

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Programa de Doctorado:  
INGENIERÍA BIOMÉDICA

**Tesis Doctoral**

**Registro Elástico de Imágenes Médicas  
Multimodales.  
Aplicación en Oncología.**

Gaspar Delso

Directora: Alicia Casals

Centre de Recerca en Enginyeria Biomèdica  
Barcelona, Octubre 2003



A mis padres,



# Agradecimientos

De ningún modo hubiera podido llevar esta tesis a buen termino sin el apoyo y la colaboración de muchas personas, a quienes quisiera desde aquí hacer llegar mi más sincero reconocimiento:

- A la Dra. Alicia Casals, que ha dirigido este trabajo de tesis, por su ayuda en todo momento, a pesar de lo mucho que ha sido preciso alejarse de la idea original.
- A la Dra. Isabelle Bloch, por darme acceso a su departamento en la École Nationale Supérieure des Télécommunications, y al proyecto Oncomatching, que se ha convertido en el eje de este trabajo.
- A Oscar Càmara, por su lealtad, por hacer de una potencial competencia una colaboración que ha fructificado más allá de lo que nadie hubiera podido imaginar.
- A los responsables de Segami, por poner a nuestra disposición unos contactos y medios que tantos años de trabajo representan.
- Al Dr. Hervé Foehrenbach, por su paciencia, por supervisar la evolución de este trabajo y por su colaboración en el desarrollo del método de evaluación.
- Y finalmente, a todos los compañeros que, tanto en Barcelona como en París, dentro y fuera de la universidad, me han acompañado durante esta etapa que toca a su fin.

A todos ellos y a muchos más, sinceramente, gracias!



# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Campos de aplicación . . . . .	2
1.2. Registro en medicina . . . . .	7
1.2.1. Tipos de imágenes médicas . . . . .	8
1.2.2. Aplicaciones del registro en medicina y biología. . . . .	17
<b>2. Métodos de registro</b>	<b>21</b>
2.1. Modelo de sistema de registro . . . . .	21
2.2. Dimensionalidad de las imágenes . . . . .	24
2.3. Transformaciones geométricas . . . . .	27
2.3.1. Transformaciones lineales . . . . .	27
2.3.2. Transformaciones no lineales . . . . .	35
2.4. Medidas de similitud . . . . .	39
2.4.1. Medidas no basada en el contenido de las imágenes . . . . .	40
2.4.2. Medidas extrínsecas . . . . .	40
2.4.3. Medidas intrínsecas . . . . .	42
2.4.4. Estimación del histograma y la fdp. . . . .	53
2.5. Optimización . . . . .	60
2.5.1. Cálculo directo . . . . .	60
2.5.2. Cálculo iterativo . . . . .	60
2.6. Modalidades de imagen implicadas . . . . .	68
2.7. Sujeto visualizado . . . . .	69
2.8. Objeto visualizado . . . . .	70
2.9. Necesidad de interacción . . . . .	70
<b>3. Diagnóstico en oncología de tórax y abdomen</b>	<b>73</b>
3.1. Motivación . . . . .	74
3.2. Antecedentes . . . . .	75
3.3. Condicionantes del método . . . . .	80
3.4. Método propuesto . . . . .	84
3.5. Inicialización mediante referencias . . . . .	89
3.5.1. Método 1: Proyección de superficies. . . . .	92
3.5.2. Método 2. Deformación RMS-FFD . . . . .	100

3.5.3. Resultados . . . . .	104
3.6. Conclusión . . . . .	106
<b>4. Segmentación de estructuras anatómicas</b>	<b>109</b>
4.1. Objetivo de la segmentación . . . . .	109
4.2. Segmentación por umbralizado . . . . .	112
4.2.1. El método de las k-medias . . . . .	113
4.2.2. Morfología matemática . . . . .	118
4.3. Modelos deformables . . . . .	120
4.3.1. Modelos deformables paramétricos 2D: Contornos activos . . . . .	121
4.3.2. Difusión del vector gradiente . . . . .	124
4.3.3. Modelos deformables paramétricos 3D: Superficies activas . . . . .	128
4.4. Conclusión . . . . .	135
<b>5. Validación</b>	<b>137</b>
5.1. Problemática . . . . .	137
5.2. Base de datos . . . . .	138
5.3. Medidas numéricas . . . . .	139
5.4. Protocolo de evaluación . . . . .	139
<b>6. Conclusiones</b>	<b>153</b>
6.1. Objetivos cumplidos . . . . .	153
6.2. Futuras líneas de investigación . . . . .	155
6.3. Contribuciones . . . . .	157
<b>A. Splines</b>	<b>159</b>
A.1. Introducción . . . . .	159
A.2. B-splines . . . . .	160
A.3. Interpolación B-spline por filtrado digital . . . . .	162
A.4. Ejemplo: Aplicación a la transformación geométrica de imágenes . . . . .	163
A.5. Otras propiedades óptimas . . . . .	164
A.5.1. Propiedades variacionales . . . . .	164
A.5.2. Suavizado mediante splines . . . . .	165
A.5.3. Regularidad máxima y mínimo soporte . . . . .	165
A.5.4. Splines fraccionarias . . . . .	166
A.6. Aplicaciones . . . . .	166
A.7. Conclusión . . . . .	166
<b>B. Interpolación</b>	<b>169</b>
B.1. Introducción . . . . .	169
B.2. Métodos de interpolación . . . . .	170



<b>C. Publicaciones</b>	<b>179</b>
<b>Bibliografía</b>	<b>183</b>



# Capítulo 1

## Introducción

Existen numerosas aplicaciones donde es necesario integrar información proveniente de distintas fuentes, sean estas distintos sensores o distintas adquisiciones efectuadas por un mismo sensor.

Se denomina *fusión* al procedimiento por el cual se combinan los datos provenientes de distintas fuentes en una única representación.

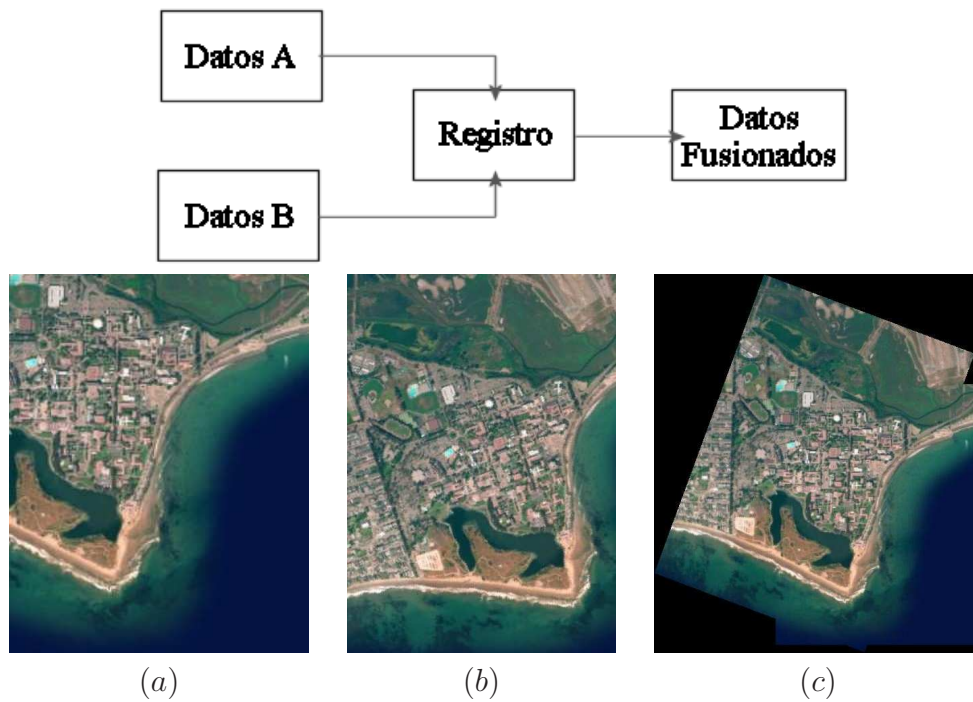
A menudo no es trivial efectuar dicha fusión de la información a causa de la divergencia entre los datos, sea debido a diferencias en los parámetros de los sensores (punto de vista, tecnología empleada, . . . ), o por diferencias del propio objeto visualizado (condiciones de iluminación, cambios en el tiempo, . . . ).

Se denomina *registro* al procedimiento mediante el cual se establece la correspondencia entre los datos provenientes de distintas fuentes. Para mayor simplicidad, en adelante nos referiremos a estos datos como *volúmenes* o *imágenes*.

Dicho procedimiento consiste en establecer una aplicación entre el espacio en que está definido un volumen de datos y el espacio en que está definido el otro, de modo que a cada elemento del volumen origen se le asigne la posición en el volumen destino que corresponde a ese mismo elemento de la escena.

El registro de imágenes proporciona una valiosa herramienta en muy diversos campos de aplicación, habiéndose extendido su uso en los últimos años debido a la proliferación de nuevas técnicas de medición, así como al aumento de la disponibilidad de la potencia de cálculo necesaria para llevarlo a cabo. En general, se pueden beneficiar de este tipo de técnicas aquellos sistemas que precisen de:

- Fusionar información proveniente de distintos sensores.
- Determinar o compensar los cambios en imágenes tomadas en distintos momentos o bajo distintas condiciones de adquisición.
- Inferir información tridimensional a partir de un conjunto de imágenes bidimensionales obtenidas del desplazamiento de la cámara respecto al objeto visualizado y viceversa.
- Efectuar reconocimiento de formas o de objetos a partir de un modelo.



**Figura 1.1:** Una operación de registro permite fusionar distintas fuentes de datos (a) y (b) en una representación común (c). (*Vision Lab, Electrical & Computer Engineering department. University of California Santa Barbara*)

El rango de aplicaciones que precisan de las prestaciones arriba descritas es muy amplio, abarcando desde la astronomía a la geología, desde la aviación a la medicina. En el siguiente apartado se revisan algunos de los campos a los que ya se está aplicando el registro de imágenes, si bien este tipo de técnicas están todavía en pleno desarrollo, y sin duda en los próximos años han de surgir procedimientos originales que permitan incorporarlas a nuevos ámbitos.

## 1.1. Campos de aplicación

El registro de imágenes se conoce y emplea desde hace tiempo en áreas tan dispares como la geografía, la robótica o la medicina. En esta sección se citan algunas de las aplicaciones más importantes de este tipo de técnicas, relacionadas con la fusión de datos provenientes de sensores remotos y la visión por computador.

A continuación se describen brevemente distintos campos de aplicación en la detección mediante sensores remotos y el procesamiento de datos. En la figura 1.2 pueden encontrarse algunos ejemplos de registro para el estudio de la variabilidad temporal de zonas urbanas, agrícolas y forestales.

- Guiado autónomo para el aterrizaje de aviones.

Los sistemas de aterrizaje autónomo (ALG) ofrecen al piloto información visual que permite aterrizar en condiciones adversas, integrada a partir de datos provenientes del espectro visible, del infrarrojo (adecuado en operaciones nocturnas) y del sistema de radar (capaz de penetrar la niebla).

- Meteorología.

La meteorología es sin duda un caso típico de aplicación que precisa de la integración de datos procedentes de muy diversas fuentes. Así, por ejemplo, para generar mapas de precipitaciones continuos a la vez en el tiempo y en el espacio, es preciso combinar la información brindada por las redes convencionales de medición y los radares meteorológicos (continua en el tiempo pero con mala cobertura en el espacio) con imágenes de satélite (esporádicas pero con buena resolución espacial).

- Agricultura.

Muy estrechamente relacionada con la meteorología, la agricultura se beneficia también de la fusión de datos. Así, por ejemplo, las imágenes térmicas nocturnas permiten localizar fácilmente las masas de agua durante la estación de las lluvias en países tropicales. Sin embargo es frecuente que las nubes oculten parte de la escena, por lo que es preciso efectuar el registro de imágenes tomadas a lo largo de unos diez días para conseguir una buena composición.

- Geología.

Dos aplicaciones típicas del registro de datos en geología son la creación de mapas a partir de un mosaico de imágenes aéreas o de satélite, y la integración en un sistema de información geográfica (GIS) de dichos mapas con datos de altitud obtenidos mediante lidar (light detection and ranging) aerotransportado.

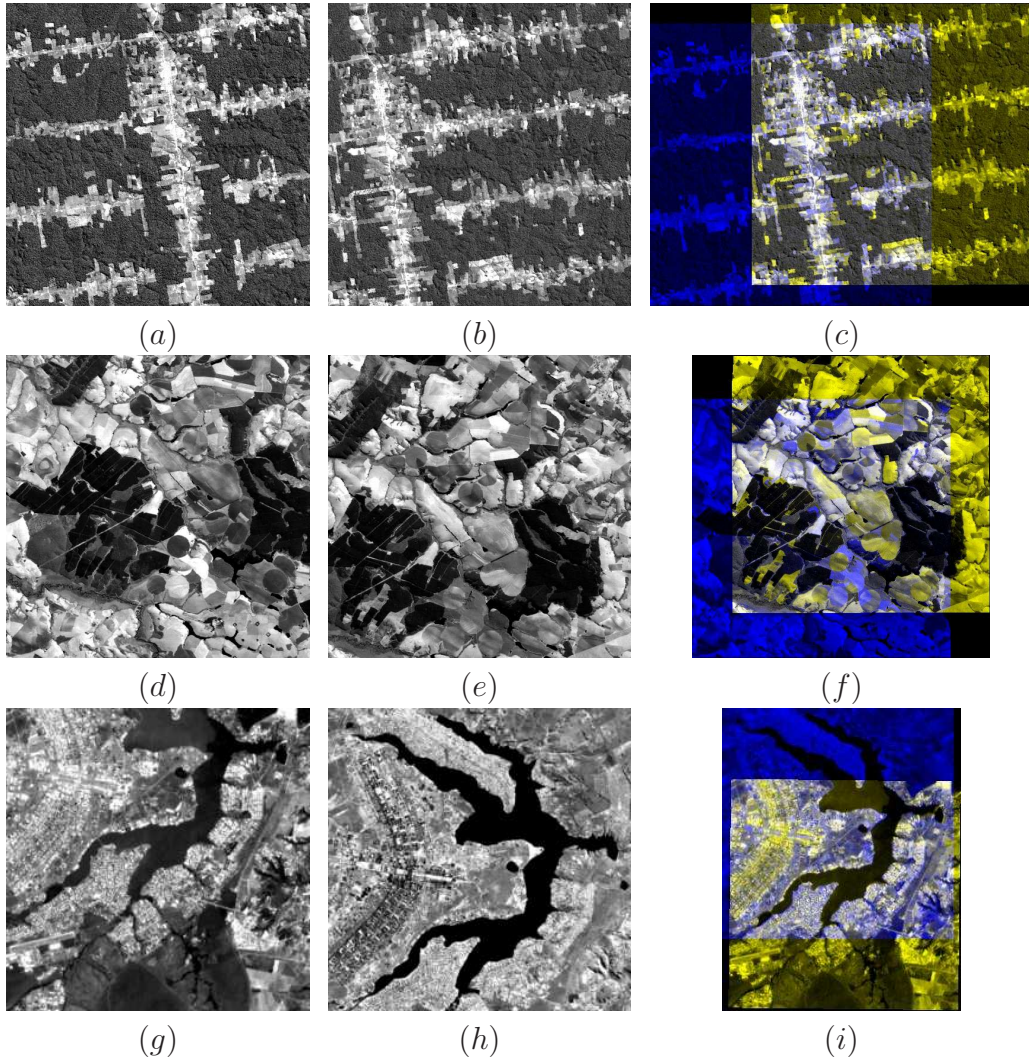
Como curiosidad podemos mencionar que se han empleado técnicas similares en la exploración de Marte, tanto para generar mapas de elevación a partir de imágenes orbitales como para componer un paisaje tridimensional a partir de los datos captados por los sensores del robot Pathfinder.

