperadas (ϕ_r) y las aplicadas (ϕ_a) para las distintas secuencias analizadas.

5.1.3. Experimento 3

Por último, en el experimento 3 se aplica una transformación aleatoria con traslación y rotación, combinación de los dos experimentos anteriores, y que reflejará de forma más precisa el tipo de deformaciones que se encuentran en los fotografías pertenecientes a secuencias *SLO* y que la técnica propuesta tendrá que corregir. Este proceso se aplica a 200 fotografías seleccionados también aleatoriamente en cada una de las 5 secuencias *SLO* analizadas. La Tabla 5.4 muestra los valores resultantes de este experimento. La medida del error se tomó a partir de la diferencia entre la transformación aplicada (T) y la recuperada (T') utilizando la ecuación 5.1, que representa la magnitud de dicha diferencia tomada como un vector con dos componentes: traslación y rotación.

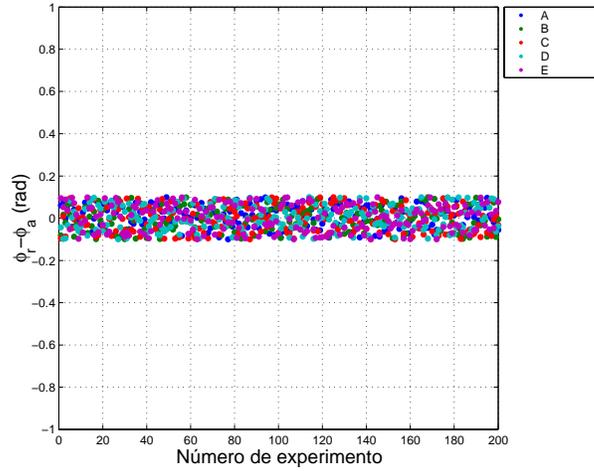


Figura 5.2: Resultados obtenidos del segundo experimento de alineamiento de fotogramas. El eje de X de la gráfica señala los 200 fotogramas alineados de cada una de las secuencias, mientras que el eje Y muestra la diferencia entre la rotación recuperada (ϕ_r) y la rotación aplicada (ϕ_a).

$$error = |T' - T| = \sqrt{(T_r - T_a)^2 + (\phi_r - \phi_a)^2} \quad (5.1)$$

donde T_r representa la traslación recuperada, T_a representa la traslación aplicada a la imagen, ϕ_r es el ángulo de rotación recuperado y ϕ_a representa el ángulo de rotación aplicado con la transformación aleatoria.

Para medir la diferencia entre dos transformaciones se utilizó la técnica descrita en 2.1.3, que consiste en seleccionar manualmente 4 puntos en la imagen dinámica registrada, de forma que estos se utilizan como puntos de referencia, y que permitirán evaluar el error del alineamiento mediante la aplicación de la ecuación 2.16.

Obsérvese que los resultados representados en la tabla 5.4 siguen siendo muy buenos, ya que el porcentaje medio de fotogramas alineados correctamente es del 92%, a pesar de que varía un poco en función de las características de las imágenes que componen la secuencia, lo cual como ya hemos comentado anteriormente es un problema difícil de tratar. Además el tratamiento individualizado de estos fotogramas implicaría un aumento en el tiempo de procesamiento de la secuencia lo cual, tal como veremos a continuación, no representaría una mejora sustancial, ya que estos

	Transformaciones con éxito (%)	ECM	\bar{t} (seg.) $\times 10^{-3}$	μ $\times 10^{-3}$	σ $\times 10^{-3}$
A	95	1,010	3,180	0,0007	0,0277
B	94	1,904	3,242	-1,0356	0,8807
C	89	3,021	3,767	-2,4187	1,7237
D	92	0,916	2,602	-1,0790	0,9690
E	94	0,872	2,915	0,9925	0,9490

Tabla 5.4: Estadísticas del experimento realizado para analizar el resultado del alineamiento de fotografías a los que se les aplica una transformación aleatoria con traslación y rotación: la primera columna indica el porcentaje de alineamientos recuperados con éxito, la segunda indica el error cuadrático medio de las transformaciones recuperadas (ecuación 5.1) y la tercera el tiempo necesario para el alineamiento de cada par de fotografías. Las columnas cuarta y quinta muestran la media (μ) y la desviación típica (σ) de la diferencia entre las transformaciones recuperadas y aplicadas.

fotogramas mal alineados son eliminados del proceso de cálculo del resultado final mediante un método de ajuste robusto.

La Figura 5.3 muestra gráficamente los resultados del experimento. El eje de las ordenadas contiene cada experimento realizado para cada una de las 5 secuencias analizadas, mientras que el eje Y representa la magnitud de la diferencia entre la transformación recuperada y la aplicada, calculada utilizando la ecuación 5.1.

Hay que destacar que los errores que aparecen en la Figura 5.3 pertenecen en su mayoría, al igual que ocurría en el primer experimento, a fotografías de la secuencia C mal alineados. Estos casos de alineamiento erróneo son provocados mayoritariamente por fotografías con errores en la adquisición. Sin embargo, existe también un pequeño número de fotografías en las otras secuencias (de media un 5%) que no son correctamente alineados. Estos errores se deben a que, a pesar de la robustez de los algoritmos de alineamiento, en determinados casos, la imagen de referencia y la transformada difieren de tal manera que el algoritmo de alineamiento no consigue ajustarlas correctamente. Estos errores son de todas formas pequeños, ya que fijándonos en el eje Y vemos que el valor máximo del error es de 5 pixels.

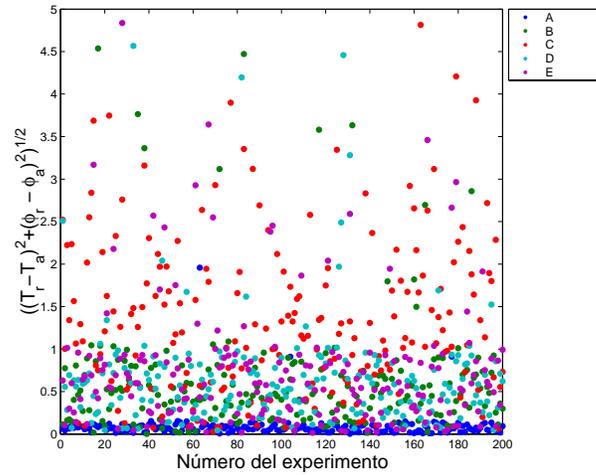


Figura 5.3: Resultados obtenidos del tercer experimento de alineamiento de fotogramas. El eje X señala los 200 fotogramas alineados de cada una de las secuencias, mientras que el eje Y muestra la diferencia entre la transformación recuperada y la aplicada, que idealmente debería ser 0, y que en la realidad se aproxima con gran precisión a este valor en la mayoría de los casos.

Una vez que hemos verificado que el algoritmo diseñado funciona correctamente con las imágenes de *SLO*, utilizaremos esta técnica para el cálculo de una de las variables hemodinámicas más usadas por los oftalmólogos, el tiempo arteria-vena. El método se validará comparando los resultados obtenidos por oftalmólogos expertos utilizando el procedimiento manual frente a los obtenidos automáticamente utilizando el método que se describe en este trabajo.

5.2. Validación clínica

Una de las herramientas que los oftalmólogos utilizan para el cálculo de los parámetros hemodinámicos son las denominadas *curvas de dilución*, que representan la evolución del nivel de gris en un punto de la imagen a lo largo de una serie de fotogramas de la secuencia. En esta sección se explica cómo se calculan estas curvas, cómo se utilizan y se validarán los resultados obtenidos.

5.2.1. Cálculo de las curvas de dilución

El estudio de la circulación retinal y su regulación ha demostrado que los pacientes con diabetes retinopática muestran anomalías en las mismas [Kohner y col., 1975, Yoshida y col., 1983, Grunwald y col., 1986, Patel y col., 1992, Bursell y col., 1992, Feke y col., 1994]. Además, diversos estudios clínicos han demostrado que el flujo sanguíneo retinal se incrementa una vez que la retinopatía diabética ha alcanzado niveles avanzados y no proliferativos. Para analizar el flujo sanguíneo retinal se utilizan parámetros como el tiempo brazo-vena, el tiempo arteria-vena o la velocidad media arterial del contraste inyectado. En este trabajo se describe cómo calcular el tiempo arteria-vena (TAV) a partir de las secuencias de *SLO*.

El tiempo arteria-vena es el período de tiempo entre que el primer influjo del colorante inyectado (fluoresceína en nuestro caso) alcanza una arteria retinal y su aparición en la correspondiente vena retinal. Es el tiempo de circulación más pequeño del indicador (colorante) a través de un área de microcirculación retinal. El procedimiento habitual en el Servicio de Oftalmología del *Complejo Hospitalario Universitario de Santiago de Compostela (CHUS)* para adquirir secuencias de *SLO* comienza con la inyección de 3 ml de fluoresceína sódica en la vena antecubital del brazo derecho del paciente. A continuación, el oftalmoscopio (SLO-101, Rodenstock con láser de Argon, 488 nm) en modo de fluoresceína, comienza la adquisición de las imágenes. Éstas se digitalizan mediante una tarjeta digitalizadora Matrox RT5000 en un PC convencional a una velocidad de 25 fotogramas por segundo durante 40 segundos, con una resolución de 720×576 píxeles y 256 niveles de gris. Las imágenes se capturan con un campo angular de visión de 40° .

A continuación el oftalmólogo alinea manualmente los fotogramas de la secuencia y extrae una distribución de puntos que representan la evolución del nivel de gris en un punto de la imagen a lo largo de una serie de fotogramas de la secuencia, a partir de la que calculará las denominadas *curvas de dilución* [Yang y col., 1995, Sperber y Alm, 1997, Pesavento y col., 2002]. Otra posibilidad a la hora de construir la distribución de puntos consiste en tomar, en lugar del valor del píxel, el valor medio en un entorno del mismo, consiguiéndose de esta forma una variabilidad mucho menor, y unas mediciones más fiables. En la Figura 5.4 se muestra una comparación de estas

posibilidades. En la figura 5.4(a) se representa la evolución del nivel de gris en un punto de una secuencia cualquiera tomando como valores los niveles de gris en un punto (señalado en la Figura 5.4(b), en la que aparece indicado con una flecha el píxel en el que se tomaron los valores), los valores medios de los 4 píxeles vecinos, y los valores medios de los 8 píxeles vecinos, junto con la desviación típica de cada una de ellas. De la desviación típica de cada curva (14,235, 12,052 y 11,854 para el valor en el punto, promedio de los 4 vecinos y promedio de los 8 vecinos respectivamente) se concluye que resulta conveniente tomar el promedio de los 8 vecinos, ya que así las mediciones ganan en robustez y fiabilidad.

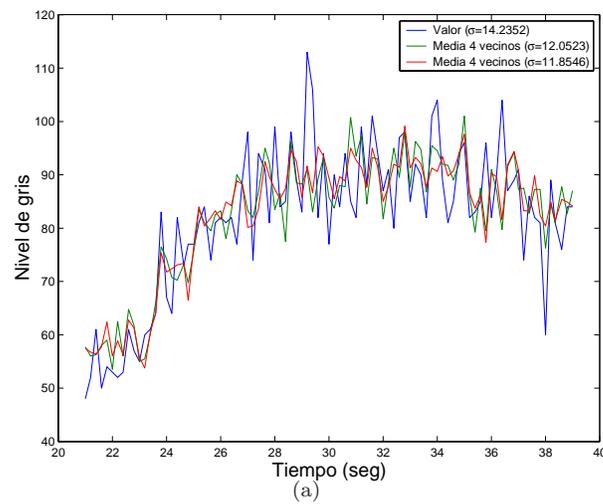


Figura 5.4: (a) Curvas obtenidas al tomar el valor de gris del píxel, el valor de gris promedio de los 4 vecinos o el valor de gris promedio de los 8 vecinos, sobre el punto mostrado en (b). Además en (a) también se incluye la desviación típica de cada una de las curvas, apreciándose claramente cómo disminuye la variabilidad al aumentar el número de píxeles considerados.

Si el alineamiento de los fotogramas ha sido realizado con éxito, la anchura de la evolución del conjunto de puntos muestreados no debe ser excesiva, debido a que los cambios de intensidad de los niveles de gris no pueden ser instantáneos, y de serlo vendrán causados por errores en el proceso de adquisición como los mostrados en la Figura 5.5. La Figura 5.5(a) indica los puntos muestreados en los fotogramas mal alineados dentro de la nube de puntos, donde el punto marcado como $P1$ se corresponde con el fotograma mostrado en 5.5(b), en el que se produjo un error de sincronismo entre el dispositivo de adquisición de imágenes y el dispositivo digitalizador, mientras que el punto $P2$ se corresponde con el error de adquisición ilustrado por la Figura 5.5(c), donde se puede ver una imagen en la que el paciente parpadeó, ocultando la estructura arterio-venosa. Ambas imágenes conducen a errores en el proceso de alineamiento que se reflejan en los valores atípicos que aparecen en las nubes de puntos muestreadas.

El método empleado para estimar las curvas de dilución a partir de las nubes de puntos muestreadas de los fotogramas está basado en el utilizado por Koyama y col. [Koyama y col., 1990] que representaron los datos obtenidos con el tiempo en el eje de abscisas y la intensidad del píxel en el eje de ordenadas, y ajustaron por mínimos cuadrados la nube de puntos a la función:

$$I_e = I_0 + I_p \cdot e^{-A \ln^2 \frac{T-T_0}{T_p-T_0}} \quad (5.2)$$

donde I_0 representa la intensidad del primer fotograma considerado, I_p es el máximo valor de intensidad obtenido, A es un parámetro de forma, T_0 representa el número del primer fotograma considerado y finalmente T_p representa el número del fotograma en el que se alcanza el máximo valor de la intensidad.