- Oceanografía.

En principio similar a la geografía, en este campo cobra especial interés la fusión de datos debido a la peor calidad de las imágenes disponibles. Algunos ejemplos podrían ser la creación de mapas submarinos a partir de múltiples imágenes sonar de barrido, o el estudio de la estratigrafía submarina mediante imágenes sonar y técnicas de reflexión sísmica.

- Exploración petrolífera y mineralógica.

Dado su extraordinario interés económico, éste fue uno de los ámbitos donde antes se planteó la fusión de datos procedentes de múltiples sensores. Así, por ejemplo, se emplean datos de espectrometría aérea visible e infrarroja (AVIRIS), radar de



**Figura 1.2:** Registro de datos en aplicaciones forestales  $(a)(b)(c)$ , agrícolas  $(d)(e)(f)$  y urbanas  $(g)(h)(i)$ . (*Vision Lab, Electrical & Computer Engineering department. University of California*)

apertura sintética (AIRSAR) y otras modalidades para estudiar la morfología y composición mineralógica de un terreno. Otra aplicación es la combinación de barridos de radar y sonar para localizar posibles yacimientos petrolíferos.

- Astronomía.

Empleando el registro de imágenes se puede sustituir la adquisición de imágenes con largo tiempo de exposición por un promediado de múltiples adquisiciones con menor tiempo de exposición. Si bien es preferible una única imagen con mayor exposición, la alternativa de promediar diversas imágenes no debe descartarse, ya que permite prescindir del sistema de guiado y previene la saturación. Del mismo modo, es posible combinar datos provenientes de distintas bandas espectrales en una única representación.

- Estudios urbanos y de polución.

El registro de imágenes aéreas o de satélite con mapas urbanos facilita considerablemente la caracterización de la infraestructura de una población, siendo empleado en la generación de vistas tridimensionales, planificación y desarrollo urbano.

- Estudios forestales y de deforestación.

Combinando imágenes satélite AVHRR (radiómetro avanzado de muy alta resolución) en las bandas de rojo visible, infrarrojo cercano e infrarrojo térmico con imágenes Landsat en las bandas de rojo visible, infrarrojo cercano e infrarrojo medio, se puede efectuar un seguimiento detallado de la superficie forestal y deforestación.

- Aplicaciones militares.

Desgraciadamente, muchas de las técnicas de registro empleadas en aplicaciones civiles fueron inicialmente concebidas con fines militares. Hoy en día la investigación en este campo sigue siendo sumamente activa, como por ejemplo para la detección, reconocimiento y seguimiento de objetivos en presencia de *clutter* o contramedidas. Actualmente ciertos aviones de combate van equipados con una base de datos que contiene modelos CAD de los posibles objetivos, empleando técnicas de registro de imágenes para determinar la mejor correspondencia.

- Aplicaciones biomédicas

Existen innumerables aplicaciones en el campo de la medicina y la biología en las que se emplean técnicas de registro de imágenes. Algunos ejemplos pueden ser la fusión de datos anatómicos y funcionales para la localización y diagnóstico de patologías, la integración de datos preoperatorios en intervenciones asistidas por computador y el alineamiento de secuencias de imágenes para estudiar la evolución de afecciones crónicas.



Un importante ámbito al que desde hace tiempo se aplican con gran éxito las técnicas de registro de imágenes es el de la visión por computador. A continuación se enumeran diferentes procedimientos comunes en visión por computador y su posible relación con el registro de imágenes.

- Segmentación.

El registro de imágenes se puede emplear para mejorar la segmentación, tanto mediante la combinación de información proveniente de distintos sensores como haciendo corresponder un modelo sobre la imagen a segmentar. De forma recíproca, la segmentación de una imagen puede facilitar en gran medida un proceso de registro.

- Reconocimiento de objetos y patrones.

El reconocimiento de objetos puede interpretarse como un problema de registro de una imagen con un conjunto de modelos con el fin de hallar aquél que presenta una mejor correspondencia. Casos particulares de especial relevancia son el reconocimiento de caracteres (OCR), el reconocimiento del habla y el análisis de huellas dactilares.

- Estereovisión.

Para inferir información tridimensional a partir de varias vistas de una imagen es preciso determinar las correspondencias entre los elementos visualizados en una y otra imagen. Hay distintas técnicas de registro que permiten llevar a cabo dicha tarea, sea segmentando referencias o elementos de interés de las escenas, sea definiendo un campo de deformación directamente entre las escenas sin segmentar.

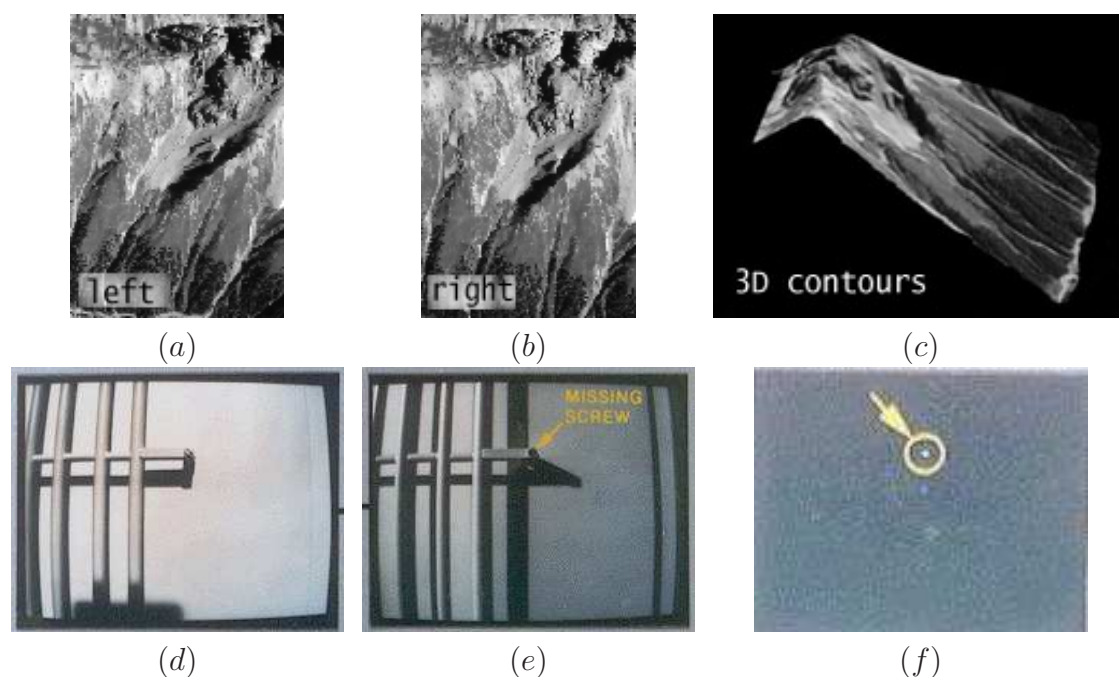
- Reconstrucción de formas.

En cierto modo relacionado con la estereovisión, aunque más general en el sentido de que puede incorporar información de otros tipos de sensores, el registro de datos para la reconstrucción de formas tridimensionales tiene aplicaciones tan importantes como el acoplamiento (*docking*), la prensión (*grasping*) y la detección de fallas.

- Seguimiento del movimiento.

Es posible analizar el movimiento de un elemento contenido en una secuencia, una vez éste ha sido registrado con un modelo. Así, el proceso de análisis y predicción del movimiento se reduce a un proceso iterativo de registro, empleándose como estimación inicial en cada iteración la posición obtenida en la iteración anterior.





**Figura 1.3:** Registro de datos en aplicaciones de procesamiento de imágenes y visión por computador: reconstrucción tridimensional por estereovisión (a)(b)(c), detección de fallos en la inspección de plataformas orbitales (d)(e)(f).

## 1.2. Registro en medicina

Gracias a la proliferación de numerosas nuevas técnicas de adquisición de imágenes médicas que han ido apareciendo en los últimos años, se han abierto nuevas vías para la detección, diagnóstico y tratamiento de numerosas patologías.

A la vez que se multiplican las modalidades que posibilitan la adquisición de información complementaria sobre el estado de un paciente, crece también el interés en disponer de métodos de registro que permitan fusionar toda esa información en una representación conjunta.

No obstante, las aplicaciones médicas presentan un conjunto de particularidades y requisitos sumamente restrictivos que las distinguen de otros campos en los que se emplea el registro de imágenes. Se trata en general de aplicaciones en las que un error puede acarrear graves consecuencias, por lo que se exige un alto grado de robustez y precisión al sistema.

Por otra parte, debido a que se priorizan la seguridad y, en menor medida, la comodidad del paciente, no es posible adquirir las imágenes en condiciones óptimas para su análisis. Un claro ejemplo de esta situación es la adquisición de imágenes mediante radiación ionizante, característica de modalidades como la CT, PET y SPECT, descritas en el siguiente apartado. Todas estas modalidades tienen la capacidad de brindar imágenes de gran calidad del paciente. En cambio, se prefiere recortar al

máximo el tiempo de exposición y la dosis de radiación a la que se somete al sujeto, sacrificando contraste y resolución en pro de la seguridad. Por esta razón en tratamiento de imágenes médicas, al contrario que en las aplicaciones de visión industrial clásicas, a menudo se debe trabajar con datos ruidosos y poco contrastados.

Se describen a continuación las principales modalidades de imágenes médicas, indicándose su formato y el tipo de información que proporcionan.

### 1.2.1. Tipos de imágenes médicas

- Rayos X

La *radiografía* es la modalidad más antigua y posiblemente también la más común de adquisición de imágenes en la práctica médica. Las imágenes radiográficas son proyecciones bidimensionales que muestran la atenuación de un haz de radiación ionizante (rayos X) al atravesar al paciente. Esta atenuación está relacionada con la naturaleza y densidad de los tejidos que son atravesados por el haz.

La *fluoroscopia* es una variante de la radiografía que se emplea para visualizar en tiempo real ciertos procesos dinámicos dentro del cuerpo. En este caso se emplea una pantalla de fluorescencia para visualizar un haz continuo de rayos X de baja potencia que circula a través del paciente. Esta técnica se puede emplear tanto para el diagnóstico (ingestión o enema de bario para el examen del sistema digestivo, angiografía para el circulatorio,...), como en el guiado durante una intervención (angioplastia,...)

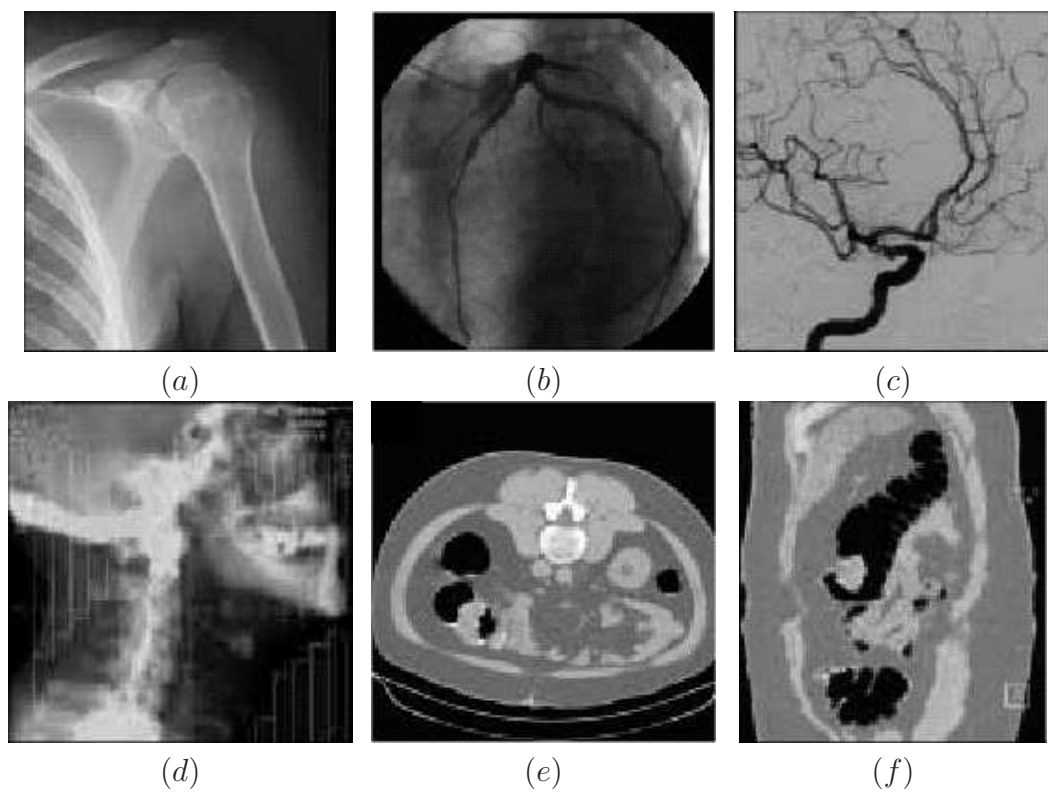
Otra variante de la radiografía, de interés en el campo de la radioterapia, son las imágenes *portal*. Estas imágenes se obtienen empleando como fuente de rayos X el propio generador de alta potencia que se utilizará para efectuar la radioablación del tejido patológico. Debido a que el rango de energías empleadas es mucho mayor que las de la radiografía convencional, la imagen obtenida tendrá poco contraste entre los distintos tejidos.