En este trabajo no hemos utilizado el ajuste por mínimos cuadrados, pues precisábamos un método de ajuste más robusto, ya que existen algunos valores atípicos causados por fotogramas con errores como los mostrados en la Figura 5.5 que podrían llevar a un ajuste poco preciso. En estas imágenes erróneas, ni la alineación basada en crestas ni la basada en información mutua funcionan bien, ya que los fotogramas de referencia son demasiado diferentes del fotograma dinámico. Además estos fotogramas no ofrecen ningún tipo de información relevante al oftalmólogo ni al proceso de

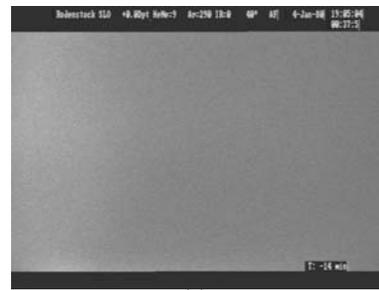
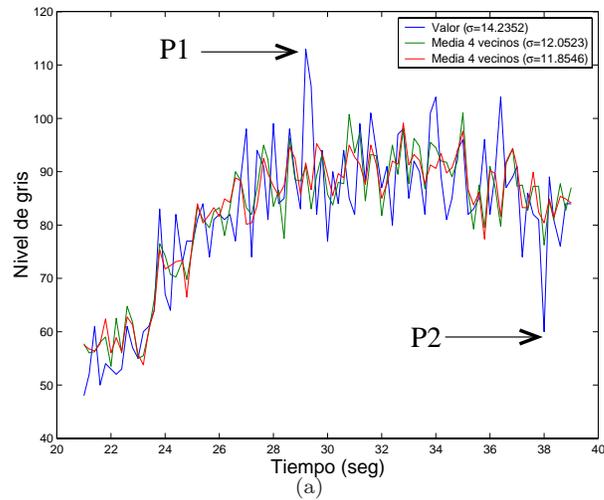


Figura 5.5: Ejemplos de fotogramas que se eliminan de las secuencias, ya que producen errores en el cálculo de las curvas de dilatación. (a) muestra la evolución del nivel de gris señalado en el fotograma de la Figura 5.4(b), señalando con los puntos $P1$ y $P2$ dos fotogramas que han sido alineados incorrectamente y que se muestran en (b) y (c). (b) es un fotograma que se capturó de forma errónea debido a problemas de sincronización entre el dispositivo de captura de las imágenes y el de digitalización de las mismas (marcado como $P1$ en (a)). (c) es un fotograma con error debido a que el paciente parpadeó mientras se estaba capturando la secuencia (marcado como $P2$ en (a)). Ambas imágenes serán eliminadas de los resultados mediante el proceso de ajuste.

alineamiento, por lo que deben ser eliminados de los resultados. El método del *Temple Simulado* [Press y col., 1992] demostró ser un buen método para el ajuste de la curva a la nube de puntos. Los cinco parámetros de la ecuación 5.2 son las componentes de nuestro espacio de búsqueda, y la función de coste viene dada por:

$$\mathcal{C} = \sum_i (I_i - I_e)^2 \quad (5.3)$$

donde i es el instante temporal al que se corresponde el fotograma dentro del proceso de adquisición, I_i representa el valor de la intensidad en este instante y I_e es el valor estimado para la intensidad calculada mediante la ecuación 5.2.

El pseudocódigo resumido en el Algoritmo 1 resume el proceso de ajuste a la nube de puntos utilizando el método del Temple Simulado, ofreciendo como resultado el valor para cada uno de los cinco parámetros que determinan la forma de la curva en la ecuación 5.2, y que en dicho pseudocódigo se denominan x_1, x_2, x_3, x_4, x_5 representando el ajuste para los parámetros I_0, I_p, A, T_0 y T_p , respectivamente. El ajuste finaliza si la diferencia entre dos iteraciones de la función de coste (ecuación 5.3) no supera un umbral ϵ determinado heurísticamente, o al llegar a tres iteraciones, que es el número mínimo de iteraciones que se comprobaron son necesarias para obtener un ajuste satisfactorio de la curva final a la nube de puntos.

A partir de los experimentos realizados con este método de ajuste, encontramos que, aunque los valores iniciales de los parámetros estuvieran alejados del resultado final, el método es lo suficientemente preciso y robusto como para seguir el camino correcto en la búsqueda. Por esto los valores iniciales de los parámetros se fijaron para todas las secuencias con los datos incluidos en la Tabla 5.5, excepto para los valores de A y T_p , que vienen dados simplemente por el número de puntos en la secuencia.

La Figura 5.6 ilustra la evolución del proceso de ajuste a una nube de puntos obtenida de una secuencia. La Figura 5.6(a) muestra la nube de puntos y la curva de partida, obtenida con los valores iniciales de los parámetros de la curva calculada mediante la ecuación 5.2, que se muestran en el Algoritmo 1. En la Figura 5.6(b) se muestran las distintas curvas que se generan con los parámetros obtenidos en cada una de las iteraciones del método de ajuste descrito. Finalmente, la Figura 5.6(c)

Algoritmo 1 Pseudocódigo para el método de ajuste robusto a la nube de puntos muestreados utilizando como técnica de optimización el *Temple Simulado*.

```

1:  $I \leftarrow \{I_i\}$ , conjunto de los puntos muestreados
2:  $N \leftarrow \text{card}(I)$ , número de elementos del conjunto  $A$ 
3:  $x \leftarrow \{x_i\}$ , conjunto de parámetros a ajustar
4:  $x_1 \leftarrow \mu_{inf}$ , donde  $\mu_{inf}$  representa la media de los puntos situados por debajo de la media del total de puntos muestreados
5:  $x_2 \leftarrow (\mu_{sup} - \mu_{inf})$ , donde  $\mu_{sup}$  representa la media de los puntos situados por encima de la media del total de puntos muestreados
6:  $x_3 \leftarrow N$ 
7:  $x_4 \leftarrow T_0$ , donde  $T_0$  es el instante temporal del primer fotograma considerado
8:  $x_5 \leftarrow (T_0 + N)$ 
9:  $k \leftarrow 0, \sigma_e \leftarrow \text{Infinite}$ 
10: while  $k < 3$  AND  $\sigma_e > \epsilon$  do
11:    $x' \leftarrow TS(I, x)$ , ajuste utilizando el Temple Simulado
12:   calcular  $I_e$  utilizando la ecuación 5.2,  $I$  y  $x'$ 
13:   calcular  $C$  utilizando la ecuación 5.3
14:   calcular  $\sigma_e$ , desviación típica de los puntos
15:   for all  $I_i \in I$  do
16:     if  $I_i \notin [I_e - \sigma_e, I_e + \sigma_e]$  then // valor atípico, se elimina
17:       eliminar  $(I_i)$  de  $I$ 
18:     end if
19:      $N = \text{card}(I)$ 
20:      $k = k + 1$ 
21:      $x = x'$ 
22:   end for
23: end while

```

Parámetro	Valor inicial
I_0	100
I_p	100
A	Número de fotogramas considerados
T_0	1
T_p	Número de fotogramas considerados

Tabla 5.5: Valores iniciales para los parámetros de la ecuación 5.2.

muestra cómo queda ajustada la curva a la nube de puntos con los valores finales de los parámetros.

Una vez calculada la curva de dilución, es necesario eliminar los valores atípicos del conjunto de puntos, que como se señaló anteriormente se consideran registros erróneos posiblemente debidos a errores en la adquisición de las imágenes. Un punto I_i es considerado un valor atípico si se cumple la ecuación 5.4:

$$I_i \in [I_e - \sigma_e, I_e + \sigma_e] \quad (5.4)$$

donde I_i representa el valor de la intensidad en el punto, I_e el valor de la curva ajustada en ese punto y σ_e representa la desviación típica de los puntos de la muestra. En la Figura 5.7 se muestran dos imágenes: en 5.7(a) aparecen dos curvas, que señalan los límites entre la zona de puntos aceptados y la zona de puntos rechazados, mientras que en 5.7(b) aparecen no sólo estas curvas, sino que también se destacan aquellos puntos que no cumplen la condición 5.4.

Finalmente, para evaluar la mejora que implica el proceso de eliminación de valores atípicos, la Tabla 5.6 muestra las estadísticas de fotogramas alineados correctamente en los experimentos 1, 2 y 3 descritos previamente, antes y después de aplicar el proceso de eliminación de valores atípicos que acabamos de comentar. Como se puede apreciar en dicha tabla, una vez eliminados los valores atípicos se produce una sensible mejora en el porcentaje de fotogramas alineados correctamente en todos los experimentos realizados, como muestra el aumento del porcentaje medio de fotogramas alineados correctamente en cada experimento (última fila de la Tabla 5.6).

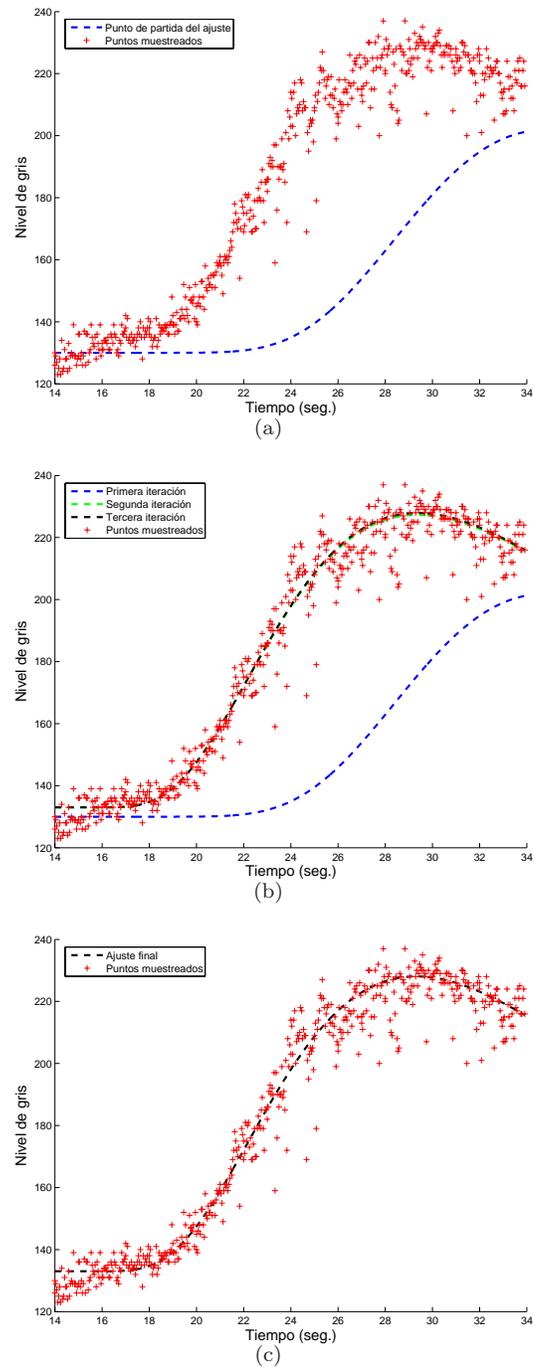
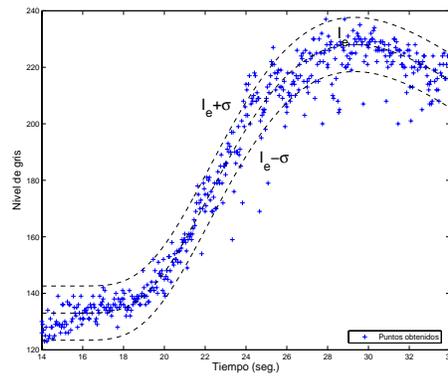
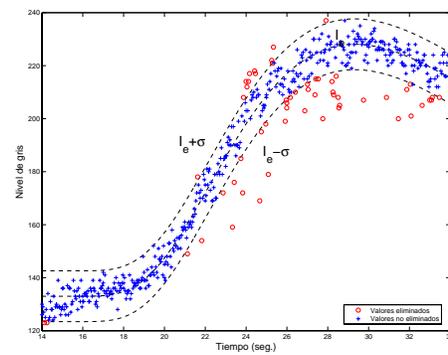


Figura 5.6: Ilustración del proceso de ajuste para estimar la curva de dilución a partir de los puntos muestreados. (a) Nube de puntos con la estimación inicial de la curva de ajuste, utilizando los valores de los parámetros mostrados en el pseudocódigo 1. (b) Varias iteraciones del proceso de ajuste. (c) Valor final de la curva ajustada a la nube de puntos.



(a)



(b)

Figura 5.7: Representación de los valores atípicos eliminados. (a) nube de puntos obtenida al muestrear un punto a lo largo de toda la secuencia, con las dos curvas (de la forma $I_e \pm \sigma_e$) que determinan los límites de los puntos que se aceptan y los que se rechazan por atípicos. (b) gráfica que resalta los puntos considerados como atípicos y que serán eliminados del proceso de cálculo de las curvas de dilución.

	Transformaciones con éxito (%)					
	Experimento 1		Experimento 2		Experimento 3	
	Antes	Después	Antes	Después	Antes	Después
A	98	99	90	92	95	98
B	94	97	84	89	94	98
C	90	93	91	94	89	93
D	96	97	86	91	92	95
E	94	96	84	89	94	96
% medio éxito	94,4	96,4	87,0	91,0	93,4	96,0

Tabla 5.6: Comparativa del porcentaje de fotogramas alineados correctamente antes y después de la aplicación del proceso de eliminación de puntos atípicos. Cada columna muestra, para cada uno de los tres experimentos realizados al principio de este capítulo, el porcentaje de fotogramas que se alinearon de forma correcta antes de la aplicación del proceso de eliminación de valores atípicos (columna *antes*) y tras la aplicación de dicho proceso (columna *después*). La última fila muestra el porcentaje medio de fotogramas alineados con éxito en cada experimento para el conjunto de los 5 pacientes analizados.

5.2.2. Validación de los resultados obtenidos

En esta sección presentaremos los resultados obtenidos en la medición automática del tiempo arteria-vena (TAV) [Hamilton y col., 1932, Wolf y col., 1989], objetivo final del trabajo presentado en esta memoria, utilizando como fuente las secuencias de *SLO* alineadas.

Para evaluar los resultados obtenidos, se han comparado las curvas de dilución obtenidas con nuestro método con aquellas obtenidas manualmente por un oftalmólogo utilizando el método habitual, es decir, eligiendo dos puntos en el fotograma de referencia (uno en una arteria y otro en una vena) e intentando seleccionar los mismos puntos a lo largo de las restantes imágenes de la secuencia, mientras se anotan los valores del nivel de gris de estos puntos en cada imagen para construir con ellos la curva de dilución (Figura 5.8). En esta gráfica, los valores obtenidos para la intensidad se dibujan en el eje de ordenadas frente al tiempo (eje de abscisas), mostrando la evolución del nivel de gris en los puntos seleccionados a lo largo del tiempo de

duración de la secuencia *SLO*. En la Figura 5.8(a) se muestra un fotograma en el que aparecen señaladas la arteria (*A*) y la vena (*V*) en las que se tomaron las muestras. La Figura 5.8(b) muestra los resultados obtenidos por los oftalmólogos empleando el método tradicional, mientras que la Figura 5.8(c) muestra el resultado obtenido utilizando la metodología descrita en este trabajo.