Finalmente, una modalidad muy extendida en el estudio de los vasos sanguíneos es la *angiografía de sustracción digital (DSA)*. Consiste en efectuar la sustracción de una angiografía (radiografía con un elemento de contraste que resalta los vasos sanguíneos) de la misma radiografía previa a la inyección del contraste. De este modo se elimina el fondo de la imagen, permaneciendo tan sólo las arterias afectadas por el contraste.

La figura 1.4 ilustra estas modalidades.

- Tomografía computarizada

Si bien se basa en el mismo principio físico que la radiografía, la *tomografía computarizada de rayos X (CT)* ha cobrado tanta importancia que merece ser considerada como una modalidad a parte.



**Figura 1.4:** Imágenes radiográficas: Radiografía convencional (a); imágenes fluoroscópicas de angiografía (b) y DSA (c); portal (d); y cortes axial (e) y sagital (f) de un volumen CT. (Siemens)

El principio de funcionamiento de esta técnica consiste en proyectar a través del paciente un haz de rayos X a la vez que el sistema emisor-receptor realiza un movimiento circular entorno al mismo. La información resultante es transmitida a un ordenador, donde un algoritmo de retroproyección permite construir una imagen del coeficiente de atenuación de los tejidos en el plano irradiado.

Repitiendo esta operación a medida que se recorre longitudinalmente la zona que se desea estudiar, se obtiene una representación tridimensional por cortes, o *volumen tomográfico*, que representa la opacidad a los rayos X de cada elemento de volumen, o *voxel*, de dicha zona.

Actualmente existen tomógrafos capaces de realizar un barrido helicoidal de la región de interés, consiguiendo así una reducción significativa del tiempo de adquisición.

Las figuras 1.4.e y 1.4.f muestran sendos cortes *axial* (perpendicular al eje del cuerpo) y *sagital* (paralelo al plano de simetría del cuerpo) de un volumen tomográfico. Otro corte clásico, no mostrado en la figura, es el *coronal*, perpendicular a los anteriores.

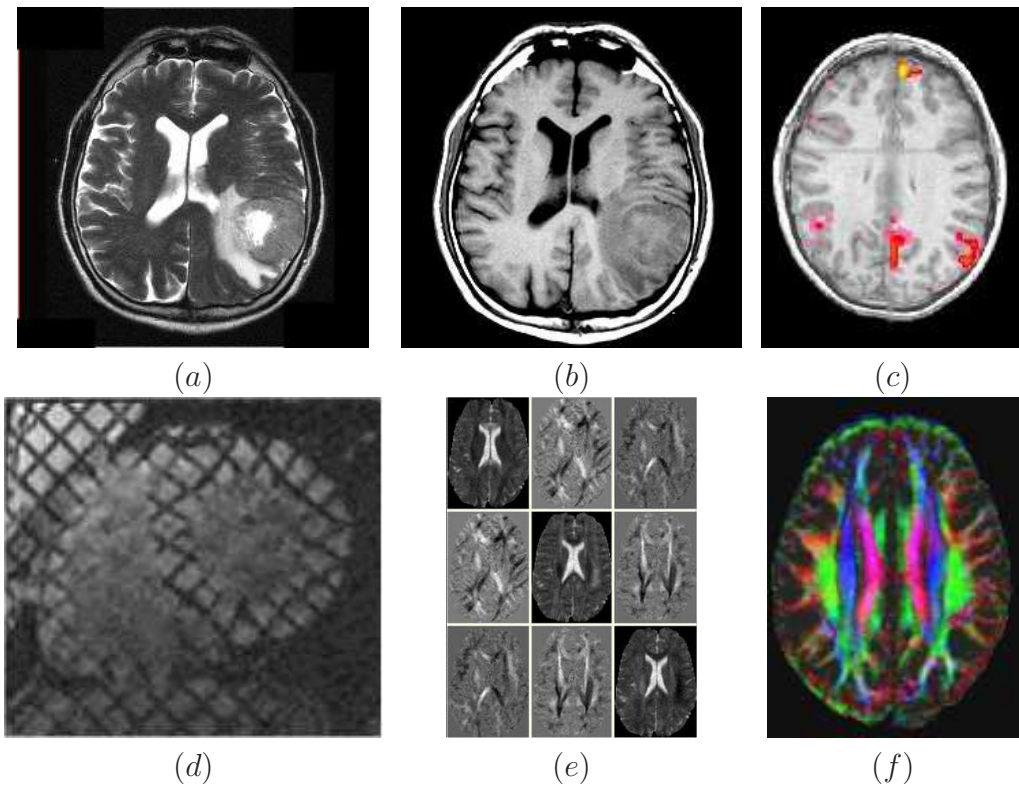
- **Resonancia magnética nuclear**

La *resonancia magnética nuclear (MRI)* permite obtener volúmenes tomográficos de datos sobre la composición química del tejido visualizado. Para ello se excitan los protones de los tejidos mediante un determinado campo magnético, observándose su comportamiento frente a cambios de dicha excitación. La respuesta obtenida depende de las propiedades locales del entorno de cada protón, tales como su movilidad y la homogeneidad local del campo magnético. La señal de excitación puede ajustarse de modo que acentúe algunas de estas propiedades, de modo que bajo el denominador común de resonancia magnética se incluye toda una gama de imágenes que reflejan distintas características del tejido (densidad de protones (PD), relajación T1 y T2, tensor de difusión (DT),...).

Otras modalidades basadas en la resonancia magnética que citamos aquí por su importancia en la práctica médica son la *tagged MR*, para el estudio de elementos en movimiento; la *resonancia magnética funcional (fMRI)*, para el estudio de la activación cortical en respuesta a determinados estímulos; la *resonancia magnética de perfusión (pMRI)* y la *resonancia magnética abierta*, que permite realizar la visualización a la vez que se interviene quirúrgicamente al paciente.

La figura 1.5 muestra algunos ejemplos comunes de imágenes obtenidas por resonancia magnética nuclear, si bien insistimos en que mediante el ajuste de las secuencias de excitación magnética se puede obtener una amplísima gama de resultados.

- **Autoradiografía**



**Figura 1.5:** Imágenes de resonancia magnética: imágenes axiales T2-*weighted* (a) y T1-*weighted* (b); corte axial de un estudio fMRI donde los potenciales evocados se representan mediante un código de colores (c); corte de un volumen *tagged* MRI cardíaco (d); corte axial de un estudio MRI DT (e) y representación mediante un código de colores de la orientación de la difusión (f).

Se entiende por *autoradiografía* la detección de la radiación emitida por un isótopo radioactivo introducido en el paciente. Así, asociando un radioisótopo a ciertas moléculas que se incorporan al metabolismo se puede efectuar un seguimiento selectivo de diversos procesos metabólicos. Las imágenes o secuencias de imágenes obtenidas son representaciones bidimensionales de la concentración del radiofármaco en el organismo, siendo a menudo más importante la localización de esta actividad metabólica que la resolución o el detalle con el que se captura la anatomía, que en general es muy bajo.

#### ■ SPECT

Este acrónimo se emplea para referirse a la *tomografía computarizada de emisión de fotón único*, una técnica similar a la autoradiografía en la que la cámara detectora gira entorno al paciente, de modo que se puede construir una imagen tomográfica de la distribución del radiofármaco en el paciente.

Mediante esta técnica se obtienen dos volúmenes de datos distintos: el volumen de *transmisión* (figura 1.6.a) permite estimar la densidad de los tejidos, siendo esta estimación un requisito para la construcción del volumen de *emisión* (figura 1.6.b), que muestra la distribución del radiofármaco en el paciente.

Si bien en la mayoría de aplicaciones se ignora el volumen de transmisión, más adelante en esta tesis se demostrará el interés para nuestros propósitos de recuperar esta información.

#### ■ PET

La *tomografía de emisión de positrones* es una técnica de generación de imágenes tomográficas basada en la detección de la radiación producida por un radiofármaco introducido en el paciente, de forma similar a la autoradiografía y la SPECT. En el caso de la PET, sin embargo, se emplean radioisótopos con un tiempo de semivida sumamente corto. En la desintegración se generan positrones, que al cancelarse con los electrones presentes en el cuerpo generan la radiación gamma que será detectada.

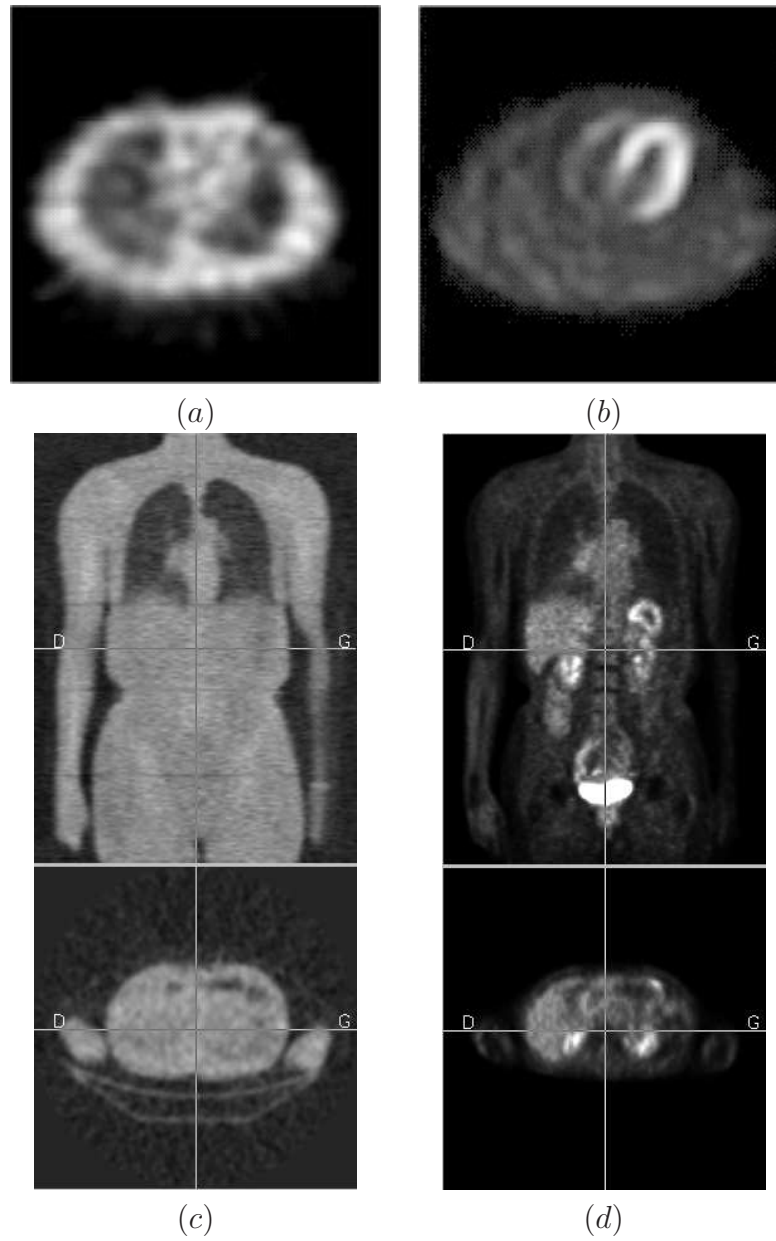
El uso de isótopos de vida tan corta permite conseguir imágenes de mejor calidad, y posibilita el estudio de procesos metabólicos demasiado rápidos para ser observados con SPECT. Sin embargo, la dosis de radiación que el paciente recibe es muy superior, unas dos veces la administrada en una radiografía.

Otro problema radica en la producción de los radioisótopos, que precisa de un ciclotrón muy próximo al lugar donde se efectuará el análisis, dado que la rápida desintegración de los isótopos complica su transporte y hace imposible su almacenamiento.

Al igual que en el caso de la SPECT, esta modalidad genera dos volúmenes de datos: un volumen de transmisión (figura 1.6.c) y otro de emisión (figura 1.6.d).

#### ■ Ultrasonidos





**Figura 1.6:** Imágenes de medicina nuclear: corte coronal de un volumen SPECT de transmisión (a) y emisión (b); cortes coronal y axial de un volumen PET de transmisión (c) y emisión (d).

Las imágenes de *ultrasonidos (US)* se consiguen mediante la aplicación sobre el paciente de ondas de presión a alta frecuencia. A partir de los ecos producidos por estas ondas se puede construir una representación bidimensional denominada sonograma, que refleja las discontinuidades de impedancia acústica, permitiendo estudiar el interior de tejidos blandos y cavidades.

Las principales ventajas de los ultrasonidos son su bajo coste, su naturaleza no invasiva, y la posibilidad de visualizar las imágenes a medida que se realiza la exploración, permitiendo un manejo interactivo. Como inconvenientes cabe citar una mala relación de señal a ruido y la existencia de un compromiso entre resolución y profundidad de penetración.

Una variante muy extendida de los ultrasonidos es la basada en el efecto *Doppler*, que permite visualizar la velocidad de los tejidos sometidos a examen, siendo empleada en estudios del movimiento cardíaco y del flujo sanguíneo.

Otra técnica derivada de los ultrasonidos que ha estado cobrando importancia en los últimos años es la *holografía de ultrasonidos (3D US)*, esto es, la reconstrucción de volúmenes tridimensionales de datos a partir de múltiples vistas bidimensionales de las que se conoce la posición del sensor.

En la figura 1.7 se pueden ver algunos ejemplos de imágenes ultrasonográficas de buena calidad. Es preciso aclarar que la imagen US típica está sumamente degradada por el ruido impulsional y artefactos que son inherentes a esta modalidad, y corresponde por tanto al usuario manipular los parámetros de la adquisición hasta obtener una visualización satisfactoria.