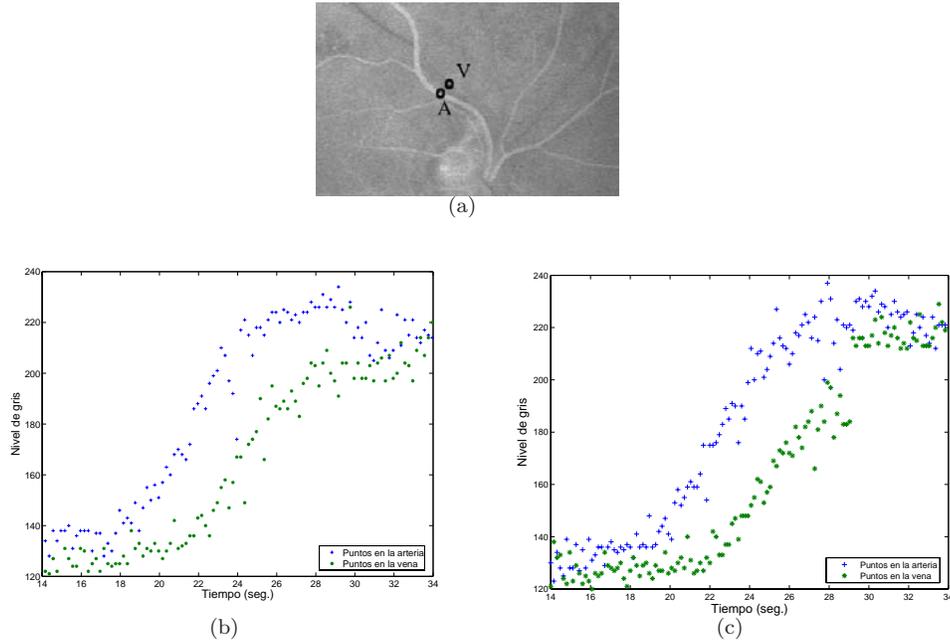


Figura 5.8: (a) Fotograma con los puntos en los que se toman las muestras, donde *A* señala la arteria y *V* señala la vena. (b) curva de dilución obtenida por un oftalmólogo experto utilizando el método tradicional, alineando manualmente los fotogramas y tomando el valor del punto a lo largo de la secuencia. (c) nube de puntos obtenida utilizando el método de alineamiento automático.

Hemos de destacar que para que los resultados sean lo más similares posible, en el cálculo automático de las curvas se han tomado, para cada secuencia, tan sólo uno de cada cinco fotogramas, ya que esta es la metodología que siguen los oftalmólogos, debido a que el alineamiento manual de los fotogramas es una tarea muy ardua. En la implantación real de la técnica se pueden utilizar todos los fotogramas que se desee. A partir de las nubes de puntos de la Figura 5.8 se estiman las curvas de dilución, utilizando el método descrito en la sección anterior. La Figura 5.9 muestra las curvas ajustadas a los puntos de dicha figura.

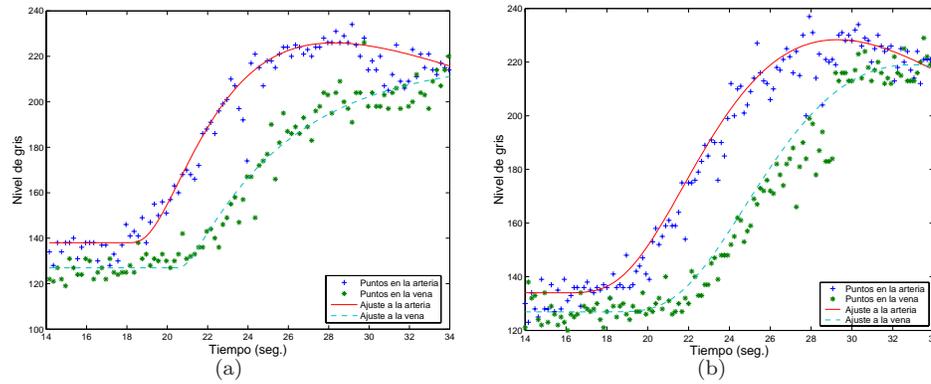


Figura 5.9: (a) Curva de dilución obtenida por un oftalmólogo experto utilizando el método tradicional, alineando manualmente los fotogramas y tomando el valor del punto a lo largo de la secuencia. (b) Nube de puntos obtenida utilizando el método de alineamiento automático.

Una vez obtenidas estas curvas, se pueden obtener los parámetros hemodinámicos, entre ellos el *Tiempo Arteria-Vena (TAV)*. Este tiempo puede estimarse utilizando la parte de la curva de dilución en la que la pendiente es ascendente, y tomando tanto para la arteria como para la vena la coordenada en X del centro de gravedad de dicha parte de la curva [Hamilton y col., 1932]. La diferencia entre los tiempos en la arteria y en la vena en dicha coordenada proporciona el TAV. La Figura 5.10 ilustra el procedimiento de cálculo del tiempo arteria-vena, tal y como acabamos de comentar, sobre los mismos conjuntos de datos utilizados en las figuras anteriores.

Es necesario añadir que los oftalmólogos suelen utilizar el TAV no como una medida en sí, sino como un indicador de respuesta ante un tratamiento aplicado a un paciente. Así, la metodología normal consiste en realizar un análisis del TAV antes de la aplicación del tratamiento, y otro análisis cierto tiempo después, cuando se considere que el tratamiento ha surtido el efecto necesario. Para un paciente con un TAV anormal, cualquier mejora causada por el tratamiento se verá reflejada en el segundo (y posteriores) análisis.

En la Figura 5.11 se muestran las curvas de dilución obtenidas para cinco secuencias de *SLO* pertenecientes a cinco pacientes distintos. En la columna de la izquierda se muestran las curvas obtenidas por los oftalmólogos utilizando el procedimiento manual, mientras que en la de la derecha se muestra el resultado obtenido utilizando la

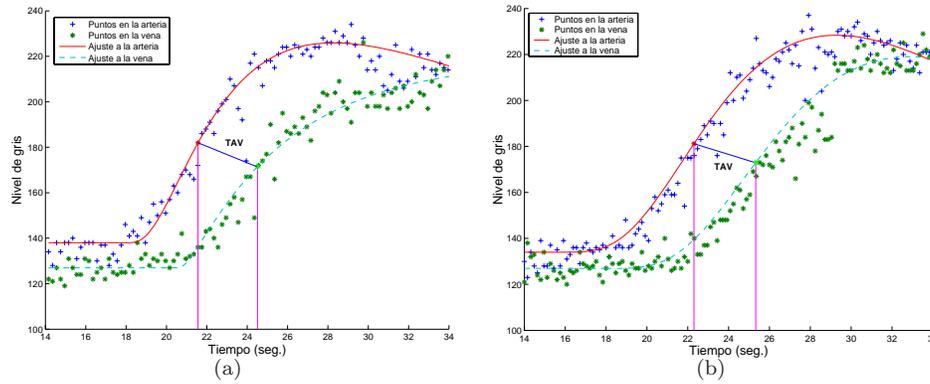


Figura 5.10: Cálculo del tiempo arteria-vena (TAV) como la diferencia entre el instante en el que la curva de dilución alcanza el centro de gravedad de la zona con pendiente ascendente, tanto para la arteria como para la vena. (a) TAV para el conjunto de datos obtenido manualmente. (b) TAV para el conjunto de datos obtenidos automáticamente.

metodología descrita en este trabajo.

La Tabla 5.7 contiene los tiempos arteria-vena obtenidos a partir de las imágenes de la Figura 5.11. La primera fila de las imágenes pertenece al paciente A, la segunda al paciente B, y así sucesivamente.

	TAV Manual (tav_m)	TAV Automático (tav_a)	$ tav_m - tav_a $
A	4,098	5,208	1,110
B	2,728	1,685	1,043
C	1,923	1,226	0,697
D	3,505	3,430	0,075
E	2,168	2,935	0,767

Tabla 5.7: Tiempos arteria-vena (en segundos) obtenidos utilizando el método manual y el método automático para las cinco secuencias analizadas.

A pesar de las diferencias evidentes entre los tiempos obtenidos por los oftalmólogos y los tiempos obtenidos por el método automático (también validados por los oftalmólogos), se puede asegurar que el sesgo en los valores obtenidos viene dado por las imprecisiones que se cometen al realizar los cálculos manualmente, sobre todo al

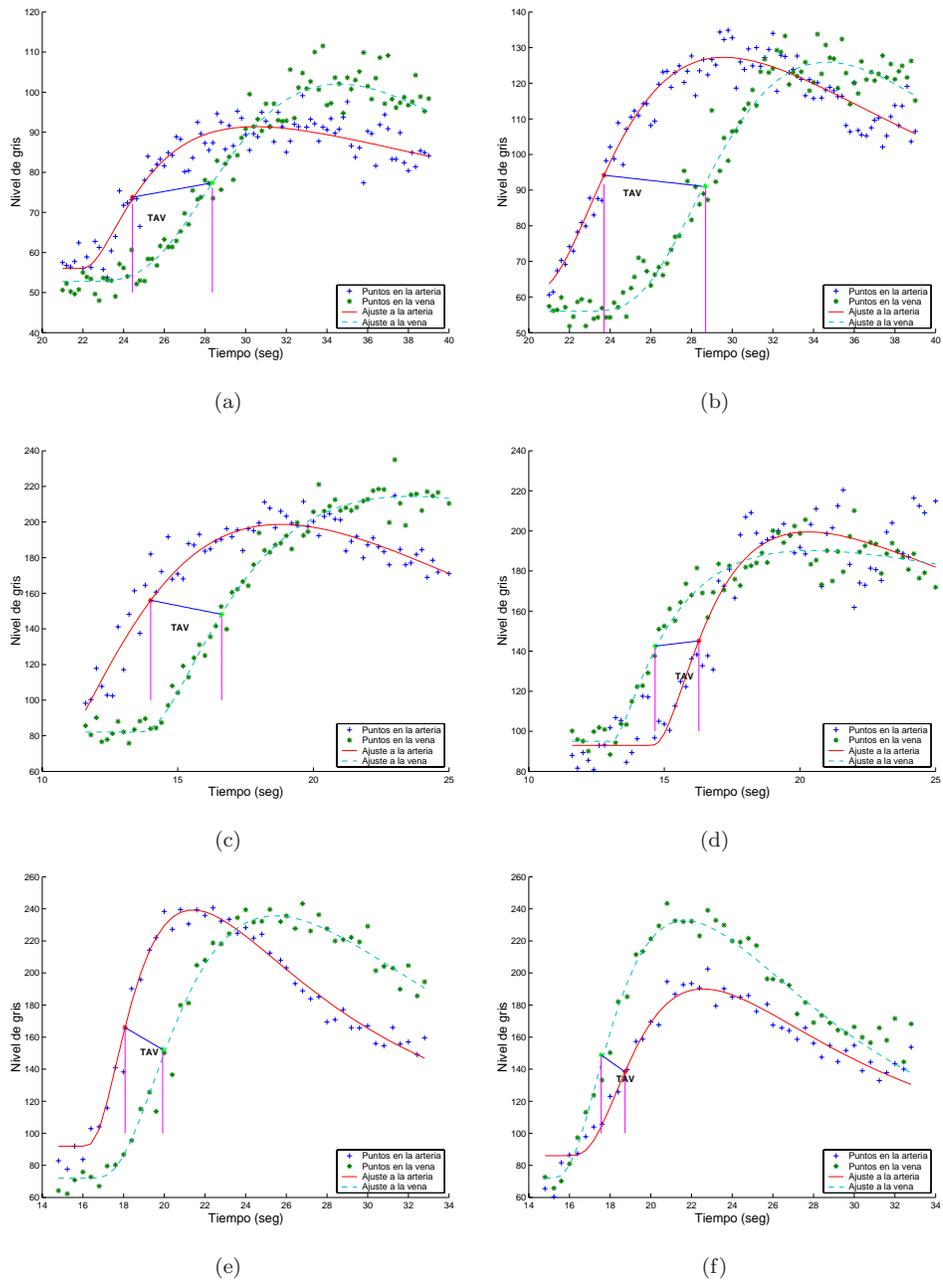


Figura 5.11: (sigue en la página siguiente)

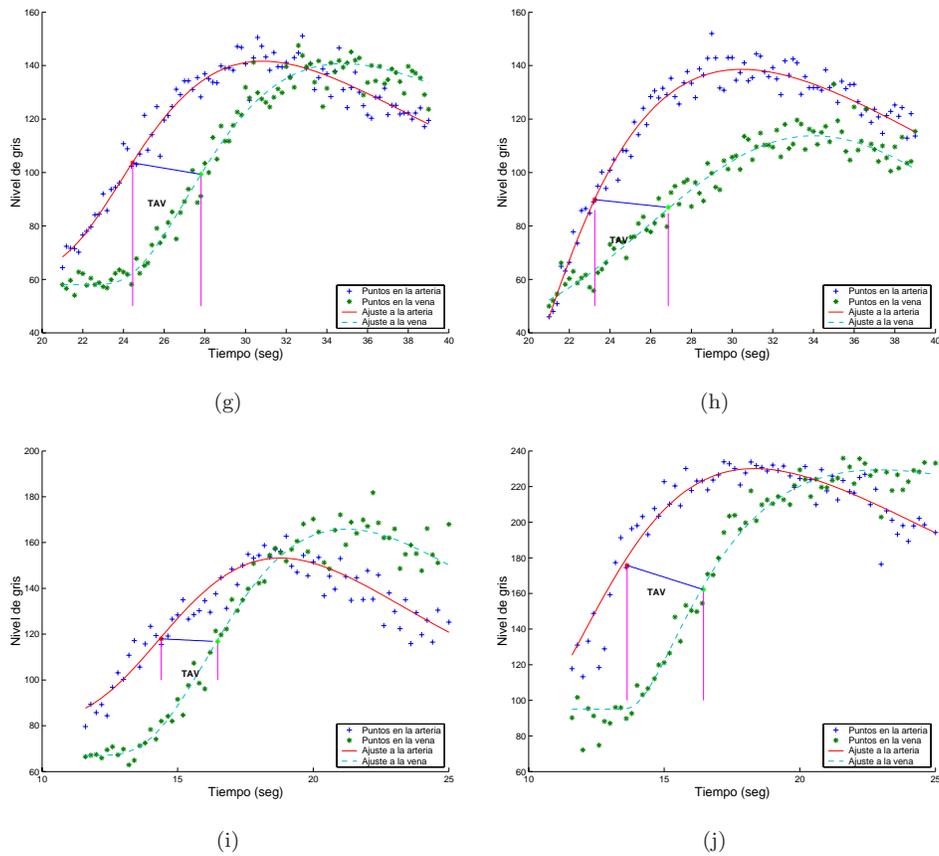


Figura 5.11: Curvas de dilución obtenidas al procesar cinco secuencias de *SLO*. En la columna de la izquierda se muestra la curva de dilución obtenida por los oftalmólogos utilizando el método manual, mientras que en la columna de la derecha aparecen las curvas de dilución obtenidas utilizando la metodología descrita en este trabajo. Los tiempos aparecen detallados en la Tabla 5.7, donde la correspondencia entre pacientes y resultados es, de arriba a abajo, A,B,C,D y E respectivamente.

alinearse manualmente los fotogramas sobre los que se toman los valores de gris para calcular las curvas de dilución. Además, esta variabilidad aparece documentada en la bibliografía, donde distintos autores han obtenido tiempos con gran variabilidad entre ellos. Por ejemplo, Bursell y col. [Bursell y col., 1996] obtienen valores de $3,1 \pm 0,7$ segundos para pacientes sin diabetes, mientras que para pacientes con diabetes los tiempos son de $5,1 \pm 1,8$ segundos. Por otro lado, Clermont y col. [Clermont y col., 1997] obtienen, para pacientes no diabéticos, valores de $3,5 \pm 0,9$ segundos, mientras que para pacientes con diabetes, dependiendo del tratamiento al que estén siendo sometidos, estos tiempos pueden variar entre $4,9 \pm 1,3$ hasta $3,3$ segundos.

Para los oftalmólogos, una de las características más relevantes del método automático es su *repetibilidad*, es decir, que cada vez que se ejecuta proporciona los mismos resultados, lo cual utilizando el método tradicional es prácticamente imposible. Sin embargo, empleando el alineamiento automático y el método de estimación de las curvas de dilución descrito anteriormente, los tiempos obtenidos son siempre los mismos. Además este método no está sujeto a las desviaciones propias de las técnicas manuales utilizadas anteriormente por los oftalmólogos: la variabilidad de los resultados dependiendo del sujeto que realiza el análisis y la influencia del cansancio acumulado tras el análisis de cientos de fotogramas.

Por otro lado, debe tenerse en cuenta que en el caso de los pacientes analizados no se ha tenido en cuenta si son diabéticos o no, o si están sometidos a algún tratamiento o no. Lo importante de este trabajo es obtener un método robusto y preciso para la medición del tiempo arteria-vena, lo cual consideramos que se ha conseguido. Este método permitirá llevar a cabo una mejora en la calidad asistencial a los pacientes, ya que gracias a él se podrá realizar el diagnóstico en un tiempo muy inferior y con una mayor fiabilidad, lo que reportará un tratamiento más temprano y eficaz. Además, gracias a la reducción del tiempo necesario para llevar a cabo la evaluación de cada paciente, el oftalmólogo verá reducido el tiempo dedicado al trabajo rutinario consistente en el análisis manual de los fotogramas de las secuencias, pudiendo así dedicar más tiempo a otras tareas. Otra de las ventajas que se obtendrán con la utilización de esta técnica es la mayor precisión y fiabilidad a la hora de evaluar los fármacos empleados. La reproducibilidad del método garantiza que se eliminan los errores cometidos por los oftalmólogos debido al cansancio que producen las tareas

repetitivas y prolongadas de análisis de muchas imágenes, como las que componen las secuencias de *SLO*, así como aquellos errores más subjetivos que aparecen cuando más de una persona analiza un conjunto de imágenes.

Capítulo 6

Conclusiones y líneas futuras de investigación

Las secuencias de imágenes obtenidas con el *SLO* suponen un gran avance en los métodos de obtención de imágenes del fondo del ojo, permitiendo un análisis más exhaustivo del mismo. También facilitan el cálculo de medidas hemodinámicas que permiten una prevención temprana de determinadas enfermedades oculares, así como un mejor seguimiento de la respuesta de los pacientes antes los tratamientos aplicados por los oftalmólogos. Sin embargo, para poder obtener este conjunto de medidas es necesario que el conjunto de los fotogramas que componen las secuencias estén alineados sobre un mismo eje de coordenadas. Los fotogramas de las secuencias presentan una serie de características que hacen que su alineamiento sea más complejo que otros tipos de imágenes: gran variabilidad del contraste a lo largo de la secuencia (debida en parte a la fluoresceína inyectada), posibilidad de que aparezcan grandes transformaciones en un breve período de tiempo debido a la alta velocidad de movimiento del ojo, además de unos requerimientos temporales más exigentes, ya que se trata de grandes volúmenes de datos (del orden de 3500 fotogramas por secuencia).

En este trabajo se han explorado dos métodos de alineamiento de imágenes con el objetivo de poder ajustar las imágenes extraídas de las secuencias de *SLO* de manera que se puedan extraer medidas cuantitativas de las mismas, lo cual es extremadamente

difícil sin que las imágenes estén ajustadas. Por otro lado, se ha resuelto el problema de combinar los resultados de las dos técnicas de forma que se complementan la una a la otra, ya que cada una permite realizar el alineamiento de las imágenes con unas características que al otro método le resultan más problemáticas. Todo esto se ha conseguido llevar a cabo en un tiempo que permite abordar el alineamiento de las secuencias entre dos y tres horas, dependiendo del tamaño de la secuencia.

El primer método que se analizó basa su funcionamiento en la extracción previa de unas características de las imágenes, las crestas y los valles. Dicha extracción se lleva a cabo utilizando un operador basado en geometría diferencial, el *MLSEC-ST*, que permite obtener las curvas cresta y valle de una forma robusta y fiable. Hay que destacar que no se trata de un método de segmentación ni de extracción de bordes, ya que las estructuras extraídas no se corresponden con los límites de los objetos, sino que se pueden interpretar como un conjunto de marcadores anatómicos que permiten definir la forma y orientación de las estructuras más relevantes presentes en las imágenes, en el caso de las imágenes de *SLO* la estructura arterio-venosa del fondo del ojo. Sin embargo, el alineamiento de las crestas por sí sólo no proporciona buenos resultados, además de exigir un nivel de cálculo bastante elevado, por lo que se combinó con una estructura en pirámide multiresolución para llevar a cabo el alineamiento. En dicha pirámide, partiendo del nivel superior en el que las imágenes con menor resolución son ajustadas utilizando una búsqueda exhaustiva en el dominio de la Transformada de Fourier, los resultados obtenidos en cada nivel sirven como semilla en el nivel inmediatamente inferior. Así, reduciendo el número de semillas que se pasan de un nivel a otro se consigue un ajuste mucho más robusto y eficiente, hasta llegar al nivel 0 en el que se alinean las imágenes de crestas originales para obtener el resultado definitivo. Como función de medida de la similitud entre las imágenes transformadas se seleccionó el valor de la correlación entre ellas, función que se optimizó empleando el método Downhill Simplex, que demostró ser el más eficaz para maximizar la correlación.

La precisión y fiabilidad de los resultados obtenidos con esta técnica, una vez ajustados los parámetros que configuran su comportamiento y que aparecen resumidos en la tabla 6.1, han permitido que el posterior alineamiento de las imágenes utilizando como referencia las imágenes de crestas y valles tuviera un gran porcentaje de acier-

tos, y unos valores de error en el alineamiento muy bajos. Estos valores, obtenidos experimentalmente, demostraron ser los que mejores resultados ofrecen, tanto a nivel de tiempo de cálculo como de precisión en el alineado. Las distintas columnas representan lo siguiente:

Escala de diferenciación (σ_D): Representa el tamaño de la Gaussiana con la que se convolucionan las imágenes, y se emplea en la estimación de las derivadas de la ecuación 2.1.

Escala de integración (σ_I): Representa el tamaño de la ventana de integración de la ecuación 2.10.

Umbral: Factor que determina el valor que ha de superar una cresta o valle para ser considerada como tal.

Factor c : Factor que se utiliza en la determinación de la similaridad de los autovectores del tensor estructural mediante la ecuación 2.9.

Número de niveles: Número de niveles de la pirámide multiresolución.

Semillas por nivel: Número de semillas que se pasan de un nivel al inferior. Los números se asocian de mayor a menor nivel de resolución.

Tamaño mínimo de la cresta: Las imágenes de crestas definidas aquí permitían un buen ajuste de las imágenes de *SLO*. Sin embargo algunas veces, y debido al gran número de características extraídas, el proceso no daba buenos resultados. Para eliminar este problema, se añade un paso más al método, consistente en el filtrado de las crestas y valles extraídos, de forma que sólo se mantienen los más relevantes, que son aquellos que superan un umbral dado. Este valor podría ser considerado como una elección arbitraria, ya que no se han probado todos los valores posibles en los diferentes experimentos.

Estos valores proporcionaron buenos resultados para todas las secuencias probadas. Una modificación razonable podría ser que los parámetros *Tolerancia* (valor que determina la convergencia del método Downhill Simplex, Algoritmo 2) y *Tamaño mínimo de la cresta* se ajustasen a la calidad de las imágenes: con imágenes de contraste alto se podrían ajustar estos parámetros para obtener más precisión.

Parámetro	Valor
Escala de diferenciación (σ_d)	3.5 píxeles
Escala de integración (σ_i)	5 píxeles
Umbral	0.25
Factor c	0.1
Número de niveles en la pirámide	5
Semillas por nivel	{1, 15, 15, 15, 20}
Longitud mínima de la cresta	125 píxeles

Tabla 6.1: Parámetros del método de alineado basado en el cálculo de líneas cresta y líneas valle.

Sin embargo los cálculos adicionales que serían necesarios añadirían mucho tiempo de procesamiento frente al reducido número de imágenes en las que se mejoraría el resultado del alineado, por lo que finalmente se pospuso este estudio para futuras mejoras del método.

Queremos destacar además que los tiempos de cómputo exigidos por esta técnica de alineamiento permiten que el procesamiento de una secuencia de *SLO* sea una tarea abordable en un tiempo razonable (2-3 horas, en lugar de días, como era necesario utilizando el método tradicional de alineamiento empleado por los oftalmólogos hasta la fecha). Sin embargo, se encontraron zonas de las secuencias en las que el comportamiento de este método no resultaba satisfactorio. En dichas zonas las imágenes presentaban un bajo contraste que no permitiría la extracción de crestas con la suficiente calidad como para que el posterior alineamiento tuviera las referencias necesarias.

Para abordar el problema del alineamiento en los fotogramas con contraste bajo y pocas estructuras vasculares se empleó un método basado en el cálculo de la *información mutua* entre las dos imágenes a ajustar. También se analizaron otros métodos de estimación de la diferencia entre imágenes basados en los valores de los píxeles, como diversas medidas que se extraen a partir de la matriz de coocurrencia de los niveles de gris, pero la *información mutua* fue la que mejores valores obtuvo pues proporciona una medida de la similitud entre ellas sin la necesidad de un paso previo de extracción de características. Esto es así porque se calcula empleando solamente los

histogramas de niveles de gris de las imágenes. También se evaluó como técnica para obtener las probabilidades con las que se calcula la *información mutua* un método no paramétrico de estimación, el denominado método de la ventana de Parzen, pero los resultados obtenidos con ella la descartaron como una buena opción para el problema que se trata de solucionar. Con este método se consiguieron buenos resultados en el alineamiento, aunque considerablemente más lentos que empleando el método anterior, por lo que sólo se empleó para alinear aquellos fotogramas que están en la zona de la secuencia en la que el nivel de fluoresceína en las estructuras vasculares es bajo. En este caso, debido a la cantidad de máximos locales que aparecen en el espacio de búsqueda, el Downhill Simplex no presenta un buen comportamiento como método de optimización, así que se utiliza el *Temple Simulado*, un método estocástico de optimización que gracias al empleo de una función de generación de números aleatorios que influye en la convergencia del método permite escapar de los máximos locales en los que se quedan atrapado los métodos determinísticos como el Downhill Simplex.

Uno de los mayores problemas que surgieron en el desarrollo de la metodología de alineamiento de toda la secuencia fue el ajuste de los alineamientos obtenidos al utilizar dos fotogramas de referencia distintos, uno para la zona con bajo contraste y otro para la zona con un contraste alto, necesarios debido a la alta variabilidad entre las imágenes de cada zona, que imposibilita el alineamiento de todos los fotogramas contra uno solo de referencia. Para resolverlo se diseñó un post-procesado de los fotogramas una vez alineados, que permite poner sobre un mismo eje de coordenadas los fotogramas que antes estaban sobre dos, cada uno de ellos determinado por el fotograma de referencia utilizado en el ajuste.

Finalmente, utilizando la metodología de alineamiento que combina los dos métodos de ajuste se consiguió alinear todos los fotogramas de las secuencias de *SLO*. Esto permite extraer el valor de un píxel a lo largo de toda la secuencia, lo que sirve para construir la curva de dilución, que es una herramienta muy útil que permite a los oftalmólogos estimar una serie de medidas hemodinámicas que permiten estudiar la evolución de un paciente ante el tratamiento al que es sometido. Una de las medidas más importantes que se pueden calcular es el tiempo arteria-vena (TAV), o tiempo que tarda la sangre en llegar desde un punto situado en una arteria de la retina a un punto en una vena de la misma, ya que con él se puede estimar la velocidad de

la sangre en el ojo, el volumen del flujo sanguíneo, y otros parámetros significativos. Tradicionalmente, para calcular estos parámetros eran necesarias muchas horas de trabajo impreciso, ya que el alineamiento de las imágenes de la secuencia se realizaba manualmente. Este procedimiento manual conduce además a resultados difícilmente reproducibles y propensos a errores. Gracias al método aquí presentado se solucionan estos problemas, y las curvas y tiempos estimados son más precisos y con ello las evaluaciones de los pacientes permiten un grado de fiabilidad mucho mayor.

En la actualidad se está tratando de obtener y validar otros parámetros hemodinámicos, como son el flujo de sangre retinal, que se calcula utilizando el tiempo arteria vena, del que ya disponemos, y el diámetro de las arterias y venas. En concreto, el flujo de sangre retinal es proporcional a la suma de los cuadrados de los diámetros de arterias y venas de la retina dividido por el tiempo arteria-vena. Este parámetro es de una importancia fundamental en el estudio de la fisiología retinal, así como en el estudio de cambios patológicos en la circulación retinal causados por enfermedades vasculares como la retinopatía diabética. Sin embargo, este cálculo implica la determinación de los calibres de las estructuras vasculares, por lo que también se está trabajando en la obtención de una técnica de segmentación de la estructura arterio-venosa de la retina que permita calcular con precisión este calibre.

Por otro lado, gracias a la experiencia adquirida tanto en el tratamiento de las imágenes de fondo de retina como en la extracción de características de las mismas, se ha abierto una nueva línea de trabajo, en la que ya se han cosechado los primeros resultados [Mariño y col., 2003, Mariño y col., 2005, Mariño y col., 2006, Ortega y col., 2006]. Los sistemas fiables de autenticación de personas han sido una importante área de trabajo durante mucho tiempo, que se vuelve más importante conforme la necesidad de un nivel de seguridad alto crece. Así el acceso a una infraestructura fiable de identificación de personas se convierte en una herramienta fundamental en muchas situaciones (seguridad en aeropuertos, control de acceso a recursos protegidos, etc.). Los sistemas convencionales de identificación basados en la posesión de tarjetas identificativas o conocimiento exclusivo no son fiables, ya que las primeras pueden perderse, ser robadas o extraviadas, mientras que las claves de acceso pueden olvidarse o comprometerse debido a descuidos. Una solución a todos estos problemas la proporciona las tecnologías de autenticación basadas en parámetros biométricos.

Un sistema biométrico es un sistema de reconocimiento de patrones que determina la autenticidad de una característica física o del comportamiento de una persona. La identificación puede ser en forma de autenticación (comprobación de la veracidad de la identidad de una persona, alguien es quien dice ser) o en forma de identificación (determinación de la identidad de una persona sin conocer su nombre).