#### ■ Vídeo

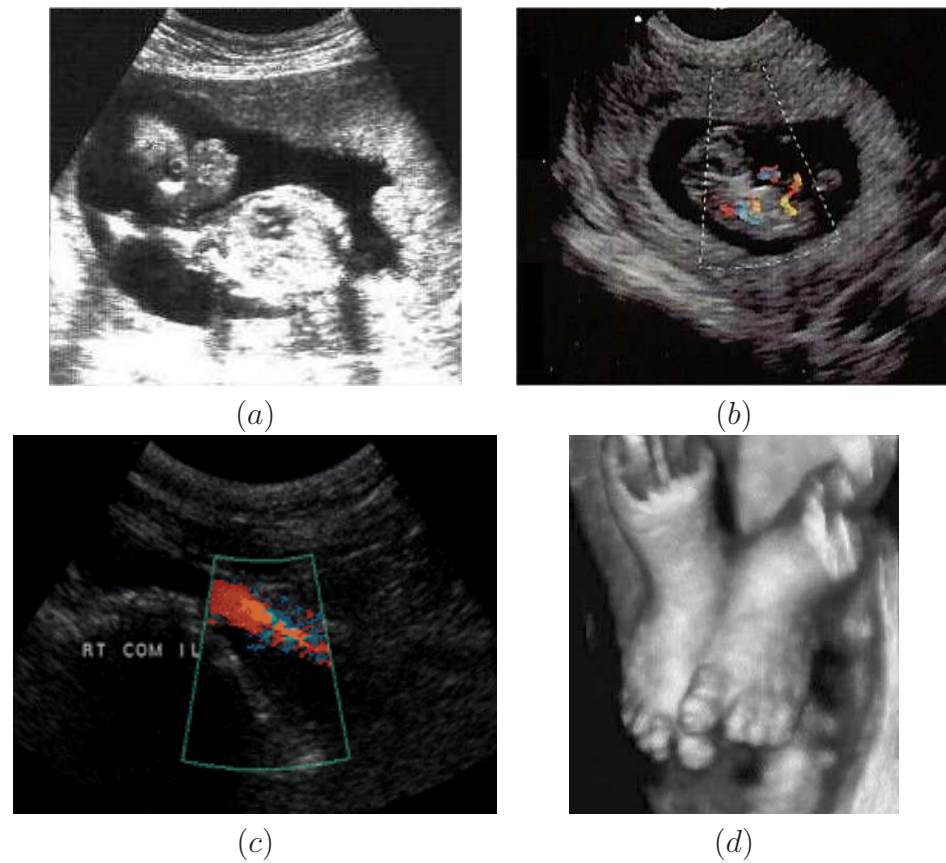
Las secuencias de imágenes ópticas convencionales también son utilizadas ocasionalmente en la práctica médica. Si bien no son tan comunes como algunas de las técnicas citadas anteriormente, puesto que en general aportan poco a la percepción del personal médico, hay dos excepciones que merece la pena mencionar: las exploraciones endoscópicas (*laparoscopia, artroscopia,...*) y los sistemas de visión para la intervención asistida por ordenador.

La figura 1.8 muestra algunos ejemplos de este tipo de imágenes. Dado el enorme impacto que ha tenido en el ámbito del tratamiento de imágenes médicas, merece una mención especial el *Visible Human Project*, de la *National Library of medicine*. Este proyecto pone a la disposición de la comunidad científica un completo atlas anatómico, del que se muestra un corte en la figura 1.8.a.

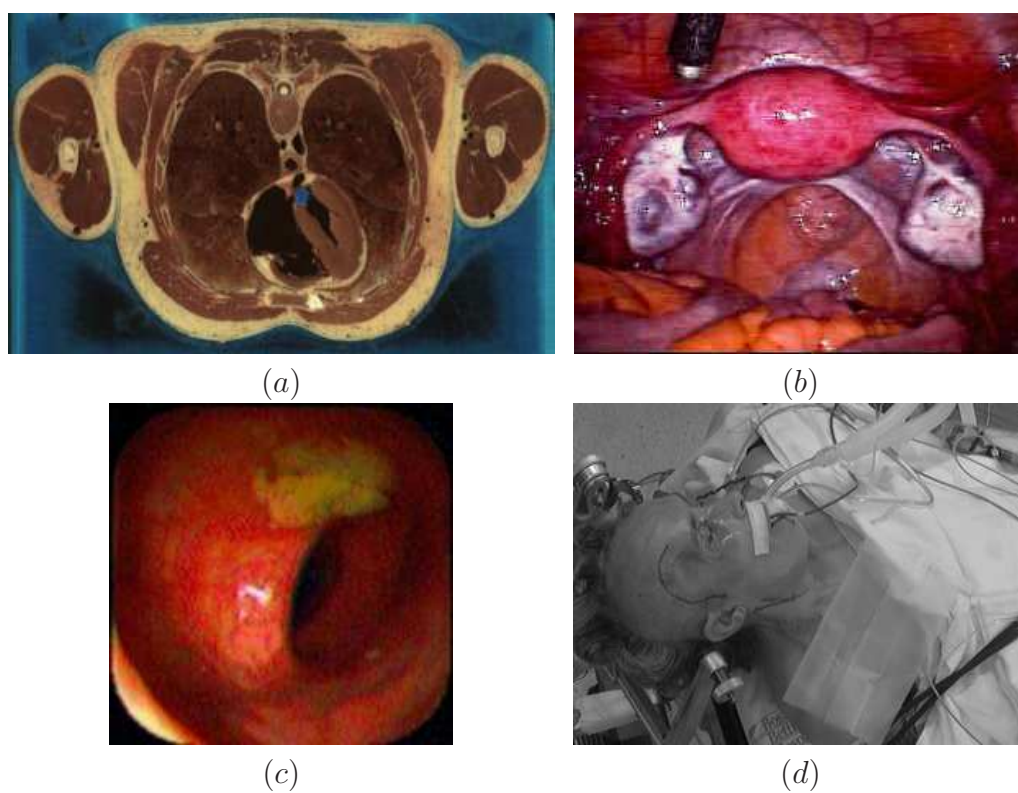
#### ■ Electroencefalografía

Si bien la *electroencefalografía (EEG)* no es exactamente una técnica de adquisición de imágenes, sino que brinda información de la actividad bioeléctrica cerebral del paciente en un conjunto de posiciones dispersas, se incluye en esta lista por haberse dado casos de aplicaciones en las que interesa registrar dicha información con imágenes de otras modalidades.





**Figura 1.7:** Imágenes de ultrasonidos: ultrasonograma de un feto de 11 semanas *(a)*; imagen de ultrasonografía Doppler en tiempo real que muestra el flujo sanguíneo en el feto *(b)*; angiografía Doppler que muestra una estenosis en la zona ilíaca *(c)*; reconstrucción de los pies de un feto a partir de un ultrasonograma 3D *(d)*.



**Figura 1.8:** Imágenes médicas de vídeo: Imagen de un corte axial del tórax *(a)*; vista laparoscópica de la pelvis *(b)*; colonoscopia que muestra el íleo terminal *(c)*; imagen intraoperatoria para una intervención de cirugía asistida *(d)*.

- Otras modalidades

Existen numerosas modalidades de adquisición de imágenes médicas que no se detallarán en este estudio, bien por estar aún en fase experimental, bien por ser de aplicación menos frecuente, bien por no emplearse en combinación con técnicas de registro de imágenes.

Citamos a continuación algunas de las más importantes, como pueden ser la *estimulación magnética transcraneal (TMS)*, la *elastografía de resonancia magnética (MRE)*, la *visualización de fuentes eléctricas (ESI)* (que generaliza la *electroencefalografía (EEG)*, la *electrocardiografía (ECG)* y la *vectorcardiografía (VCG)*), la visualización de fuentes magnéticas (MSI) (que incluye la *magnetoencefalografía (MEG)* y la *magnetocardiografía (MCG)*), la *resonancia de spin electrónico (ESR)*, la *tomografía por impedancia eléctrica (EIT)*, la *tomografía por microondas (MT)*, *visualización térmica por infrarrojos (TIR)* y la *microscopía confocal (CM)*.

La figura 1.9 muestra algunos ejemplos de estas modalidades. Nótese la significativa diferencia de calidad que presentan con respecto a las modalidades más comunes citadas anteriormente.

### 1.2.2. Aplicaciones del registro en medicina y biología.

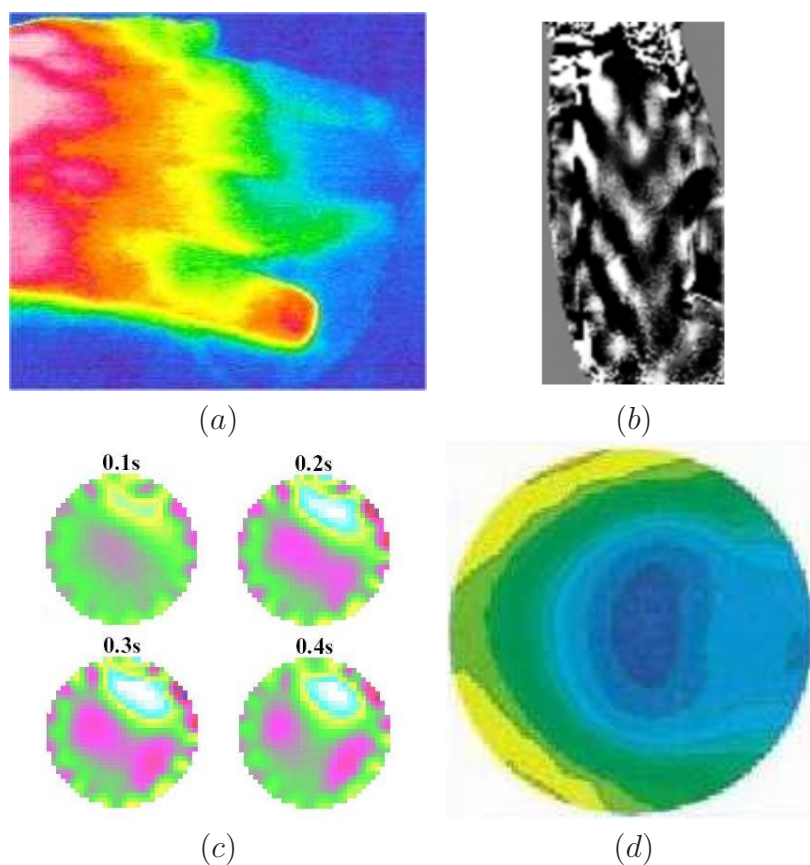
El registro de imágenes es un componente vital en todo el proceso médico, desde el diagnóstico, pasando por la planificación hasta la intervención y evaluación de los resultados. A continuación se describen distintos campos de aplicación dentro del ámbito de la medicina que precisan de la integración de múltiples paquetes de datos.

- Detección de tumores y localización de patologías.

A menudo no es posible capturar con una sola modalidad de imagen toda la información necesaria para localizar, diagnosticar o tratar una determinada patología. Las imágenes médicas que muestran la actividad metabólica del paciente ofrecen en general muy poca información sobre el soporte anatómico de la misma, e inversamente aquellas que permiten capturar con detalle información anatómica del paciente, a menudo no permiten detectar cierto tipo de lesiones. Es por tanto evidente el beneficio de combinar distintas modalidades de imágenes que contengan información complementaria.

- Cirugía asistida por computador.

La cirugía asistida es un campo que puede beneficiarse enormemente de la incorporación de técnicas de registro que permitan emplear datos preoperatorios durante la intervención. Ello permite aumentar la especificidad de la intervención, reduciendo el riesgo de dañar accidentalmente el tejido sano. Por otra parte, los requisitos de rapidez, precisión y seguridad para dichas técnicas son enormemente restrictivos, cosa que ha limitado en cierta medida su utilización.



**Figura 1.9:** Ejemplos de otras modalidades de imágenes médicas: Imágen térmica por infrarrojos de la pérdida de circulación sanguínea en los dedos debida a la enfermedad de Reynaud (*a*); corte de una elastografía de resonancia magnética del bíceps (*b*); secuencia de imágenes de tomografía de impedancia eléctrica en la que se muestran el corazón y los pulmones (*c*); permitividad relativa de una muestra ex-vivo de tejido mamario obtenida mediante tomografía de microondas (*d*).

- Investigación biomédica.

Dentro del campo de la investigación en biología y medicina son también numerosas las aplicaciones que precisan fusionar datos provenientes de distintos sensores o alinear secuencias temporales de imágenes. Algunos ejemplos podrían ser la clasificación de imágenes microscópicas (células sanguíneas, raspado cervical,...) y de electroforesis. Otro campo de especial importancia es la microscopía electrónica, en cuyo caso se usan técnicas de registro para construir volúmenes de datos apilando múltiples imágenes bidimensionales.

La figura 1.10 muestra algunos ejemplos de aplicaciones médicas en las que se emplea el registro de imágenes.

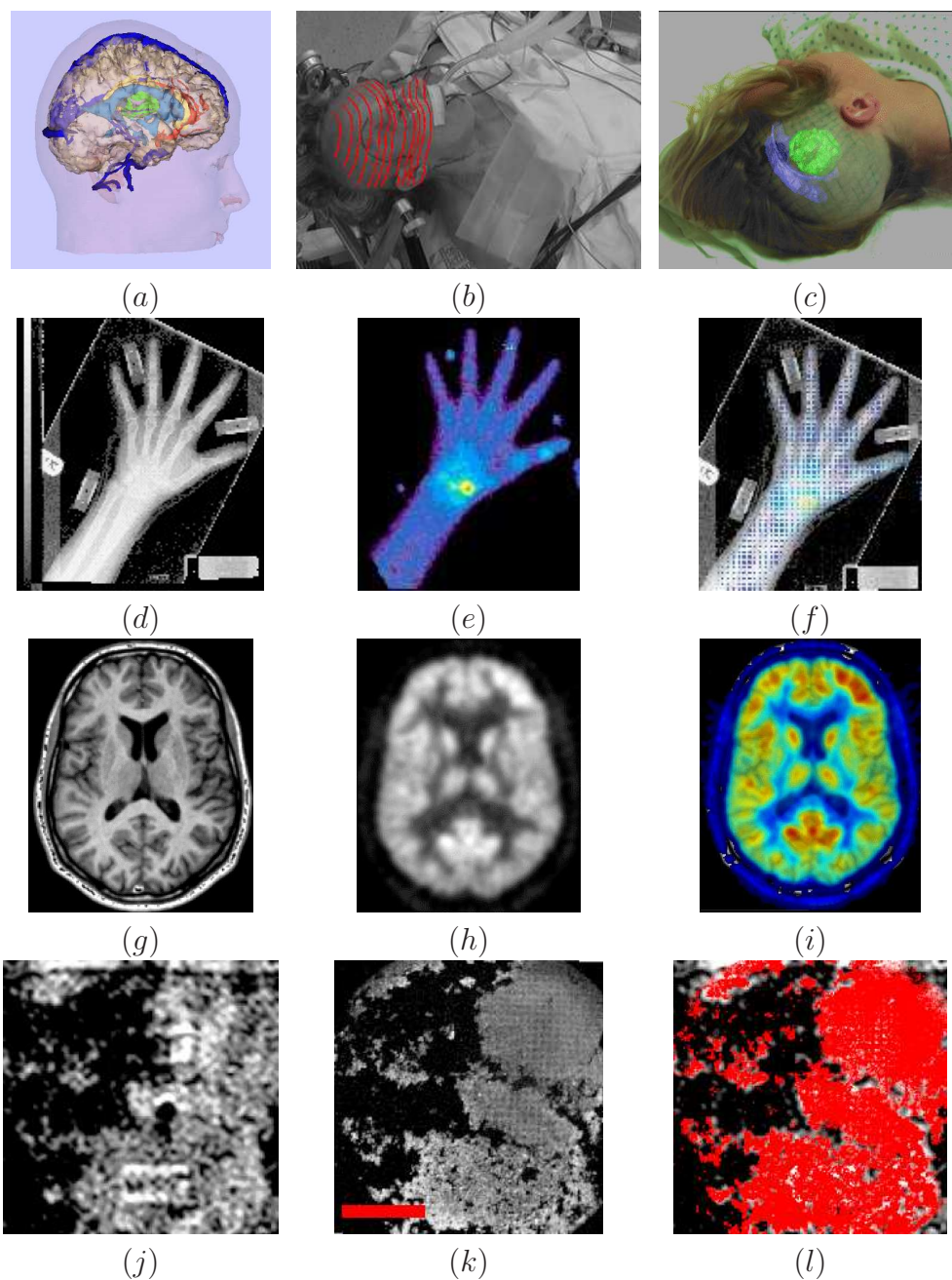
El primer ejemplo corresponde a un entorno de realidad aumentada para la neurocirugía, en el que un atlas anatómico generado a partir de imágenes preoperatorias es registrado con el paciente para facilitar la localización de las estructuras de interés durante la intervención.

A continuación se muestra un ejemplo de registro de imágenes anatómicas y funcionales (radiografía y medicina nuclear, respectivamente), destinada a localizar con precisión una cierta patología, en este caso una fractura. En los próximos capítulos se describirá con mayor detalle el interés del registro de imágenes de naturaleza anatómica con imágenes de naturaleza funcional, dada su gran importancia en el ámbito de la oncología cubierto por este trabajo de tesis.

El tercer ejemplo corresponde también a un registro anatómico-funcional, en este caso de resonancia magnética con medicina nuclear. El objetivo del registro es aquí la localización de la actividad cerebral anormal inducida por el consumo de drogas.

Finalmente, una aplicación de registro poco convencional: el registro de imágenes microscópicas para estudios histológicos.





**Figura 1.10:** Registro en medicina: Registro de estructuras anatómicas segmentadas de un volumen MRI (a) y mediciones laser (b) para generar un entorno de realidad aumentada para aplicaciones de cirugía asistida (c); registro de imágenes de radiografía (d) y SPECT (e) que permite detectar una fractura de escafoide (f); datos de MRI (g) y PET (h) fusionados para un estudio del efecto sobre el cerebro del abuso de drogas (i); registro de imágenes de microscopía óptica (j) y de resonancia magnética (k) para un estudio celular (l).

# Capítulo 2

## Métodos de registro

En este capítulo se presenta el esquema genérico de un sistema de registro de imágenes, describiéndose cada una de sus partes desde un punto de vista funcional.

En las secciones sucesivas se describirá con mayor profundidad cada una de dichas partes, detallándose las dificultades técnicas asociadas a las mismas y las soluciones más comunes en función de las particularidades de cada aplicación.

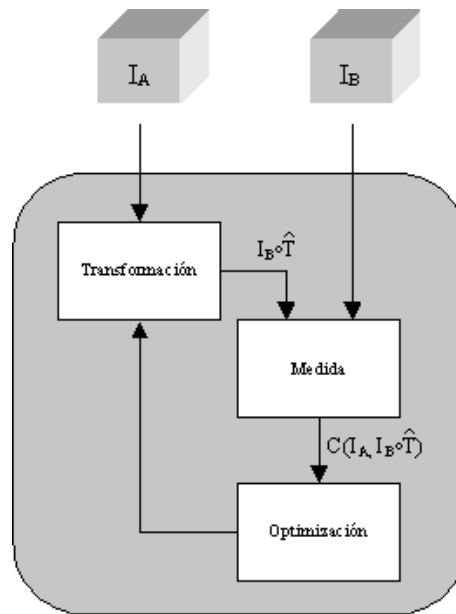
### 2.1. Modelo de sistema de registro

Un proceso de registro de dos imágenes consiste en aplicar una transformación geométrica a una de ellas de forma que se haga coincidir las posiciones de los pixels de una y otra imagen que corresponden a un mismo elemento de la escena. Con tal fin, un algoritmo de optimización se encarga de ajustar los parámetros de dicha transformación a fin de maximizar una función de similitud, en base a algún criterio establecido, entre la imagen de referencia y la imagen transformada. La figura 2.1 muestra el esquema de un sistema típico para el registro de dos imágenes,  $I_A$  e  $I_B$ . Este mismo esquema puede expresarse mediante la ecuación:

$$\hat{T} = \operatorname{argmin}_{T \in E} C(I_A, I_B \circ T), \quad (2.1)$$

en que se busca dentro de un determinado espacio de transformaciones  $E$  la transformación  $\hat{T}$  que minimiza un criterio  $C(I_A, I_B)$ . Donde  $I_A$  es la imagen estática de referencia y  $I_B$  la imagen que deseamos poner en correspondencia con la primera. La operación  $I_B \circ T$  indica que la imagen  $I_B$  ha sido afectada por la transformación geométrica  $T$ .

En función de la particularización de cada uno de estos componentes (a saber, transformación geométrica, criterio de similitud y método de optimización), las prestaciones del sistema varían en gran medida. Es preciso pues considerar detenidamente cada problemática individual antes de abordar el diseño de un sistema de registro de imágenes.



**Figura 2.1:** Sistema genérico de registro de imágenes.

A causa del gran número de posibles formas de enfocar el problema del registro se han definido distintas clasificaciones que permiten seleccionar el procedimiento que mejor se ajuste a las necesidades de cada aplicación. Dado que el eje central de este trabajo de tesis es el registro de imágenes médicas y su aplicación en oncología, en adelante nos centraremos en los métodos que son de aplicación en dicho campo, si bien en su mayor parte son generalizables a otros ámbitos.

A continuación se presenta una clasificación de los métodos de registro de imágenes médicas basada en las pautas propuestas por Van den Elsen(1993) y adoptado por Maintz(1998) [1].

Dichas pautas constituyen una importante referencia a la hora de seleccionar o desarrollar uno de estos métodos, permitiendo sopesar las necesidades de la aplicación y formar una idea clara de las opciones disponibles para construir el sistema de registro.

Se consideran en este caso ocho criterios a la hora de clasificar un sistema de puesta en correspondencia de imágenes médicas:

- Dimensionalidad de las imágenes

Comprende tanto las dimensiones espaciales de los datos a tratar como su dimensión temporal, caso de tratarse de secuencias de imágenes. De especial interés son los casos en que dicha dimensionalidad no coincide en una y otra imagen.

- Tipo de transformación

Considera la transformación geométrica empleada para desplazar la imagen



objetivo hacia su posición de registro con la imagen de referencia. Los tipos de transformación más comunes son, en grado creciente de complejidad:

- Rígida
- Afín
- Proyectiva
- Elástica

■ Medida de similitud

Se entiende por medida de similitud al método empleado para determinar la bondad del registro de dos imágenes. En función de los datos que se emplean para realizar dicha medida, se consideran:

- Medidas no basada en las imágenes
- Medidas extrínsecas
- Medidas intrínsecas
  - basadas en referencias
  - basadas en segmentación
  - basadas en intensidades

■ Procedimiento de optimización

Considera el método empleado para hallar el valor de los parámetros de la transformación que maximiza la medida de similitud entre las imágenes

■ Modalidades de imagen implicadas

Las modalidades de las imágenes a registrar (CT, MRI, PET, ...) son otro factor a tener en cuenta durante el diseño de un método de registro, especialmente al seleccionar la medida de similitud que se debe emplear. Así, en función de si ambas imágenes pertenecen a una misma modalidad, a modalidades distintas, o se desea registrar una imagen médica con otro tipo de datos, distinguimos:

- Registro monomodal
- Registro multimodal
- Registro de una modalidad a un modelo anatómico
- Registro de una modalidad a la posición del paciente

■ Sujeto visualizado

Este criterio considera si ambas imágenes pertenecen a un mismo paciente u objeto visualizado, o si por el contrario se pretende establecer un registro entre sujetos distintos con algún tipo de rasgo en común. De este modo, distinguimos:

- Registro intrasujeto
- Registro intersujeto
- Registro de sujeto a atlas

En este último caso, se entiende por atlas anatómico cualquier conjunto de datos que contenga información sobre la anatomía típica de una determinada región, generalmente en forma de imágenes segmentadas y etiquetadas.

- Objeto visualizado

Dado que los métodos de registro de imágenes son, en general, específicos para cada aplicación, tiene sentido clasificarlos en función del objeto visualizado en las imágenes. En el caso del registro de imágenes médicas esto equivale a considerar el órgano o región anatómica sobre el que se trabaja.

- Necesidad de interacción

Finalmente, tanto durante el diseño como en la práctica clínica es importante considerar el grado de autonomía de un sistema. En efecto, la utilidad de un método dependerá en gran medida de la necesidad de intervención por parte del usuario, sea en la inicialización, guiado o validación de los resultados.

A continuación se analizan con mayor detalle cada uno de los mencionados criterios. Merecen especial atención las distintas medidas de similitud y tipos de transformación geométrica, pues su elección es posiblemente la etapa más importante del proceso de diseño de un método de registro de imágenes.

## 2.2. Dimensionalidad de las imágenes

La dimensión de las imágenes a registrar es un factor importante a tener en cuenta en el diseño de un método de registro, pues afecta de forma directa no solo al volumen de datos que se deberán manejar, y por tanto al tiempo de cálculo necesario, sino también al número de grados de libertad que tendrán las transformaciones.

El criterio de dimensionalidad de las imágenes puede referirse a dimensiones espaciales únicamente, o puede incluir también una dimensión temporal, como es el caso de las secuencias endoscópicas de vídeo.

Algunos ejemplos de aplicaciones sin dimensión temporal son:

- Registro 2D/2D

Fusión de imágenes intrínsecamente bidimensionales (X-ray, portal,...), creación de mosaicos 2D, alineamiento de cortes consecutivos de un volumen 3D,...

El registro de imágenes bidimensionales es en general mucho más simple que su contrapartida 3D, tanto por el menor número de parámetros y volumen de los

datos implicados en la optimización, como por la mayor facilidad para visualizar y evaluar los resultados.

La figura 2.2.a muestra un ejemplo de registro de imágenes 2D para una aplicación de separación electroforética de proteínas.

- Registro 2D/3D

Volumen preoperatorio con imagen 2D intraoperatoria, imagen 2D con atlas o datos espaciales,...

Este es un caso de especial interés, puesto que al no coincidir las dimensiones de las imágenes a registrar se complica considerablemente su registro. Cabe añadir que en el caso de aplicaciones intraoperatorias, las restricciones de tiempo de cálculo son extremadamente importantes.

En la figura 2.2.b se pueden ver las imágenes empleadas por un sistema de registro de datos fluoroscópicos con un modelo tridimensional de la pelvis.

- Registro 3D/3D

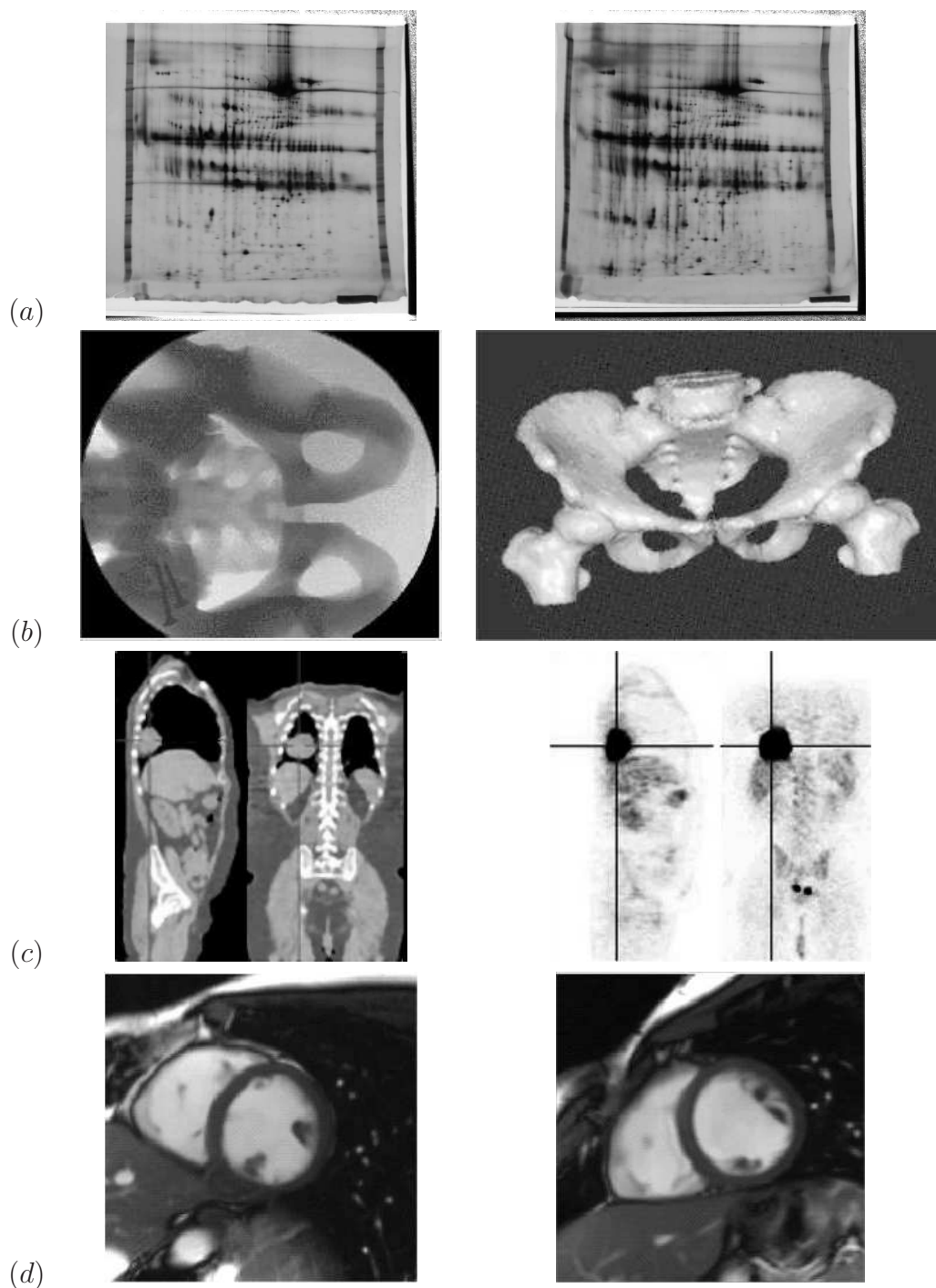
Registro de dos volúmenes tomográficos o de un volumen tomográfico con un conjunto de datos distribuidos en el espacio (EEG, puntero láser),...