Existen en la literatura muchas técnicas de autenticación biométricas, como las basadas en el patrón del iris, la huella dactilar (quizás la más común de las técnicas de autenticación biométrica), geometría de la mano, reconocimiento de voz, reconocimiento del rostro, etc. Incluso existen técnicas en las que se combina más de uno de estos parámetros para dotar de una mayor robustez y fiabilidad al sistema.

Sin embargo, hoy en día se tiende a desarrollar sistemas con un alto nivel de seguridad, en los que es muy complicado, o idealmente imposible, crear una copia de las características que se emplean en el sistema para discriminar entre un usuario autorizado y uno que no lo es, de forma que un impostor podría ser aceptado por el sistema como un usuario autorizado.

En este sentido, aprovechando el trabajo realizado con imágenes de fondo de ojo, y la experiencia con las técnicas desarrolladas para la extracción de características (crestas y valles), hemos empezado a desarrollar un sistema en el que se emplea el patrón arterio-venoso de la retina como parámetro biométrico, pues este es único para cada persona. Además, la configuración macroscópica de las venas y arterias de la retina se mantiene estable a lo largo de la vida de una persona, y la identifica unívocamente. Por otro lado, debido a su posición interna, la retina está protegida de variaciones causadas por la exposición al entorno (hecho que no se da en otros parámetros como la huella dactilar, rostro o la geometría de la mano). Estos factores, junto con la robustez de los métodos de alineamiento que hemos desarrollado en este trabajo, que permiten realizar el ajuste incluso ante la desaparición de alguna vena o arteria debido al movimiento del ojo durante la captura de la muestra, hace que este método de autenticación sea un buen candidato para entornos de alta seguridad.

Apéndice A

Fundamentos de teoría de la información

En éste apéndice introduciremos los fundamentos de la *teoría de la información* necesarios para comprender el funcionamiento del método de alineamiento de imágenes basado en *información mutua*. La teoría de la información se utiliza en numerosos campos, como son la criptografía, la compresión de datos, ... En esta sección introduciremos una serie de conceptos relacionados con la *información mutua*. Los conceptos se introducirán sólo para *variables aleatorias* discretas. Para un tratamiento más extenso sobre el tema puede consultarse [Cover y Thomas, 1991].

A.1. Entropía

La entropía es una medida de la incertidumbre o aleatoriedad de una *variable aleatoria*. Muchos de los estudios sobre entropía se enmarcan en la transmisión de datos, en los cuales la aportación de Shannon [Shannon, 1948, Shannon y Warren, 1963] con respecto a este tema ha sido fundamental.

Sea X una *variable aleatoria* discreta con un alfabeto \aleph y función de distribución de probabilidad $p(x) = P(X = x)$, $x \in \aleph$. Se define la entropía $H(X)$ de una *variable*

aleatoria discreta como:

$$H(x) = - \sum_{x \in \mathfrak{N}} p(x) \log p(x) \quad (\text{A.1})$$

También puede definirse la entropía como una esperanza matemática:

$$H(x) = -E_x[\log(p(x))] \quad (\text{A.2})$$

Si el logaritmo es en base 2 entonces la entropía se expresa en bits. La entropía de una *variable aleatoria* no depende de los valores que ésta pueda tomar sino de sus probabilidades.

Una propiedad importante de la entropía es que $H(x) \geq 0 \forall x$.

A.2. Entropía condicional, entropía conjunta e información mutua

La definición de entropía se extiende para el caso de dos *variables aleatorias*. La *entropía conjunta* $H(X, Y)$ de dos *variables aleatorias* discretas X e Y con una distribución conjunta de probabilidad $p(x, y)$ se define como:

$$H(X, Y) = - \sum_{x \in \mathfrak{N}} \sum_{y \in \mathfrak{S}} p(x, y) \log p(x, y) \quad (\text{A.3})$$

La entropía condicionada de una variable con respecto a otra se define como:

$$H(Y|X) = \sum_{x \in \mathfrak{N}} p(x) H(Y|X = x) = - \sum_{x \in \mathfrak{N}} p(x) \sum_{y \in \mathfrak{S}} p(y|x) \log p(y|x) \quad (\text{A.4})$$

La entropía condicionada $H(Y|X)$ proporciona la entropía media de la *variable aleatoria* Y dado que se conoce la *variable aleatoria* X .

Una propiedad interesante de la entropía conjunta de dos *variables aleatorias* es la siguiente: $H(X, Y) = H(X) + H(Y|X)$.

A.3. Información mutua

Consideremos dos *variables aleatorias* X e Y con función de distribución conjunta $p(x, y)$ y funciones de densidad $p(x)$ y $p(y)$ respectivamente. La *Información Mutua* $IM(X, Y)$ se define como:

$$IM(X, Y) = \sum_{x \in \mathfrak{X}} \sum_{y \in \mathfrak{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (\text{A.5})$$

La *información mutua* entre dos *variables aleatorias* X e Y representa la cantidad de información que una de las variables proporciona sobre la otra.

La *información mutua* entre dos variables se puede expresar de varias formas realizando una serie de manipulaciones matemáticas:

$$\begin{aligned} IM(X, Y) &= & IM(Y, X) &= \\ &= & \sum_{x,y} p(x, y) \log \frac{p(x,y)}{p(x)p(y)} &= \\ &= & \sum_{x,y} p(x, y) \log \frac{p(x|y)}{p(x)} &= \\ &= & - \sum_{x,y} p(x, y) \log p(x) + \sum_{x,y} p(x, y) \log p(x|y) &= \\ &= & - \sum_x p(x) \log p(x) - \left(\sum_{x,y} p(x, y) \log p(x|y) \right) &= \\ &= & H(X) - H(X|Y) &= \\ &= & H(X) + H(Y) - H(X, Y) &= \end{aligned} \quad (\text{A.6})$$

La relación entre la entropía y la *información mutua* dada por la ecuación A.6 será la que utilizemos para calcular el valor de la *información mutua* en este trabajo. Además se puede demostrar que se cumple la siguiente relación:

$$H(X, Y) \leq H(X) + H(Y) \quad (\text{A.7})$$

De la ecuación A.7 se deduce que la *información mutua* entre dos *variables aleatorias* es siempre mayor o igual que 0. La *información mutua* entre dos *variables aleatorias* X e Y se puede ver como la reducción en la incertidumbre de la *variable aleatoria* X debido a que se conoce la *variable aleatoria* Y . Se cumple la relación:

$$IM(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (\text{A.8})$$

Si X e Y son *variables aleatorias* independientes entonces el valor de $IM(X, Y)$ es 0. La relación entre entropía e *información mutua* se puede ver claramente en el diagrama de Venn de la Figura A.1. En ella aparece representada la idea de la *información mutua* como la entropía conjunta de las *variables aleatorias* con respecto a las entropías marginales.

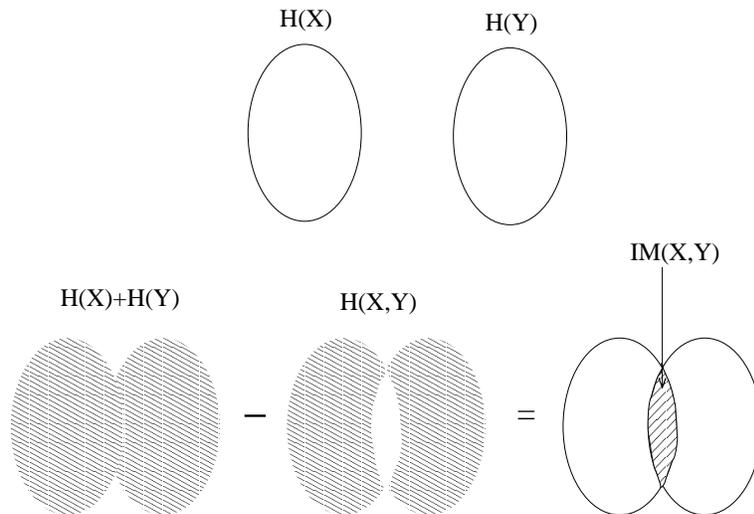


Figura A.1: Esquematación de la relación entre entropía e *información mutua* utilizando diagramas de Venn. En la parte superior se muestran las entropías marginales de cada una de las *variables aleatorias*, mientras que en la parte inferior se representa la ecuación A.6 que indica cómo calcular la *información mutua* a partir de las entropías marginales y de la entropía conjunta de las *variables aleatorias*.

Apéndice B

Modelización de transformaciones afines

Dado que el dominio de aplicación de nuestro trabajo son los fotogramas de secuencias *SLO*, presentaremos aquí sólo la formulación de las transformaciones para puntos en el plano bidimensional. Además veremos la representación en *coordenadas homogéneas* que se utilizó en la implementación de las matrices de transformación, que nos proporcionan una forma compacta de representar los valores de vectores y puntos de una imagen.

B.1. Coordenadas homogéneas

Para representar matemáticamente los puntos y los vectores, recordemos que se utiliza un **marco de coordenadas**, que viene dado por un punto ϑ denominado origen y un trío (aunque nuestras imágenes estén en el espacio bidimensional, realizaremos esta explicación para el caso más general de 3D, siendo la formulación para el caso bidimensional idéntica) de vectores mutuamente excluyentes \mathbf{i} , \mathbf{j} y \mathbf{k} , que son los ejes de coordenadas.

Así, para representar un vector \mathbf{v} se utilizan tres cantidades (v_1, v_2, v_3) de forma que:

$$\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k} \quad (\text{B.1})$$

Por otro lado, para representar un punto P , vemos su posición como un desplazamiento con respecto al origen. Se representa el vector $P - \vartheta$ utilizando tres valores (p_1, p_2, p_3) de forma que:

$$P - \vartheta = p_1 \mathbf{i} + p_2 \mathbf{j} + p_3 \mathbf{k} = \vartheta + p_1 \mathbf{i} + p_2 \mathbf{j} + p_3 \mathbf{k} \quad (\text{B.2})$$

La representación de P no es un trío, sino un trío con un origen. La idea básica es hacer el origen de coordenadas *explícito*. Resulta muy útil poder representar puntos y vectores utilizando el mismo conjunto de objetos de la forma (a, b, c, ϑ) . De las ecuaciones B.1 y B.2, el vector $\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}$ necesita los cuatro coeficientes $(v_1, v_2, v_3, 0)$, mientras que el punto $P = \vartheta + p_1 \mathbf{i} + p_2 \mathbf{j} + p_3 \mathbf{k}$ necesita los cuatro coeficientes $(p_1, p_2, p_3, 1)$. El cuarto componente designa si el objeto necesita o no incluir ϑ . Se pueden escribir formalmente \mathbf{v} y P utilizando notación matricial como:

$$\mathbf{v} = (\mathbf{i}, \mathbf{j}, \mathbf{k}, \vartheta) \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ 0 \end{pmatrix} \quad (\text{B.3})$$

y

$$P = (\mathbf{i}, \mathbf{j}, \mathbf{k}, \vartheta) \begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ 1 \end{pmatrix} \quad (\text{B.4})$$

Utilizando esta notación, el vector fila indica la naturaleza del marco de coordenadas, y el vector columna indica la representación del objeto específico de interés. Así,

los vectores y los puntos tienen diferentes representaciones: los vectores tienen un 0 como cuarta componente mientras que los puntos tienen un 1 como cuarta componente. Las ecuaciones B.3 y B.4 son ejemplos de la **representación en coordenadas homogéneas** de vectores y puntos.

Además, es muy sencillo pasar de un sistema de coordenadas ordinarias a un sistema de coordenadas homogéneas:

Para pasar de coordenadas ordinarias a homogéneas

Si el objeto es un punto, añadir un 1.

Si el objeto es un vector, añadir un 0.

Para pasar de coordenadas homogéneas a ordinarias

Si el objeto es un punto, su última coordenada es un 1. Elimínese el 1.

Si el objeto es un vector, su última coordenada es un 0. Elimínese el 0.

Para más detalles sobre las coordenadas homogéneas, consultar [Faugeras, 1993, Foley y col., 1990]

B.2. Transformaciones rígidas

En general, una transformación T mapea un punto arbitrario en el plano P en otro punto Q .

La transformación de un objeto se lleva a cabo transformando cada uno de sus puntos mediante la función de mapeo T .

En el caso 3D tenemos que los puntos P y Q tienen por representación:

$$P = \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad (\text{B.5})$$

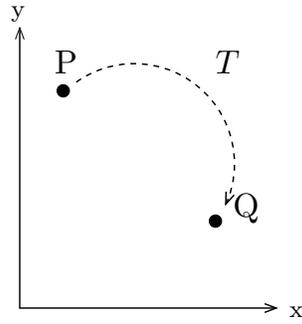


Figura B.1: Representación de la transformación en 2D del punto P en el punto Q mediante la función T .

y

$$Q = \begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} \quad (\text{B.6})$$

respectivamente. Algebraicamente, esto significa que P está en la posición $P = P_x \mathbf{i} + P_y \mathbf{j} + P_z \mathbf{k} + \vartheta$, e igual para Q . Una transformación opera sobre la representación P y produce la representación Q según alguna función de mapeo T . Esto es:

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = T \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad (\text{B.7})$$

o de forma más compacta:

$$Q = T(P) \quad (\text{B.8})$$

En nuestro caso, como ya se ha comentado, modelamos estas transformaciones utilizando *transformaciones afines*. Una transformación afín puede representarse cómo-

damente utilizando notación matricial. Las coordenadas de Q son una combinación lineal de las de P , de la forma:

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11}P_x + m_{12}P_y + m_{13}P_z + m_{14} \\ m_{21}P_x + m_{22}P_y + m_{23}P_z + m_{24} \\ m_{31}P_x + m_{32}P_y + m_{33}P_z + m_{34} \\ 1 \end{pmatrix} \quad (\text{B.9})$$

para el conjunto de constantes $m_{11}, m_{12}, \dots, m_{34}$. O representado de otra forma:

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad (\text{B.10})$$

Es importante notar que para las transformaciones afines, la tercera fila de la matriz de transformación es siempre $(0, 0, 0, 1)$.

B.2.1. Componentes de una transformación afín

Las transformaciones afines producen combinaciones de tres elementos fundamentales: traslación, escalado y rotación. A continuación describiremos la representación matricial y operaciones sobre ellas.

B.2.1.1. Traslación

El desplazamiento de un punto a una posición diferente viene dado por la componente de traslación de la transformación. La parte de traslación de una transformación afín viene dada por la cuarta columna de la matriz:

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & m_{14} \\ 0 & 1 & 0 & m_{24} \\ 0 & 0 & 1 & m_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} = \begin{pmatrix} P_x + m_{14} \\ P_y + m_{24} \\ P_z + m_{34} \\ 1 \end{pmatrix} \quad (\text{B.11})$$

Así, en coordenadas ordinarias, $Q = P + \mathbf{d}$, donde $\mathbf{d} = (m_{14}, m_{24}, m_{34})$ es el *vector desplazamiento*.

B.2.1.2. Escalado

La componente de escalado de una transformación afín varía el tamaño de un objeto, e implica dos factores de escala: S_x, S_y y S_z para el escalado en los ejes x, y y z respectivamente:

$$(Q_x, Q_y, Q_z) = (S_x P_x, S_y P_y, S_z P_z)$$

Así la matriz de transformación para un escalado es:

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = \begin{pmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad (\text{B.12})$$

De forma más precisa, esta matriz representa el *escalado con respecto al origen*.

B.2.1.3. Rotación

La tercera componente de una transformación afín que vamos a ver es la rotación, dados un punto y un ángulo.

Cuando $T()$ es una rotación sobre el origen, el vector de desplazamiento \mathbf{d} es cero, y $Q = T(P)$, la matriz de transformación tiene la siguiente forma, según la rotación sea sobre el eje x , el y o el z respectivamente:

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad (\text{B.13})$$

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad (\text{B.14})$$

$$\begin{pmatrix} Q_x \\ Q_y \\ Q_z \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 & 0 \\ \sin(\theta) & \cos(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} P_x \\ P_y \\ P_z \\ 1 \end{pmatrix} \quad (\text{B.15})$$

Para ver la rotación sobre un punto arbitrario distinto del centro, es necesario estudiar la *composición de transformaciones afines*, ya que dicha rotación será la composición de la transformación del punto hasta el origen de coordenadas seguido de la rotación tal y como la acabamos de ver.

B.2.1.4. Composición de transformaciones afines

Normalmente una transformación afín no va a estar formada por uno sólo de los componentes que acabamos de ver, sino que vendrá dada por la aplicación de varios de ellos. El proceso de aplicar varias transformaciones sucesivas para formar una transformación afín conjunta se denomina *composición* (o *concatenación*) de transformaciones. Se puede demostrar matemáticamente que el resultado de componer transformaciones afines es también una transformación afín.

Sean las transformaciones en 2D $T_1()$ y $T_2()$, de forma que, como muestra la Figura B.2, la transformación T dada por la composición de $T_1()$ y $T_2()$ lleva directamente el punto P a R sin pasar por Q , es decir, $R = T_2(Q) = T_2(T_1(P))$.

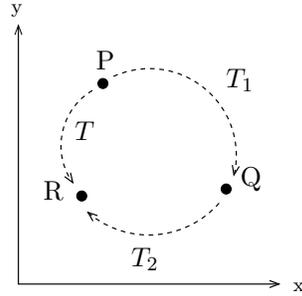


Figura B.2: Representación de la composición de dos transformaciones.

Suponiendo que las dos transformaciones vienen dadas por las matrices M_1 y M_2 , P se transforma primero en el punto M_1P , que después es transformado en $M_2(M_1P)$. Por asociatividad, esto último es igual a $(M_2M_1)P$, y así:

$$R = MP = M_2M_1P \quad (\text{B.16})$$

Es decir, utilizando coordenadas homogéneas, las transformaciones afines se componen por la simple multiplicación de las matrices de las transformaciones componentes, en orden inverso, eso sí, al orden en el que se aplican las transformaciones. Este razonamiento puede hacerse extensible a un número indefinido de transformaciones.

Apéndice C

Algoritmos de optimización

En cualquier técnica de alineamiento de imágenes es necesario obtener el óptimo de la función que mide el grado de solapamiento de las imágenes que se está procesando. En este trabajo se han empleado dos métodos de optimización: en la técnicas de alineamiento de imágenes basadas en la extracción de características se empleó el método de minimización multidimensional denominado *Downhill Simplex*, desarrollado por Nelder y Mead [Nelder y Mead, 1965]. Por otro lado, para la maximización de la Información Mutua se optó por la utilización de método denominado *Temple Simulado*, que permite solventar el problema de los máximos locales que aparecen al utilizar la información mutua como medida del ajuste, ante los cuales el método Downhill Simplex presenta problemas.

C.1. Downhill Simplex

El *simplex* es una figura geométrica con $N + 1$ vértices, donde N es el número de dimensiones. En el caso bidimensional el simplex es un triángulo, en el caso tridimensional un tetraedro, etc.

Cada vértice del triángulo (vamos a centrarnos en nuestro caso, espacio bidimensional) tiene asignado el valor, devuelto por la función a maximizar (la correlación),

correspondiente a sus coordenadas. De esta forma vamos a tener ordenados los tres vértices según el resultado obtenido.

El método sigue una serie de pasos, la mayoría de los cuales consisten en mover el punto del simplex en el que la función a optimizar tiene el valor máximo. A continuación pasamos a describir los movimientos definidos para el triángulo, que son: **reflexión, expansión, contracción en un vértice y contracción total.**

La idea de la reflexión es cambiar el sentido e incluso la dirección de la búsqueda. Para ello realizamos una especie de *reflexión compuesta*, consistente en una reflexión horizontal y a continuación una vertical (ver figura C.1). De esta manera intentamos encaminar el nuevo vértice hacia la zona en donde puede haber mejores valores, debido a la proximidad con el vértice que contiene el mejor valor hasta ese momento, y en sentido opuesto al vértice que contiene el peor resultado.

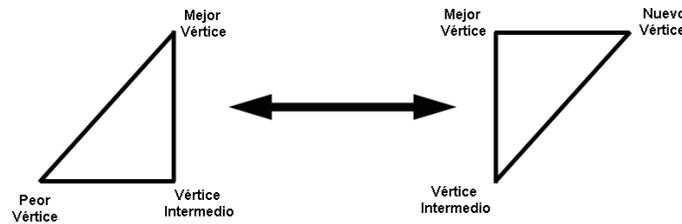


Figura C.1: Operación de reflexión del simplex en 2D.

Cuando encontramos un nuevo punto cuyo resultado supera a todos los vértices anteriores, tratamos de prolongar el cambio en el sentido en el cual se ha obtenido ese nuevo vértice ejecutando una expansión (ver figura C.2). Este cambio es el único que permite que el triángulo crezca, cuestión necesaria cuando nuestro triángulo es demasiado pequeño y estamos lejos de la solución. En este algoritmo la expansión sólo puede aplicarse después de una reflexión. Si la reflexión obtiene un buen resultado se expande el triángulo en ese sentido.

Cuando intentamos reflejar el triángulo y expandirlo no encontramos vértices prometedores, puede ocurrir que la solución buscada esté en el interior del área descrita por dicho triángulo. Este es el momento de contraer uno de los vértices, el que peor resultado esté proporcionando hasta ese momento, y observar los nuevos resultados

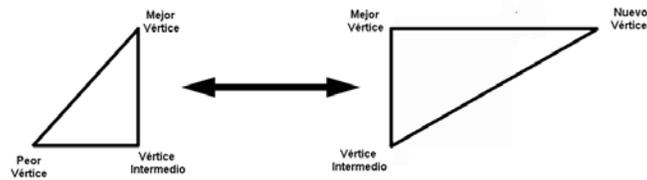


Figura C.2: Operaciones de reflexión y expansión del simplex en 2D.

(ver figura C.3). Este cambio, junto con la contracción total son los que permiten al triángulo hacerse más pequeño, cuestión necesaria cuando la solución está en su interior.

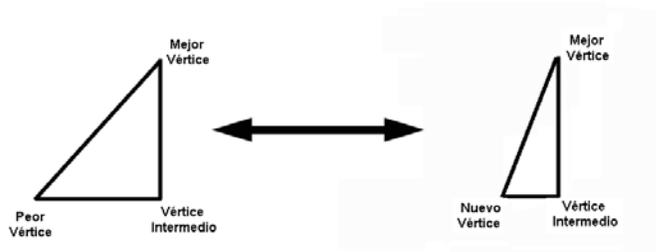


Figura C.3: Operación de contracción en un vértice del simplex en 2D.

Si probados todos los movimientos anteriores y no encontramos el camino de búsqueda, podemos recurrir a contraer completamente el triángulo (ver figura C.4).

Todos estos movimientos son aceptados cuando el *nuevo vértice*, una vez evaluado con la función a maximizar, obtiene un valor mejor que el valor del *peor vértice*.

La inicialización del triángulo es un factor determinante que puede decidir el funcionamiento posterior del método de búsqueda. Los vértices del primer triángulo generado en la búsqueda responden a la siguiente fórmula:

$$P_i = P_0 + \lambda e_i \quad (\text{C.1})$$

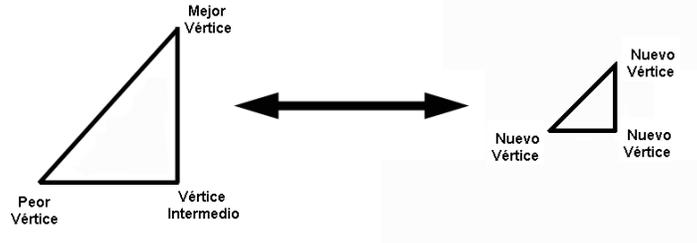


Figura C.4: Contracción total del simplex en 2D.

En la ecuación C.1, P_i simboliza cada uno de los vértices del triángulo. P_0 referencia el punto de partida de la inicialización, a la cual se le suma una cantidad λ en el sentido marcado por el vector unitario e_i . Esta fórmula puede generalizarse permitiendo el uso de un λ diferente para cada vértice. En nuestro caso el triángulo de partida siempre será equilátero, factor que potencia la búsqueda del máximo en cualquier dirección, dado que a priori no va a existir en nuestros espacios de búsqueda una dirección preferente que nos dé la solución.

El pseudocódigo mostrado en el Algoritmo 2 representa la implementación de este método de búsqueda del máximo de los valores de la correlación.

C.2. Temple Simulado

El *Temple Simulado* tiene su origen en una analogía con la termodinámica, concretamente con el proceso de congelación y cristalización de líquidos, o el enfriamiento y templado de metales. A altas temperaturas, las moléculas de un líquido se mueven libremente unas con respecto a otras. Si el líquido se enfría lentamente entonces la movilidad de las partículas debida a la temperatura se pierde progresivamente y las partículas se alinean formando cristales ordenados. El cristal es el estado de mínima energía del sistema. Cuando las partículas se enfrían lentamente, el sistema es capaz de evolucionar al estado de mínima energía. Por otra parte, si un metal líquido se enfría rápidamente no alcanza su estado de equilibrio, sino que evoluciona a un estado policristalino o amorfo con una energía mayor que la del estado cristalino. La principal característica del proceso es el enfriamiento progresivo, dejando el tiempo

Algoritmo 2 Pseudocódigo para el algoritmo Downhill Simplex

```
1:  $ntemps \leftarrow$  número máximo de intentos a realizar
2:  $ftol \leftarrow$  umbral de tolerancia para la convergencia del método
3:  $ctol \leftarrow$  valor actual de la tolerancia
4:  $terminar \leftarrow$  false
5: inicialización del triángulo
6: repeat
7:   ordenar los vértices según el valor devuelto por la correlación
8:   calcular  $ctol$ 
9:   if  $ctol < ftol$  then
10:      $terminar = true$ 
11:   else
12:     aplicar reflexión sobre el vértice con mayor valor de correlación
13:     if el nuevo vértice supera al mejor vértice then
14:       almacenar el cambio
15:       aplicar expansión
16:       if el nuevo vértice supera al mejor vértice then
17:         almacenar el cambio
18:       end if
19:     else if es peor resultado que el vértice intermedio then
20:       aplicar contracción en un vértice
21:       almacenar el cambio
22:       if el nuevo vértice obtiene un peor resultado que el peor vértice then
23:         aplicar contracción total
24:         almacenar el cambio
25:       end if
26:     end if
27:   end if
28:   incrementar intentos
29:   if  $intentos > ntemps$  then
30:      $terminar = true$ 
31:   end if
32: until  $terminar = true$ 
```

suficiente al sistema para que evolucione al estado de mínima energía. La distribución de Boltzmann:

$$P(E) \propto \exp\left(\frac{-E}{kT}\right) \quad (\text{C.2})$$

expresa la idea de que la energía de un sistema en equilibrio térmico a una temperatura T está distribuida en diferentes estados de energía. Así, incluso a bajas temperaturas hay una pequeña probabilidad de que el sistema se encuentre en un estado de alta energía. Esto permite que el sistema pueda evitar caer en mínimos locales encontrando el mínimo global. La constante k (constante de Boltzmann) relaciona la energía con la temperatura. El sistema evoluciona hacia estados de menor energía aunque algunas veces aumenta su energía para evitar mínimos locales.

El uso de este algoritmo en la minimización de una expresión requiere el uso de tres funciones:

1. Una función generadora de números aleatorios.
2. Una función de aceptación.
3. Una función de enfriamiento.

Explicaremos ahora brevemente el algoritmo de maximización (que se puede entender como una minimización cambiando de signo los valores de la función) para el caso que nos ocupa.

Sea $IM(x, y)$ la función bidimensional a minimizar (recordemos que tan solo se consideran los parámetros de traslación en x e y , no la rotación). Cada parámetro de la función se minimiza independientemente del otro. Llamemos (x_c, y_c) al valor que tenemos como principal candidato a máximo global con la configuración actual del sistema. El siguiente paso a realizar consiste en generar un nuevo valor de prueba (x_n, y_n) . Este valor se obtiene usando la función generadora de números aleatorios, que nos da el salto que se realiza desde la posición actual, (x_c, y_c) , hasta la posición de prueba, (x_n, y_n) . La función de generación de números aleatorios nos da valores siguiendo una distribución con forma de campana, de tal forma que generan saltos

más pequeños con una mayor probabilidad. A medida que la temperatura se reduce, usando la función de enfriamiento, la anchura de la función generadora de números aleatorios también se reduce. Si $IM(x_n, y_n) < IM(x_c, y_c)$ entonces (x_n, y_n) reemplaza a (x_c, y_c) como máximo provisional, $x_c = x_n, y_c = y_n$. Por el contrario, si $IM(x_n, y_n) > IM(x_c, y_c)$, para escapar de mínimos locales habrá ocasiones en las que se acepte el valor de (x_n, y_n) como máximo provisional. La aceptación o rechazo de un valor de la *información mutua* menor que el mínimo provisional viene dado por la función de aceptación.

Hay varias opciones para elegir la función generadora de números aleatorios, la función de enfriamiento y la función de aceptación. La elección de las dos primeras está relacionada, para asegurar una convergencia al máximo global. Como función generadora de números aleatorios se ha elegido la función de *Cauchy*:

$$G_k(d) = \frac{t_k}{\pi(d^2 + t_k^2)} \quad (\text{C.3})$$

donde t_k es la temperatura actual del sistema y d es un parámetro generado aleatoriamente. Esta función se puede ver en la Figura C.5.

La función de enfriamiento que asegura la convergencia del algoritmo al mínimo es la siguiente:

$$t_k = \frac{t_0}{k} \quad (\text{C.4})$$

donde t_0 es la temperatura inicial del sistema que se elige empíricamente y t_k es la temperatura en la iteración k . La combinación de estas dos funciones da lugar a lo que se conoce como *Temple Simulado Rápido*.