Dada la extraordinaria proliferación en estos últimos años de distintas técnicas médicas de tomografía computarizada, este tipo de métodos hoy en día ocupan la mayor parte de trabajos de investigación. Esto se debe también al aumento de la disponibilidad de los medios informáticos precisos para gestionar el gran volumen de datos que comportan.

En la figura 2.2.c se muestran sendos cortes de dos volúmenes tomográficos CT y PET, provenientes de una aplicación de registro 3D/3D muy similar a la que se abordará en este trabajo de tesis.

Por otra parte, algunos ejemplos de aplicaciones con dimensión temporal serían la monitorización del crecimiento de tejido óseo en niños o pacientes sometidos a una operación de alargamiento de hueso, la monitorización del crecimiento de un tumor o su evolución tras el tratamiento, el seguimiento de la distribución en el flujo sanguíneo o respiratorio de un elemento marcador, estudios del movimiento cardíaco (mediante *tagged MRI*), etcétera.

Cabe decir que en muchos casos el registro de series temporales con imágenes estáticas se puede abordar como una sucesión de registros estáticos donde el resultado para una determinada imagen se emplea para inicializar el registro de la siguiente. Un ejemplo de esto último puede verse en la figura 2.2.d, en que se emplea el registro elástico de secuencias MRI para un estudio comparativo del movimiento cardíaco (en las imágenes se muestra el instante correspondiente a la diástole ventricular).



**Figura 2.2:** Ejemplos de aplicaciones de registro de distinta dimensionalidad: Par de imágenes 2D de electroforesis (a); Imagen fluoroscópica 2D y modelo tridimensional de la cadera (b); Volúmenes tomográficos CT y PET (c); Imágenes cardíacas de MRI correspondientes a un mismo instante de sendas secuencias temporales (d)

## 2.3. Transformaciones geométricas

Denominamos transformación geométrica a una correspondencia que asocia a cada punto de un espacio, un punto y sólo uno del mismo espacio.

El objetivo del registro de imágenes es hallar la transformación que pone en correspondencia los elementos equivalentes de una y otra imagen. Con tal fin, es preciso seleccionar una transformación paramétrica que permita compensar las diferencias entre las imágenes. Nótese que en este punto es preciso establecer un compromiso entre la precisión con que se podrá aproximar la solución deseada y el número de parámetros a ajustar. Dicha elección es de suma importancia en el diseño del sistema, y debe fundarse en un cuidadoso estudio de la naturaleza de las deformaciones entre las imágenes con que se deberá trabajar, así como de la magnitud de los errores que es posible aceptar en el resultado del registro. Al final de esta sección se detallan algunas de las transformaciones más comunes en registro de imágenes médicas, así como su número de grados de libertad.

Dentro de las transformaciones empleadas en aplicaciones de registro de imágenes pueden distinguirse aquellas *lineales*, que incluyen las transformaciones denominadas rígidas, afines y proyectivas, de las transformaciones *no lineales*, mucho más complejas y aún hoy en día campo de numerosos trabajos de investigación.

A continuación detallamos cada uno de dichos casos, centrándonos en los diversos aspectos de su parametrización.

### 2.3.1. Transformaciones lineales

Dentro del conjunto de las transformaciones lineales se incluyen las transformaciones rígidas (tan sólo incluyen traslaciones y rotaciones), las rígidas con reescalado isótropo o anisótropo, las afines (conservan la colinealidad y las relaciones de distancias) y las proyectivas (aplican rectas sobre otras rectas en un espacio de dimensión menor).

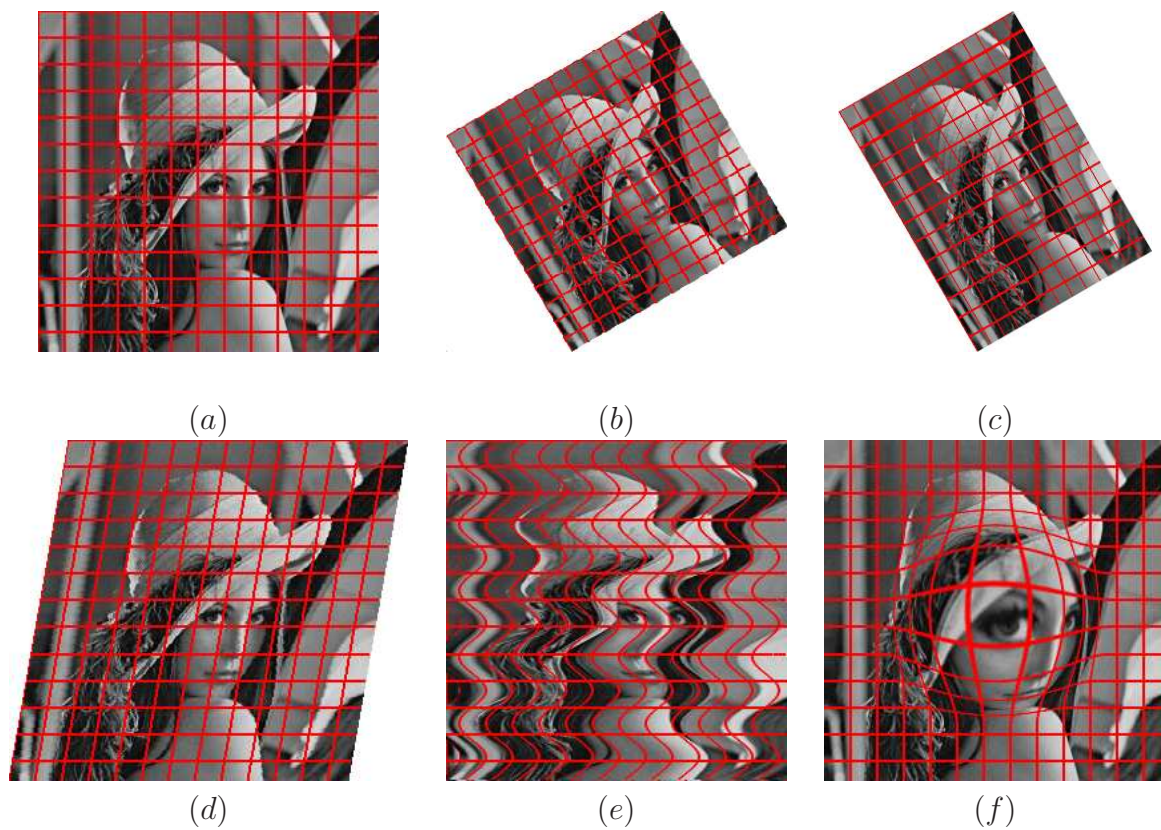
Este tipo de transformaciones pueden expresarse matricialmente de la siguiente forma:

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & b_0 \\ a_{10} & a_{11} & a_{12} & b_1 \\ a_{20} & a_{21} & a_{22} & b_2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (2.2)$$

donde  $(x, y, z)$  y  $(x', y', z')$  son respectivamente las coordenadas iniciales y transformadas de un elemento de la imagen. Esta expresión matricial puede comprender traslaciones, rotaciones, cambios de escala y corte (*shear*). A continuación se describen brevemente dichos efectos, así como su representación en la matriz de transformación.

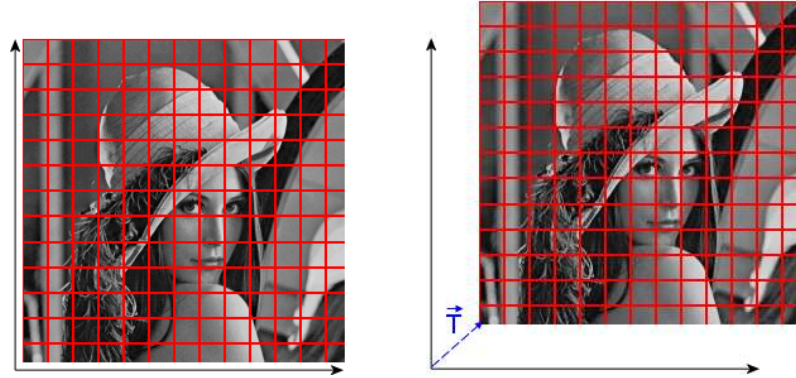
#### Traslaciones

Una traslación es una transformación consistente únicamente en un desplazamiento constante de todos los elementos de la imagen. En el espacio euclídeo tridimensional,



**Figura 2.3:** Tipos de transformaciones geométricas. Imagen original (a). Transformaciones lineales: Transformación rígida con reescalado isotrópico (b); transformación rígida con reescalado anisotrópico (c); transformación afín (d). Transformaciones no lineales: Transformación no lineal global (e); transformación no lineal local (f).





**Figura 2.4:** Ejemplo de traslación.

una traslación queda definida por los valores de dicho desplazamiento en cada eje,  $t_x, t_y, t_z$ , o, en notación matricial:

$$\begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{pmatrix} \quad (2.3)$$

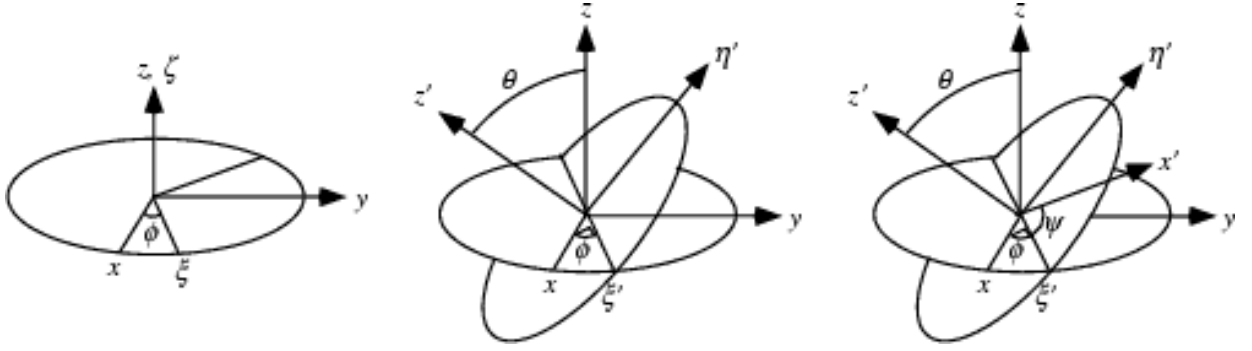
## Rotaciones

Existen distintas formas de expresar una rotación, que ofrecen prestaciones distintas a nivel de precisión numérica y facilidad de implementación. Las más comunes en aplicaciones de registro de imágenes son los ángulos de Euler, la expresión ángulo/eje y el cuaternión. Presentamos a continuación cada una de estas expresiones, resumiendo sus principales ventajas e inconvenientes.

### ■ Ángulos de Euler

El Teorema de Euler afirma que es posible expresar una rotación arbitraria mediante únicamente tres parámetros. En general se emplean con tal fin tres ángulos, denominados ángulos de Euler, que definen tres rotaciones consecutivas entorno a distintos ejes de coordenadas. Existen numerosas convenciones acerca de los ángulos de Euler, en función de los ejes sobre los que se efectúa cada rotación. La figura 2.5 ilustra paso a paso la composición de una rotación mediante rotaciones consecutivas de ángulos  $\phi$ ,  $\theta$  y  $\psi$  entorno cada eje.

La *x-convention* es una de las más comunes. Según ésta, la rotación definida por los ángulos  $(\phi, \theta, \psi)$  correspondería a una primera rotación de ángulo  $\phi$  entorno al eje  $z$ , una segunda rotación de ángulo  $\theta \in [0, \pi]$  entorno al nuevo eje  $x$  y una última rotación de ángulo  $\psi$  entorno al nuevo eje  $z$ .



**Figura 2.5:** Ejemplo de rotación mediante ángulos de Euler.

$$R_\phi = \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.4)$$

$$R_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix} \quad (2.5)$$

$$R_\psi = \begin{pmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.6)$$

$$R = R_\psi R_\theta R_\phi \quad (2.7)$$

Con lo que la matriz de rotación resulta:

$$\begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & \sin\theta\sin\phi & 0 \\ \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & -\sin\theta\cos\phi & 0 \\ \sin\psi\sin\theta & \cos\psi\sin\theta & \cos\theta & 0 \end{pmatrix} \quad (2.8)$$

Si bien este tipo de representación es conceptualmente sencilla, presenta un importante problema conocido como el *Gimbal lock*, un efecto que se da cuando dos ejes quedan alineados, perdiéndose el control sobre los mismos.

#### ■ Rotación angulo/eje

Otra posible parametrización de la rotación es mediante la definición de un cierto eje de rotación  $n = (n_x, n_y, n_z)$  y un ángulo  $\theta$  de rotación entorno a dicho eje. En este caso la *fórmula de Rodrigues* permite calcular de forma fácil la matriz de rotación:



$$R = Id + \sin\theta \cdot S_n + (1 - \cos\theta) \cdot S_n^2 \quad (2.9)$$

donde  $Id$  es la matriz identidad y  $S_n$  denota la matriz:

$$S_n = \begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{pmatrix} \quad (2.10)$$

Otra forma muy similar de expresar la rotación es mediante un único vector de rotación  $r = (r_x, r_y, r_z)$ , multiplicando el ángulo por el eje unitario sobre el que se desea efectuar el giro. En tal caso la fórmula de Rodrigues es:

$$R = Id + \frac{\sin\theta}{\theta} \cdot S_r + \frac{(1 - \cos\theta)}{\theta^2} \cdot S_r^2 \quad (2.11)$$

donde  $S_r$  se calcula de forma análoga a  $S_n$ .

En caso de trabajarse con ángulos pequeños, lo que ocurre con frecuencia en aplicaciones de registro de imágenes, se pueden aplicar las siguientes aproximaciones de Taylor:

$$\frac{\sin\theta}{\theta} = 1 - \frac{1}{6}\theta^2 + O(\theta^4) \quad (2.12)$$

$$\frac{(1 - \cos\theta)}{\theta^2} = \frac{1}{2} - \frac{1}{24}\theta^2 + O(\theta^4) \quad (2.13)$$

donde  $O(\theta^4)$  denota el orden del residuo de la aproximación.