La tercera función que se usa en este algoritmo es la función de aceptación, que se define para escapar de mínimos locales. Para ello hay ocasiones en las que se acepta un salto a una posición con una energía menor. Así, la función de aceptación es la que determina en qué ocasiones se debe aceptar un salto con un valor mayor de energía. La expresión que se ha usado en la implementación de este algoritmo es la siguiente:

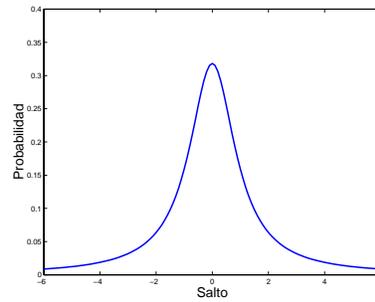


Figura C.5: Forma de la distribución de Cauchy utilizada como función generadora de números aleatorios.

$$A_k = \begin{cases} 1 & A_k \geq 1 \\ \frac{e^{IM(x_n, y_n) - IM(x_c, y_c)}}{C t_k} & A_k < 1 \end{cases} \quad (\text{C.5})$$

donde t_k es la temperatura actual del sistema y C es una constante que se determina empíricamente. El proceso de aceptación del salto a un estado de mayor energía es el siguiente: se genera un número aleatorio comprendido en el intervalo $[0, 1)$ y se compara este número con la probabilidad de aceptación que proporciona la función de aceptación. Si la probabilidad de aceptación es mayor que el número aleatorio entonces se acepta el salto a un estado de mayor energía, y en caso contrario se rechaza este salto. Por otro lado, para evitar que el algoritmo se escape poco a poco de la zona del mínimo global debido a pequeños aumentos de energía, al cabo de cada iteración se hace una comparación del valor de energía más bajo obtenido y se compara con el valor de energía del mínimo provisional; si este último valor es mucho mayor que el valor mínimo obtenido entonces el valor mínimo de energía pasa a ser el mínimo provisional. El pseudocódigo 3 muestra el algoritmo descrito.

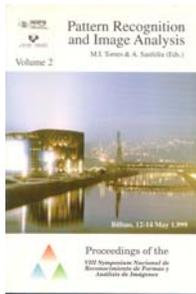
Dentro de este algoritmo hay que elegir un criterio de parada. Éste debe ser tal que la temperatura del sistema sea lo suficientemente baja como para que el sistema se encuentre en el estado de equilibrio. En este caso, como criterio de parada tomamos que el sistema no evolucione a un estado de energía menor durante 3 iteraciones y, además, que los desplazamientos que se produzcan no varíen mucho durante esas 3 iteraciones. Este último criterio se introduce para evitar que el sistema oscile entre puntos de energía similar, pero que producen alineamientos muy diferentes.

Algoritmo 3 Pseudocódigo para el algoritmo *Temple Simulado*.

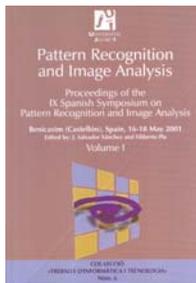
```
1:  $\Delta$  factor de reducción de la temperatura
2:  $n_{temps} \leftarrow N$ , número máximo de intentos a realizar
3:  $n_{limit} \leftarrow N_t$ , número de intentos con cada temperatura
4:  $g_{limit} \leftarrow S_e$ , número de intentos con éxito (o intercambios)
5:  $t_0 \leftarrow$ , temperatura inicial
6:  $t_k \leftarrow t_0$ , temperatura en el instante  $k$ .
7: for  $k=1, n_{temps}$  do
8:    $t_k = t_k \Delta$ 
9:   for  $j=1, n_{limit}$  do
10:    intentar intercambiar un par de puntos aleatorios
11:     $\delta = \text{coste\_actual} - \text{coste\_intento}$ 
12:    if  $\delta > 0$  then
13:      hacer el cambio permanente
14:       $\text{numero\_intercambios} = \text{numero\_intercambios} + 1$ 
15:    else
16:       $p =$  número aleatorio  $\in [0, 1)$ 
17:       $m = \exp(\frac{\delta}{t_k})$ 
18:      if  $p < m$  then
19:        hacer el cambio permanente
20:         $\text{numero\_intercambios} = \text{numero\_intercambios} + 1$ 
21:      end if
22:    end if
23:    if  $\text{numero\_intercambios} > g_{limit}$  then
24:      terminar
25:    end if
26:  end for
27: end for
```

Apéndice D

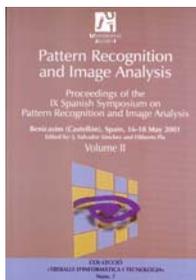
Publicaciones seleccionadas



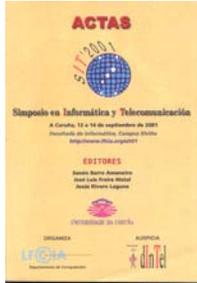
C.Mariño, M.G. Penedo, and M.J. Carreira. Modified Contour Fitting applied to Detection of Retinal Vessels. *Proceedings of the 5th Iberoamerican Symposium of Pattern Recognition (SIARP'2000)*, pp. 559-569, 2000



D.Lloret, C.Mariño, J.Serrat, A.M. López and J.J.Villanueva. Landmark-based Registration of Full SLO Video Sequences. *Pattern Recognition and Image Analysis (Proceedings of the IX Spanish on Pattern Recognition and Image Analysis, Vol I)*, pp. 189-194, 2001



F. Pardo, V. Leborán, C.Mariño, M.G. Penedo, M.J. Carreira, A. Mosquera, D. Cabello F. Gómez-Ulla and F. González. Retinal Angiography Image Registration applied to Hemodynamic Variable Measurement. *Pattern Recognition and Image Analysis (Proceedings IX Spanish on Pattern Recognition and Image Analysis, Vol II)*, pp. 139-144, 2001

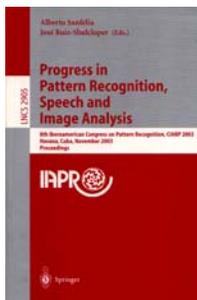


C.Mariño, V. Gulías, M. Penas, M.G. Penedo, V. Leborán, M.J. Carreira, A. Mosquera and D. Lloret. Sistema de interpretación automática de secuencias SLO basado en un servidor VoD. *Simposio en Informática y Telecomunicación (SIT'2001)*, pp. 319-329, 2001

C.Mariño, M. Penas, M.G. Penedo, M.G. Lloret and M.J. Carreira. Integration of Mutual Information and Creaseness Methods for the Automatic Registration of SLO sequences. *6th Iberoamerican Symposium on Pattern Recognition (SIARP'2001)*, pp. 77-83, 2001



C.Mariño, M.G. Penedo, J.M.Barja, V.Leborán, M.J.Carreira and F.Gómez-Ulla. Methodology for the Registration of whole SLO Sequences. *Proceedings 16th International Conference on Pattern Recognition (ICPR'2002)*, Vol. 1, pp. 79-783, 2002



C.Mariño, M.G. Penedo, M.J.Carreira and F.González. Retinal Angiography Based Authentication *Progress in Pattern Recognition, Speech and Image Analysis (Proceedings 8th Iberoamerican Congress on Pattern Recognition, CIARP'2003. Lecture Notes in Computer Science*, Vol. 2905, 306-313, 2003



I. Caderno, M. Penedo, N. Barreira, C. Mariño and F. González. Precise Detection and Measurement of the Retina Vascular Tree *Proceedings 7th International Conference on Pattern Recognition and Image Analysis: New Information Technologies*, Vol. II, pp. 639-642, 2004



C. Mariño, M. G. Penedo and M. Penas. Retinal Based Authentication via Distributed Web Application *Computer Aided Systems Theory (EUROCAST'05)*, *Lecture Notes in Computer Science*, Vol. 3643, pp. 286-391, 2005



M. Penas, M. J. Carreira, M.G. Penedo and C. Mariño. Segment extraction using Burns Principles in a pseudo-color fuzzy Hough transform. *Pattern Recognition and Image Analysis (Proceedings 2nd Iberian Conference, IBPRIA '2005)*. *Lecture Notes in Computer Science*, LNCS 3523, pp. 175-182, 2005

C. Mariño., M. Penas, M.G. Penedo and M. J. Carreira. An Algorithm for Registration of Full Scanning Laser Ophtalmoscope Video Sequences. *Image Vision and Computing*, pendiente de aceptación.



C. Mariño, M. G. Penedo, M. Penas, M. J. Carreira and F. González. Personal authentication using retinal angiographies *Pattern Analysis and Applications*. Vol. 9 (en prensa). 2006.

Bibliografía

- [Aitken, 2003] Aitken, V. (2003). Measurement of retinal microvascular pulsation from SLO image sequences. En *Canadian Conference on Electrical and Computer Engineering*, volumen 3, páginas 1473–1476.
- [Ayer-Siegel y col., 2002] Ayer-Siegel, R., Bourla, D., Yassur, Y., y Weinberger, D. (2002). Angiographic and flow patterns of retinal choroidal anastomoses in age-related macular degeneration with occult choroidal neovascularization. *Ophthalmology*, 109(9):1726–1736.
- [Becker y col., 1998] Becker, D. E., Can, A., Turner, J. N., Tanenbaum, H. L., y Roysan, B. (1998). Image processing algorithms for retinal montage synthesis, mapping and real time location determination. *IEEE Transactions on Biomedical Engineering*, 45(1):346–360.
- [Belongie y col., 2001] Belongie, S., Malik, J., y Puzicha, J. (2001). Matching shapes. En *Proceedings of International Conference on Computer Vision (ICCV 2001)*, páginas 454–463.
- [Bigun y col., 1991] Bigun, J., Granlund, G., y Wiklund, J. (1991). Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:775–790.
- [Bonafonte y García, 1996] Bonafonte, S. y García, C. A. (1996). *Epidemiología de la retinopatía diabética*. Harcourt Brace, Madrid.
- [Bro-Nielsen, 1998] Bro-Nielsen, M. (1998). Rigid registration of CT,MR and Cryosection Images Using a GLCM Framework. *Lecture Notes in Computer Science*, 1205:171–180.

- [Brown, 1992] Brown, L. G. (1992). A survey of image registration techniques. *ACM Computer Surveys*, 24(4):325–376.
- [Bursell y col., 1996] Bursell, S. E., Clermont, A. C., Kinsley, B. T., Simonson, D. C., Aiello, L. M., y Wolpert, H. A. (1996). Retinal blood changes in patients with insulin-dependent diabetes mellitus and no diabetic retinopathy. *Investigative Ophthalmology and Visual Science*, 37(5):886–897.
- [Bursell y col., 1992] Bursell, S. E., Clermont, A. C., Shiba, T., y King, G. L. (1992). Evaluating retinal circulation using video fluorescein angiography in control and diabetic rats. *Current Eye Research*, 11(4):287–295.
- [Canny, 1986] Canny, J. (1986). A computational approach to edge-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.
- [Chen y col., 1998] Chen, M., Kanade, T., y Pomerleau, D. (1998). Medical image atlas registration using deformable models for anomaly detection. En *Proceedings of the Image Understanding Workshop*, páginas 1085–1090.
- [Cheng y col., 1994] Cheng, Q., Defrise, M., y Donnck, F. (1994). Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(12):1156–1168.
- [Christensen y Johnson, 2001] Christensen, G. E. y Johnson, H. J. (2001). Consistent image registration. *IEEE Transactions on Medical Imaging*, 20(7):568–582.
- [Chua y Jarvis, 1996] Chua, C. S. y Jarvis, R. (1996). 3-D free form surface registration and object recognition. *International Journal of Computer Vision*, 17(1):77–99.
- [Clermont y col., 1997] Clermont, A. C., Aiello, L. P., Mori, F., Aiello, L. M., y Bursell, S. E. (1997). Vascular endothelial growth factor and severity of nonproliferative diabetic retinopathy mediate retinal hemodynamics in vivo: A potential role for vascular endothelial growth factor in the progression of nonproliferative diabetic retinopathy. *American Journal of Ophthalmology*, 124(4):433–446.
- [Cohen y Cohen, 1993] Cohen, L. y Cohen, I. (1993). Finite-element methods for active contour models and balloons for 2D and 3D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1131–1147.

- [Cover y Thomas, 1991] Cover, T. M. y Thomas, J. A. (1991). *Elements in Information Theory*. John Wiley and Sons, New York.
- [Domingo y col., 1997] Domingo, J., Ayala, G., Simó, A., de Ves, E., Martínez-Costa, L., y Marco, P. (1997). Irregular motion recovery in fluorescein angiograms. *Pattern Recognition Letters*, 18(9):805–821.
- [Duda y Hart, 1973] Duda, R. O. y Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. John Wiley and Sons.
- [Ettinger y col., 1994] Ettinger, G. J., Grimson, W. E. L., Lozano-Pérez, T., Wells III, W. M., White, S. J., y Kikinis, R. (1994). Automatic registration for multiple sclerosis change detection. En *Proceedings of IEEE Workshop on Biomedical Image Analysis*, páginas 297–306.
- [Evans y col., 1996] Evans, A. C., Collins, D. L., Neelin, P., y Marrett, T. S. (1996). *Corrective Analysis of Three Dimensional Brain Images*, chapter 6, páginas 99–114. Computer-Integrated Surgery, Technology and Clinical Applications. MIT Press, Cambridge, MA.
- [Fang y col., 1996] Fang, S., Raghavan, R., y Richtsmeier, J. T. (1996). *Volume Morphing Methods for Landmark based 3D Image Deformation*, volumen 2710. Bellingham, WA.
- [Faugeras, 1993] Faugeras, O. (1993). *Three-dimensional Computer Vision. A Geometric Viewpoint*. MIT Press, Cambridge, MA.
- [Feke y col., 1994] Feke, T., Buzney, S. M., Ogasawara, H., Fujio, N., Goger, D. G., Spack, N. P., y Gabbay, K. H. (1994). Retinal circulatory abnormalities in type I diabetes. *Investigative Ophthalmology and Visual Science*, 35:2968–2975.
- [Foley y col., 1990] Foley, J. D., van Dam, A., Feiner, S. K., y Hughes, J. F. (1990). *Computer Graphics, Principles and Practice*. Addison Wesley, Reading, MA.
- [Fujii y col., 2003] Fujii, G., De Juan, E., Humayun, M., Sunness, J., Chang, T., y Rossi, J. (2003). Characteristics of visual loss by scanning laser ophthalmoscope microperimetry in eyes with subfoveal choroidal neovascularization secondary to age-related macular degeneration. *American Journal of Ophthalmology*, 136(6):1067–1078.

- [Gee y col., 1993] Gee, J. C., Reivich, M., y Bajcsy, R. (1993). Elastically deforming a three-dimensional atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17(2):225–236.
- [Glavinović, 1996] Glavinović, M. I. (1996). Comparison of Parzen density and frequency histogram as estimators of probability density functions. *Pflügers Archiv European Journal of Physiology*, 443(1-2):174–179.
- [González y Woods, 1992] González, R. C. y Woods, R. E. (1992). *Digital Image Processing*. Addison Wesley.
- [Grunwald y col., 1986] Grunwald, J. E., Riva, C. E., Sinclair, S. H., Brucker, A. J., y Petrig, B. L. (1986). Laser doppler velocimetry study of retinal circulation in diabetes mellitus. *Archives of Ophthalmology*, 104(7):991–996.
- [Hamilton y col., 1932] Hamilton, W. F., Moore, J. W., Kinsman, J. M., y Spurling, R. G. (1932). Studies on the circulation, IV. further analysis of the injection method and of changes in hemodynamics and physiological ana pathological conditions. *American Journal of Physiology*, 99:534–551.
- [Hart y Goldbaum, 1994] Hart, W. E. y Goldbaum, M. H. (1994). Registering retinal images using automatically selected control point pairs. En *Proceedings of the IEEE Conference on Image Processing*, páginas 576–580.
- [Hsu y Loew, 2001] Hsu, L.-Y. y Loew, M. H. (2001). Fully automatic 3D feature based registration of multi-modality medical images. *Image and Vision Computing*, 19(1-2):75–85.
- [Ingber, 1993] Ingber, A. L. (1993). Simulated annealing: Practice versus theory. *Mathematical and Computer Modelling*, 18(11):29–57.
- [Jähne, 1993] Jähne, B. (1993). Spatio-temporal image processing. *Lecture Notes in Computer Science*, 751:153–152.
- [Jähne, 2002] Jähne, B. (2002). *Digital Image Processing*, chapter 13. Springer.
- [Jones y Metaxas, 1997] Jones, T. N. y Metaxas, D. (1997). Segmentation using deformable models with affinity-based localization. *Lecture Notes in Computer Science*, 1205:53–62.

- [Kanal y col., 1981] Kanal, L. N., Lambird, B. A., Lavine, D., y Stockman, G. C. (1981). Digital registration of images from similar and dissimilar sensors. En *Proceedings of the International Conference on Cybernetics and Society*, páginas 347–351.
- [Kass y col., 1988] Kass, M., Witkin, A., y Terzopoulos, D. (1988). Active contour models. *International Journal of Computer Vision*, 1(2):321–331.
- [Kohner y col., 1975] Kohner, E. M., Hamilton, A. M., y Saunders, S. J. (1975). The retinal blood flow in diabetes. *Diabetologica*, 11(1):23–33.
- [Koyama y col., 1990] Koyama, T., Matsuo, N., Shimizu, K., Mihara, M., Tsuchida, Y., Wolf, S., y Reim, M. (1990). Retinal circulation times in quantitative fluorescein angiography. *Graefe's Archive of Clinical and Experimental Ophthalmology*, 228:442–446.
- [Kuhl y Giardina, 1982] Kuhl, F. P. y Giardina, C. R. (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, 18(1):293–300.
- [Laliberté y col., 2003] Laliberté, F., Gagnon, L., y Sheng, Y. (2003). Registration and fusion of retinal images – an evaluation study. *IEEE Transactions on Medical Imaging*, 22(5):661–673.
- [Leventon, 2000] Leventon, M. E. (2000). *Statistical Models in Medical Image Analysis*. PhD thesis, MIT.
- [Lewis, 1995] Lewis, J. P. (1995). Fast normalized cross-correlation. *Vision Interface*, páginas 120–123.
- [Likar y Pernus, 2001] Likar, B. y Pernus, F. (2001). A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing*, 19(1-2):33–44.
- [Lloret y col., 2001] Lloret, D., Mariño, C., Serrat, J., López, A. M., y Villanueva, J. J. (2001). Landmark-based registration of full SLO video sequences. En *Proceedings of the IX Spanish Symposium on Pattern Recognition and Image Analysis*, volumen I, páginas 189–194.

- [Lois y col., 1999] Lois, N., Halfyard, A. S., Bunce, C., Bird, A. C., y Fitzke, F. W. (1999). Reproducibility of fundus autofluorescence measurements obtained using a confocal scanning laser ophthalmoscope. *British Journal of Ophthalmology*, 83:276–279.
- [López Peña y col., 1999] López Peña, A., Lumbreras, F., Serrat, J., y Villanueva, J. J. (1999). Evaluation of methods for ridge and valley detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):327–335.
- [López y col., 2000] López, A., Lloret, D., Serrat, J., y Villanueva, J. J. (2000). Multilocal creaseness based on the level-set extrinsic curvature. *Computer Vision and Image Understanding*, 77(1):111–144.
- [Maes y col., 1997] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., y Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198.
- [Maintz y col., 1996] Maintz, J. B. A., Van den Elsen, A. P., y Viergever, M. A. (1996). Evaluation of ridge seeking operator for multimodality medical image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):353–365.
- [Maintz y Viergever, 1998] Maintz, J. B. A. y Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36.
- [Mariño y col., 2001] Mariño, C., Penas, M., Penedo, M. G., Lloret, D., y Carreira, M. J. (2001). Integration of mutual information and creaseness based methods for the automatic registration of SLO sequences. En *Proceedings del VI Simposio Ibero-Americano de Reconocimiento de Patrones (SIARP 2001)*, páginas 77–83.
- [Mariño y col., 2006] Mariño, C., Penedo, M. G., Carreira, M. P. M. J., y González, F. (2006). Personal authentication using digital retinal images. *Pattern Analysis and Applications*. DOI 10.1007/s10044-005-0022-6.
- [Mariño y col., 2003] Mariño, C., Penedo, M. G., Carreira, M. J., y González, F. (2003). Retinal angiography based authentication. *Lecture Notes in Computer Science*, 2905:306–313.

- [Mariño y col., 2005] Mariño, C., Penedo, M. G., y Penas, M. (2005). Retinal based authentication via distributed web application. *Lecture Notes in Computer Science*, 3643:286–391.
- [Markov y col., 1993] Markov, M. S., Rylander, H. G., y Welch, A. J. (1993). Real-time algorithm for retinal tracking. *IEEE Transactions on Biomedical Engineering*, 40(12):1269–1281.
- [McInerney y Terzopoulos, 1996] McInerney, T. y Terzopoulos, D. (1996). Deformable models in medical image analysis: A survey. Technical report, Department of Computer Science, University of Toronto.
- [Moseley y Munro, 1994] Moseley, J. y Munro, P. (1994). A semiautomatic method for registration of portal images. *Medical Physics*, 21(4):551–558.
- [Nack, 1977] Nack, M. L. (1977). Rectification and registration of digital images and the effect of cloud detection. En *Proceedings of Machine Processing of Remotely Sensed Data*, páginas 12–23.
- [Nagin y col., 1985] Nagin, P., Schwartz, B., y Raynolds, G. (1985). Measurement of fluorescein in angiograms of the optic disc and retina using computerized image analysis. *Ophthalmology*, 92:547–552.
- [Nelder y Mead, 1965] Nelder, J. A. y Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7:308–313.
- [Niessen y col., 1997] Niessen, W., López, A., Van Enk, W., Van Roermund, P., ter Haar Romeny, B., y Viergever, M. (1997). In vivo analysis of trabecular bone architecture. *Lecture Notes in Computer Science*, 1230:435–440.
- [Noack y Sutton, 1994] Noack, J. y Sutton, D. (1994). An algorithm for the fast registration of image sequences obtained with a scanning laser ophthalmoscope. *Physics in Medicine and Biology*, 59(5):907–915.
- [Oppenheim y col., 1996] Oppenheim, A. V., Willsky, A. S., Hamid, S., y Hamid Nawab, S. (1996). *Signals and Systems*. Prentice Hall.
- [Ortega y col., 2006] Ortega, M., Mariño, C., Penedo, M. G., Blanco, M., y González, F. (2006). Personal authentication based on feature extraction and optic nerve location in digital retinal images. *Wseas Transactions on Computers*, 5:1169–1176.

- [Ott y col., 1990] Ott, D., Lades, M., Holthoff, K., y Eckmiller, R. (1990). A general numerical method evaluating three dimensional eye rotations by scanning laser ophthalmoscopy. *Ophthalmic Physiological Optics*, 10:286–290.
- [Pardo y col., 2001] Pardo, F., Leborán, V., Mariño, C., Penedo, M. G., Carreira, M. J., Mosquera, A., Cabello, D., Gómez-Ulla, F., y González, F. (2001). Retinal angiography image registration applied to hemodynamic variable measurement. En *Proceedings of the IX Spanish Symposium on Pattern Recognition and Image Analysis*, volumen II, páginas 139–144.
- [Patel y col., 1992] Patel, V., Rassam, S., Newsom, R., Wiek, J., y Kohner, E. (1992). Retinal blood flow in diabetic retinopathy. *BMJ (Clinical research ed.)*, 305(6855):678–683.
- [Peli y col., 1987] Peli, E., Augliere, R. A., y Timberlake, G. T. (1987). Feature-based registration of retinal images. *IEEE Transactions on Medical Imaging*, 6(3):272–278.
- [Pesavento y col., 2002] Pesavento, C., Harris, A., Cole, C., y Kagemann, L. (2002). Improving the analysis of arteriovenous passage times. *Graefe's Archive of Clinical and Experimental Ophthalmology*, 240(11):955–959.
- [Pinz y col., 1998] Pinz, A., Bernögger, S., Datlinger, P., y Kruger, A. (1998). Mapping the human retina. *IEEE Transactions on Medical Imaging*, 17(4):606–619.
- [Pinz y col., 1995] Pinz, A., Prantl, M., y Ganster, H. (1995). Affine matching of intermediate symbolic representation. En *Proceedings of Computer Analysis of Images and Patterns (CAIP 95)*, páginas 359–367.
- [Pluim y col., 2001] Pluim, J. P. W., Maintz, J. B. A., y Viergever, M. A. (2001). Mutual information matching in multiresolution contexts. *Image and Vision Computing*, 19(1-2):46–52.
- [Pratt, 1991] Pratt, W. K. (1991). *Digital Image Processing*. John Wiley and Sons, New York.
- [Press y col., 1992] Press, W., Teukolsky, S., Vetterling, W., y Flannery, B. (1992). *Numerical Recipes in C*. Cambridge University Press, 2nd edition.

- [Ritter y col., 1999] Ritter, N., Owens, R., y Cooper, J. (1999). Registration of stereo and temporal images of the retina. *IEEE Transactions on Medical Imaging*, 18(5):404–418.
- [Rohr, 1999] Rohr, K. (1999). Extraction of 3D anatomical point landmarks based on invariance principles. *Pattern Recognition*, 32(1):3–15.
- [Savi y col., 1995] Savi, A., Gilardi, M. C., Rizzo, G., Pepi, M., Landomi, C., Rossetti, C., Lucignami, G., Bartorelli, A., y Fazio, F. (1995). Spatial registration of echocardiographic and positron emission tomographic heart studies. *European Journal of Nuclear Medicine*, 22(3):243–247.
- [Schey, 1973] Schey, H. M. (1973). *DIV, GRAD, CURL and all that*. Norton, New York.
- [Shannon, 1948] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656.
- [Shannon y Warren, 1963] Shannon, C. E. y Warren, W. (1963). *The mathematical theory of communication*. Urban University of Illinois Press.
- [Sperber y Alm, 1997] Sperber, G. y Alm, A. (1997). Retinal mean transit time determined with an impulse-response analysis from video fluorescein angiograms. *Acta Ophthalmologica Scandinavica*, 75:532–536.
- [Stewart y col., 2003] Stewart, C., Tsai, C., y Roysam, B. (2003). The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transactions on Medical Imaging*, 22(11):1379–1394.
- [Studholme y col., 1996] Studholme, C., Hill, D. L. G., y Hawkes, D. J. (1996). Automated 3-D registration of MR and CT images of the head. *Medical Image Analysis*, 1(2):163–175.
- [ter Haar Romeny y Florack, 1993] ter Haar Romeny, B. y Florack, L. (1993). *The Perception of Visual Information (W.Hendee and P Well, Eds.)*, páginas 73–114. Springer-Verlag, Berlin/New York.
- [Toga y Thompson, 2001] Toga, A. W. y Thompson, P. M. (2001). The role of image registration in brain mapping. *Image and Vision Computing*, 19(1-2):3–24.

- [Vajda, 1989] Vajda, I. (1989). *Theory of Statistical Inference and Information*. Kluwer Academic Publishers.
- [Van den Elsen y col., 1995] Van den Elsen, P. A., Antoine Maintz, J. B., D. Pol, E.-J., y Viergever, M. A. (1995). Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Transactions on Medical Imaging*, 14(2):384–396.
- [Van den Elsen y col., 1993] Van den Elsen, P. A., Pol, E. J. D., y Viergever, M. A. (1993). Medical image matching - a review with classification. *IEEE Engineering in Medicine and Biology*, 1(2):26–39.
- [Viola y Wells III, 1995] Viola, P. y Wells III, W. M. (1995). Alignment by maximization of mutual information. *Proceedings of the International Conference on Computer Vision*, páginas 16–23.
- [Viola, 1995] Viola, P. A. (1995). Alignment by maximization of mutual information. Technical Report 1548, Massachusetts Institute of Technology.
- [Wade y Fitzke, 1998] Wade, A. R. y Fitzke, F. W. (1998). A fast, robust pattern recognition system for low level image registration and its application to retinal imaging. *Optics Express*, 5(3):190–197.
- [Wells y col., 1996] Wells, W. M., Viola, P., Atsumi, H., Nakajima, S., y Kikinis, R. (1996). Multimodal volume registration by maximization mutual information. *Medical Image Analysis*, 1(1):35–51.
- [Wolf y col., 1989] Wolf, S., Jung, F., Kiesewetter, H., Körber, N., y Reim, M. (1989). Video fluorescein angiography: method and clinical application. *Graefe's Archive of Clinical and Experimental Ophthalmology*, 227:145–151.
- [Yang y col., 1995] Yang, Y., Moon, S., Lee, S., y Kim, J. (1995). Measurement of retinal blood flow with fluorescein leucocyte angiography using a scanning laser ophthalmoscope in rabbits. *British Journal of Ophthalmology*, 80:475–479.
- [Yoshida y col., 1983] Yoshida, A., Feke, G. T., Morales-Stoppello, J., Collas, G. D., Goger, D. G., y McMeel, J. W. (1983). Retinal blood flow alterations during progression of diabetic retinopathy. *Archives of Ophthalmology*, 101(2):225–227.

-
- [Zana y Klein, 1999] Zana, F. y Klein, J. C. (1999). A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform. *IEEE Transactions on Medical Imaging*, 18(5):419–428.
- [Zitová y Flusser, 2003] Zitová, B. y Flusser, J. (2003). Image registration methods: a survey. *Image Vision and Computing*, 21(11):977–1000.