El paso inverso de una matriz de rotación a su expresión ángulo/eje se puede realizar como sigue:

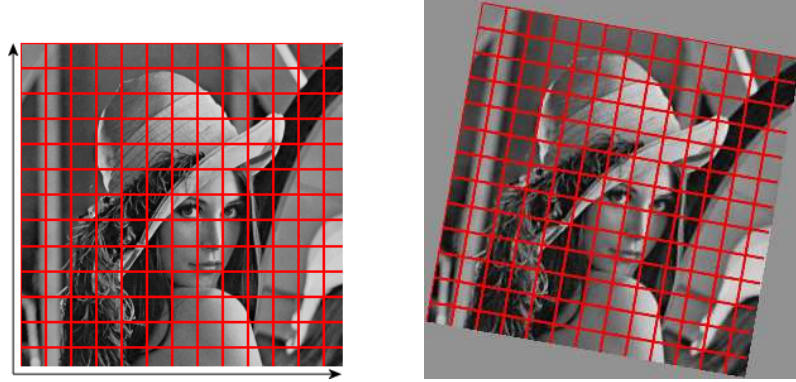
$$\begin{cases} \theta &= \arccos\left(\frac{\text{Tr}(R)-1}{2}\right) \\ S_n &= \frac{R-R^T}{s \cdot \sin\theta} \end{cases} \quad (2.14)$$

donde  $\text{Tr}(R)$  denota la traza de la matriz de rotación.

El principal inconveniente de este tipo de parametrización es que para ángulos próximos a  $\theta \approx 0$  y  $\theta \approx \pi$  el eje de rotación está mal definido y pueden darse problemas de precisión numérica.

#### ■ Cuaternión

En la parametrización mediante cuaterniones una rotación se describe mediante un vector de cuatro elementos  $q = (q_0, q_1, q_2, q_3)$ . Dado que según el teorema de Euler la rotación queda siempre definida por tres parámetros, los elementos del



**Figura 2.6:** Rotación de una imagen entorno a su punto central.

cuaternión son forzosamente dependientes, por lo que se incluye la restricción adicional  $|q|^2 = 1$ . De este modo, la matriz de rotación se puede expresar como:

$$\begin{pmatrix} q_0q_0 + q_1q_1 - q_2q_2 - q_3q_3 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_2q_1 + q_0q_3) & q_0q_0 - q_1q_1 + q_2q_2 - q_3q_3 & 2(q_2q_3 - q_0q_1) \\ 2(q_3q_1 - q_0q_2) & 2(q_3q_2 + q_0q_1) & q_0q_0 - q_1q_1 - q_2q_2 + q_3q_3 \end{pmatrix} \quad (2.15)$$

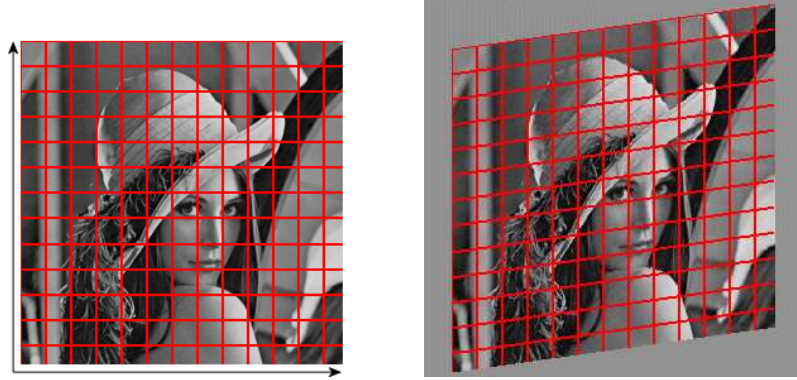
Se puede relacionar fácilmente el cuaternión con la notación ángulo-eje:

$$q = \begin{pmatrix} \cos(\theta/2) \\ n_x \sin(\theta/2) \\ n_y \sin(\theta/2) \\ n_z \sin(\theta/2) \end{pmatrix} \quad (2.16)$$

Una propiedad sumamente interesante de esta representación es que permite combinar fácilmente varias rotaciones. Así, dada una rotación  $R_3 = R_1 R_2$ , con  $q_1 = q_{1i}$  y  $q_2 = q_{2i}$ , la representación en cuaternión de la rotación compuesta es:

$$q_3 = \begin{pmatrix} q_{10}q_{20} - q_{11}q_{21} - q_{12}q_{22} - q_{13}q_{23} \\ q_{10}q_{21} + q_{11}q_{20} + q_{12}q_{23} - q_{13}q_{22} \\ q_{10}q_{22} + q_{12}q_{20} + q_{13}q_{21} - q_{11}q_{23} \\ q_{10}q_{23} + q_{13}q_{20} + q_{11}q_{22} - q_{12}q_{21} \end{pmatrix} \quad (2.17)$$

La parametrización de la rotación mediante cuaterniones es sin duda la más potente y robusta de las aquí descritas. No presenta singularidades como la representación ángulo/eje, y evalúa todos los parámetros simultáneamente, por lo que no se puede dar un *Gimbal lock*. Es sin embargo conceptualmente menos intuitiva que las anteriores, cosa que puede representar un problema en aplicaciones que requieran una interacción por parte del usuario.



**Figura 2.7:** Imagen afectada por un cierto factor de corte (*skew*).

### Corte

Se denomina corte (*shear*) a la transformación mediante la cual todos los puntos de un determinado plano  $P$  permanecen inmóviles, mientras que los puntos pertenecientes a otros planos se desplazan paralelamente a  $P$  una distancia, denominada *factor de corte*, proporcional a su distancia a  $P$ . En notación matricial, podemos expresar un corte en función de sus parámetros ( $s_{xy}, s_{yx}, s_{xz}, s_{zx}, s_{yz}, s_{zy}$ ) como:

$$\begin{pmatrix} 1 & s_{yx} & s_{zx} & 0 \\ s_{xy} & 1 & s_{zy} & 0 \\ s_{xz} & s_{yz} & 1 & 0 \end{pmatrix} \quad (2.18)$$

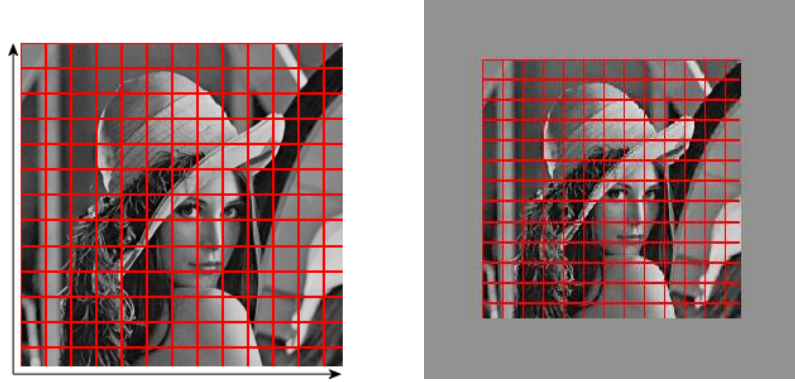
### Escala

Uno de los ajustes más habituales a realizar para establecer el registro de imágenes médicas de distintas modalidades es aumentar o reducir el tamaño de una de ellas por un factor ( $s_x, s_y, s_z$ ). Ello se consigue modificando los elementos de la diagonal de la matriz de transformación:

$$\begin{pmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \end{pmatrix} \quad (2.19)$$

Así pues, dada una imagen las dimensiones de cuyos voxels son  $I_{dx} \times I_{dy} \times I_{dz}$ , es posible reescalar dicha imagen para que sus elementos midan  $J_{dx} \times J_{dy} \times J_{dz}$ , mediante la matriz:

$$\begin{pmatrix} \frac{I_{dx}}{J_{dx}} & 0 & 0 & 0 \\ 0 & \frac{I_{dy}}{J_{dy}} & 0 & 0 \\ 0 & 0 & \frac{I_{dz}}{J_{dz}} & 0 \end{pmatrix} \quad (2.20)$$



**Figura 2.8:** Reescalado isotrópico de una imagen.

### Transformaciones proyectivas

Dada su importancia las transformaciones proyectivas merecen una mención especial. Se emplean en el registro de datos procedentes de volúmenes tridimensionales (por ejemplo un escáner CT) con imágenes planas (por ejemplo una radiografía).

En este caso es todavía posible emplear la notación matricial descrita, pero deben emplearse coordenadas homogéneas, con lo que el sistema toma la siguiente forma:

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & b_0 \\ a_{10} & a_{11} & a_{12} & b_1 \\ a_{20} & a_{21} & a_{22} & b_2 \\ a_{30} & a_{31} & a_{32} & b_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \\ w' \end{pmatrix} \quad (2.21)$$

Al añadir esta nueva variable  $w$ , podemos trabajar en el espacio proyectivo, un superconjunto del espacio real. De este modo, para  $w = 1$  las coordenadas homogéneas de un punto corresponden con sus coordenadas reales, mientras que para  $w = 0$  se obtiene la proyección de dicho punto en el infinito.

### Conclusión

Como se ha visto, las transformaciones lineales permiten compensar de forma simple y rápida una amplia variedad de diferencias de las imágenes. Ello se debe a una formulación matricial capaz de modelar efectos tales como rotaciones, traslaciones, cambios de escala, etc., hasta un total de doce grados de libertad.

Debemos insistir sin embargo en la importancia de restringir lo más posible dicho número de grados de libertad en función de las necesidades de la aplicación. La razón de esto es que con la transformación se define el espacio de búsqueda sobre el que deberá trabajar el algoritmo encargado de hallar los parámetros óptimos de registro, por lo que una transformación demasiado general dificulta y ralentiza la localización de dicho óptimo.

Las transformaciones lineales más comunes en registro de imágenes son:

- Transformación rígida:  
6 parámetros (3 traslaciones y 3 rotaciones)
- Transformación rígida + reescalado isótropo:  
7 parámetros (3 traslaciones, 3 rotaciones y 1 factor de escala)
- Transformación rígida + reescalado anisótropo:  
9 parámetros (3 traslaciones, 3 rotaciones y 3 factores de escala)
- Transformación afín:  
12 parámetros (3 traslaciones, 3 rotaciones, 3 factores de escala y 3 factores de corte)

Otros factores a tener en cuenta para facilitar la búsqueda de los parámetros óptimos de transformación son las cotas que se establecen para el espacio de búsqueda (por ejemplo, traslaciones entre  $[+10, -10]$ mm, rotaciones entre  $[+5, -5]^\circ$ ) y la precisión máxima que se puede conseguir en función de la resolución de las imágenes.

### 2.3.2. Transformaciones no lineales

Las deformaciones que no se ajustan a modelos tan simples como el rígido o afín, pueden conseguirse mediante una transformación no lineal. Se presenta a continuación una relación de los métodos más comúnmente empleados en registro de imágenes para conseguir dichas transformaciones. Éstos incluyen, entre otros, las transformaciones no lineales aplicadas globalmente a toda la imagen, la composición de deformaciones locales y la aplicación de modelos físicos:

- Transformaciones polinómicas

La generalización más inmediata de las transformaciones lineales descritas en el capítulo anterior es emplear polinomios de orden superior para conseguir deformaciones no lineales de las imágenes. Así, para el caso de emplear polinomios de segundo grado, se tiene la siguiente transformación con treinta grados de libertad (correspondientes a los parámetros de tres ecuaciones de segundo grado con tres incógnitas):

$$\begin{pmatrix} a_{00} & a_{01} & \dots & a_{09} \\ a_{10} & a_{11} & \dots & a_{19} \\ a_{20} & a_{21} & \dots & a_{29} \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} x^2 \\ y^2 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} \quad (2.22)$$

Este mismo esquema se puede emplear con polinomios de ordenes superiores para conseguir transformaciones con más grados de libertad. Este método, sin

embargo, no es muy utilizado, pues debido a la naturaleza global de las funciones polinómicas éstas no permiten modelar correctamente las deformaciones locales y tienden a provocar efectos oscilatorios indeseados.

- Funciones de base

Una alternativa a la formulación polinómica es expresar la transformación como una combinación lineal de otras funciones de base  $\theta_i(x, y, z)$ . La transformación puede escribirse como:

$$\begin{pmatrix} a_{00} & a_{01} & \dots & a_{0n} \\ a_{10} & a_{11} & \dots & a_{1n} \\ a_{20} & a_{21} & \dots & a_{2n} \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \theta_1(x, y, z) \\ \vdots \\ \theta_n(x, y, z) \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix}, \quad (2.23)$$

Las funciones empleadas a menudo forman una base ortonormal, como es el caso de bases de Fourier o wavelets.

- Splines de placa delgada

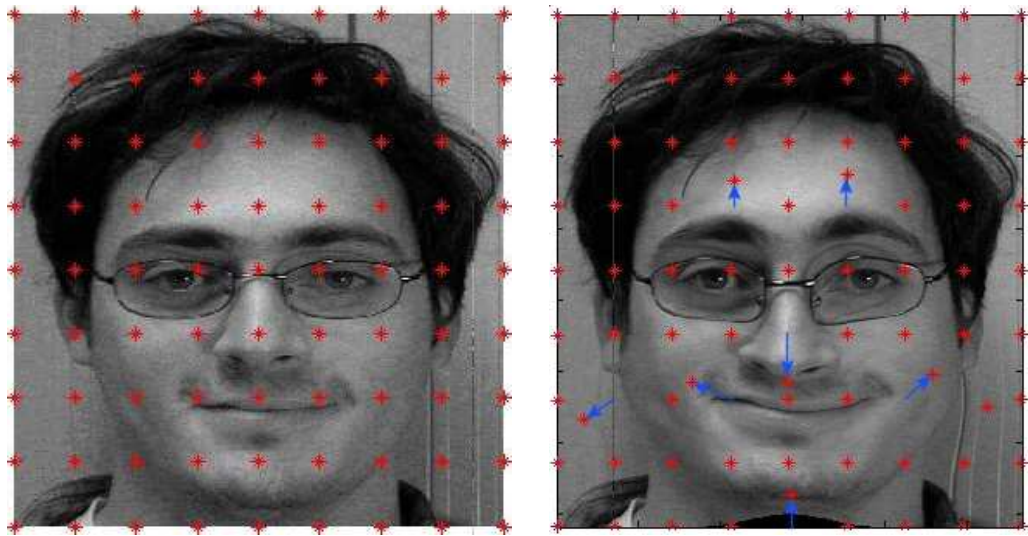
Este tipo de transformación (*thin plate splines*) es un caso particular del anterior que emplea splines como funciones de base. La transformación se expresa separadamente para las tres dimensiones mediante combinaciones lineales de splines monodimensionales  $\theta(s)$  de la forma:

$$t(x, y, z) = a_1 + a_2x + a_3y + a_4z + \sum_{j=1}^n b_j \theta(|\phi_j - (x, y, z)|), \quad (2.24)$$

donde los coeficientes  $a_i$  definen una transformación afín, mientras que los coeficientes  $b_j$  controlan la deformación no afín. La función de base empleada es generalmente  $\theta(s) = |s| \log(|s|)$  para transformaciones bidimensionales y  $\theta(s) = |s|$  en el caso tridimensional, si bien en algunos casos se han empleado Gaussianas u otras funciones apropiadas.

El cálculo de los coeficientes de la transformación precisa que se conozcan las posiciones en ambas imágenes de un conjunto de *puntos de control*. De este modo, forzando que la transformación haga corresponder dichos puntos (sea de forma exacta a través de una condición de interpolación o por aproximación mediante mínimos cuadrados), se consiguen fijar  $3n$  de los  $3(n+4)$  grados de libertad existentes. El resto de ecuaciones necesarias se obtienen forzando que los coeficientes  $b_i$  tengan media cero, y su producto escalar con las coordenadas  $x$ ,  $y$  y  $z$  de los datos sea también nulo.

- Deformaciones de forma libre



**Figura 2.9:** Ejemplo de FFD donde se aprecia como los desplazamientos de los puntos de control generan deformaciones locales en la imagen.

Las deformaciones de forma libre o *free form deformations* (FFD) son otro ejemplo de transformación no lineal construida como combinación lineal de funciones de base. En este caso, sin embargo, las deformaciones para cada eje se obtienen empleando dichas funciones de base para interpolar los desplazamientos definidos para una red regular de puntos de control. Así, el campo de desplazamientos generado se puede expresar como:

$$FFD(x, y, z) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 \theta_l(u) \theta_m(v) \theta_n(w) \phi_{i+l, j+m, k+n}, \quad (2.25)$$

donde  $u$ ,  $v$  y  $w$  indican la distancia relativa de la posición  $(x, y, z)$  a su correspondiente punto de control, de índices  $i$ ,  $j$  y  $k$ . Las funciones  $\theta_0$  a  $\theta_3$  componen una B-Spline cúbica monodimensional definida por tramos.

Las funciones de base más comúnmente empleadas en este caso son las B-splines, dado que al tener soporte finito permiten controlar localmente la deformación, a la vez que impiden que el coste computacional de calcular la transformación crezca exponencialmente con las dimensiones de la imagen.

La figura 2.9 muestra un ejemplo de deformación FFD de una imagen, donde la transformación no lineal se consigue gracias a un campo de desplazamiento creado por interpolación B-spline de los desplazamientos (representados por flechas) de una malla regular de puntos de control (representados por asteriscos).

- Modelos elásticos



Este tipo de transformaciones se basan en modelos físicos del comportamiento de materiales elásticos sometidos a fuerzas externas. El comportamiento de la deformación responde pues a la acción de dos términos: un conjunto de fuerzas externas aplicadas sobre la imagen, introducidas en nuestro caso para forzar el alineamiento de elementos correspondientes de las imágenes, y un conjunto de fuerzas internas de elasticidad, ejercidas por el propio material en respuesta a las anteriores. La transformación será pues aquella que defina un estado de equilibrio entre estas fuerzas, y puede calcularse como un campo de deformación  $u(x, y, z)$  solución de la ecuación diferencial lineal de Navier:

$$\mu \nabla^2 u(x, y, z) + (\lambda + \mu) \nabla (\nabla \cdot u(x, y, z)) + f(x, y, z) = 0, \quad (2.26)$$

donde  $f$  indica el conjunto de fuerzas externas actuando sobre el cuerpo elástico y  $\lambda$  y  $\mu$  son las constantes de elasticidad de Lamé. El operador *nabla*  $\nabla$  indica el gradiente del campo, y  $\nabla^2$  su laplaciana.

#### ■ Modelos fluidos

Las transformaciones de este tipo se basan también en leyes de la física, en este caso las que rigen el comportamiento de los fluidos. Estos modelos permiten compensar deformaciones locales muy severas, cosa que no es posible en el caso de los modelos elásticos ya que las energías internas lo impiden.

El comportamiento de un fluido viscoso se puede describir mediante la ecuación diferencial de Navier-Stokes:

$$\mu \nabla^2 v(x, y, z) + (\lambda + \mu) \nabla (\nabla \cdot v(x, y, z)) + f(x, y, z) = 0, \quad (2.27)$$

equivalente a la ecuación que describía la dinámica de un material elástico, excepto por el hecho de que aquí la incógnita es el campo de velocidades del fluido,  $v(x, y, z)$ .

El campo de deformaciones debe resolverse en cada iteración mediante la ecuación:

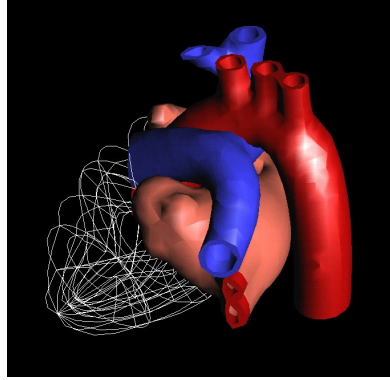
$$v(x, y, z, t) = \frac{\partial u(x, y, z, t)}{\partial t} + v(x, y, z, t) \cdot \nabla u(x, y, z, t). \quad (2.28)$$

Este esquema puede resolverse mediante sobrerelajación sucesiva (SOR), si bien se han propuesto implementaciones más rápidas para el caso de fluidos de viscosidad constante.

#### ■ Modelos de elementos finitos

Los modelos de elementos finitos (FEM) se han empleado también en la creación de transformaciones no lineales. Si bien estos métodos tienden a ser lentos y requieren de un cuidadoso mallado de las estructuras a deformar, son sumamente





**Figura 2.10:** Ejemplo de modelo cardíaco de elementos finitos.

útiles cuando se desea incorporar al modelo información sobre las propiedades físicas de los tejidos (rigidez, elasticidad, ...).

- Flujo óptico

El flujo óptico, un concepto empleado en visión por computador para el análisis del movimiento, es equivalente a las leyes físicas que describen el comportamiento de un fluido incompresible. En el caso de imágenes, esto se traduce en la hipótesis de que los valores de intensidad no varían al desplazarse. Esto se puede escribir para una imagen  $I$  como:

$$I(x, y, z, t) = I(x + dx, y + dy, z + dz, t + dt). \quad (2.29)$$

Si se ignoran los términos de orden superior en el desarrollo de esta expresión, esto puede reformularse como:

$$\frac{\partial I}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial I}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial I}{\partial z} \frac{\partial z}{\partial t} + \frac{\partial I}{\partial t} = 0. \quad (2.30)$$

Normalmente a esta ecuación del flujo óptico se le añaden ciertas restricciones para garantizar la suavidad de la solución obtenida.

## 2.4. Medidas de similitud

Todo método de registro de imágenes precisa de un criterio, explícito o no, que permita determinar la bondad de la correspondencia dado un cierto mapeo entre los espacios de las imágenes. Dicho criterio se expresa como una medida de la similitud entre las imágenes una vez puestas en el mismo espacio de referencia.

La elección de una medida de similitud es una de las decisiones más críticas en el desarrollo de un sistema de registro de imágenes, condicionando en gran medida

el coste computacional y la complejidad de la optimización requerida. Es en este punto, en definitiva donde se determina qué elementos de las imágenes se están poniendo en correspondencia y cuales no. No es extraño ver sistemas para los que se ha seleccionado una medida más general de lo estrictamente necesario, sin embargo no es esta una práctica aconsejable, pues la mejora en la precisión conseguida raramente compensa el impacto negativo en la robustez y tiempo de cálculo (debido a que la mayor complejidad de la medida a menudo acarrea una computación más lenta, y a medida que se prescinde de suposiciones sobre el problema aumenta la dificultad de su minimización). Otro factor a tener en cuenta es que cuanto más general es una medida menor aprovechamiento se hace de las restricciones implícitas del problema, con lo que la probabilidad de converger hacia un resultado erróneo aumenta significativamente [2].

Las medidas de similitud están íntimamente relacionadas con las características que se deseen emplear para establecer la correspondencia. Pueden basarse bien en los datos que se consiga extraer de las imágenes, bien en el conocimiento que se posee del sistema de adquisición, prescindiendo del contenido de las mismas. Dentro de las medidas basadas en las imágenes, hablamos de medidas extrínsecas cuando se toman como referencia ciertos elementos introducidos en la escena con tal fin. Caso de no introducirse elementos ajenos al objeto visualizado, se habla de medidas intrínsecas.

#### **2.4.1. Medidas no basada en el contenido de las imágenes**

Este tipo de registro se puede dar cuando se conoce el valor de aquellos parámetros de la adquisición que influyen en los efectos que se desea compensar.

De este modo las imágenes tomadas por una cámara fijada en el extremo de un brazo robótico pueden situarse fácilmente en un sistema de referencia común, sin reparar en su contenido, gracias a la información posicional proporcionada por dicho brazo.

Generalmente, sin embargo, estos métodos requieren que la relación entre los elementos sensores sea conocida, y suponen que el elemento visualizado permanece inmóvil. La difícil verificación de estas restricciones los hace inviables en la mayoría de aplicaciones prácticas.

Algunas aplicaciones que permiten este tipo de registro son aquellas en las que es posible fijar ambos elementos sensores entre sí, mediante una estructura rígida o un brazo robótico. También es posible referenciar de este modo la posición de las herramientas quirúrgicas con una imagen del paciente, siempre y cuando la relación entre el sistema de coordenadas de la imagen y el del instrumental pueda conocerse en todo momento (esto se consigue bien fijando el paciente a una estructura inmóvil y efectuando un seguimiento de las herramientas mediante elementos posicionadores o un brazo robotizado, bien efectuando un seguimiento de ambos).

#### **2.4.2. Medidas extrínsecas**

Como ya se ha dicho, este tipo de medidas se basan en comparar la posición en una y otra imagen de ciertos elementos de referencia introducidos artificialmente en la

escena con tal fin.

Puesto que los elementos de referencia son diseñados específicamente para resultar fácilmente localizables en las modalidades de imágenes implicadas, la etapa de registro para ponerlos en correspondencia resulta a menudo rápida y sencilla, siendo innecesario recurrir a costosos métodos iterativos. Hoy en día, este tipo de métodos son los más utilizados para aquellas aplicaciones que precisan de un registro en tiempo real, como es el caso de la cirugía asistida por computador.

Dado que existe un límite práctico al número de marcas que se pueden implantar, los grados de libertad de las transformaciones que se emplean están restringidos en consecuencia, siendo las rígidas o como mucho las afines las más comunes. Cabe destacar que también se han realizado con estas técnicas estudios que precisaban de un registro elástico, mediante la implantación de gran número de elementos marcadores, si bien dichos estudios se restringen habitualmente a animales.

El inconveniente principal de este tipo de métodos es la necesidad de prever la implantación de dichos elementos de referencia en ambas adquisiciones, que descarta la utilización de cualquier imagen no adquirida específicamente para esta aplicación. Asimismo, cualquier cambio en la posición de dichas referencias entre una y otra adquisición conllevará un error en el resultado final. Finalmente, destacar los problemas que conlleva en la práctica clínica mantener estas referencias inmóviles sobre un paciente cuando el periodo entre una y otra adquisición supera a menudo las veinticuatro horas.

Algunos de los elementos marcadores que han sido empleados en el registro de imágenes tomográficas son, por ejemplo, esferas de algún material fácilmente detectable (cristal, gelatina, cromo, vitamina E, ...), grapas o contenedores de diversas formas repletos de un líquido de contraste (sulfato de cobre, aceite, tecnecio 99m, una solución acuosa de gadolinio, ...).

Según la naturaleza de los elementos marcadores empleados se pueden clasificar estos métodos en invasivos y no invasivos, en función de si su implantación agrede o no al objeto visualizado, el tejido sano en el caso de aplicaciones médicas. La figura 2.11 muestra sendos ejemplos de este tipo de técnicas.

### Medidas no invasivas

Dentro del marco de los sistemas de registro de imágenes en aplicaciones médicas, este tipo de medidas se basan en elementos de referencia implantados sobre el paciente sin necesidad de dañar el tejido. Tal es el caso de estructuras estereotácticas, marcadores adhesivos, moldes de espuma personalizados y adaptadores dentales.

Este tipo de referencias adolecen sin embargo de una cierta falta de estabilidad debida a no emplazarse de forma solidaria a un soporte rígido (ello requeriría necesariamente de perforar la piel). Es por ello que en muchas aplicaciones se acepta el uso de elementos marcadores invasivos, especialmente cuando el traumatismo producido es pequeño en comparación con la intervención a realizar.



**Figura 2.11:** Ejemplos de elementos extrínsecos de referencia. Imagen intraoperatoria que muestra una estructura de referencia invasiva fijada a la columna (a). Detalle de la colocación de elementos marcadores no invasivos (b).

### Medidas invasivas

Sin duda los mejores resultados a nivel de precisión del registro, en particular para aplicaciones donde se trata de compensar desplazamientos rígidos, se consiguen mediante el uso de elementos de referencia implantados de forma invasiva en el paciente. Éste es el caso de marcadores y estructuras estereotácticas atornillables sobre el hueso (el cráneo, la espina dorsal o la cadera son emplazamientos típicos).

Si bien esto es aceptable para cierto tipo de aplicaciones, presenta importantes inconvenientes prácticos, especialmente cuando el intervalo entre una y otra adquisición es elevado.

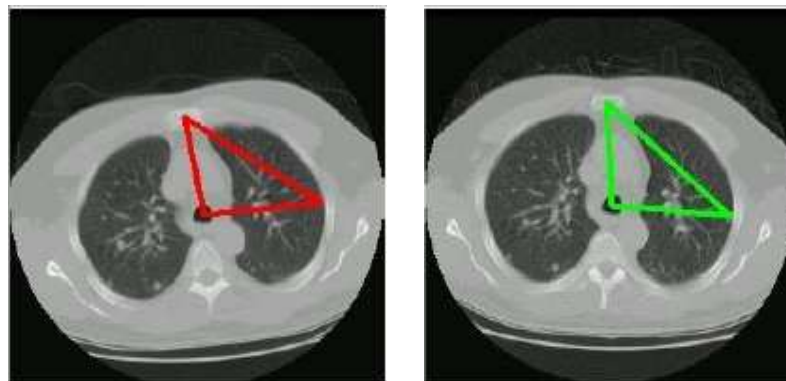
### 2.4.3. Medidas intrínsecas

Este tipo de medidas se basan en el contenido de las imágenes, sin que éstas tengan que haber sido específicamente preparadas para realizar el registro. Ello permite utilizar estos métodos de forma retrospectiva, sobre imágenes inicialmente no destinadas a tal fin, o sobre series de adquisiciones demasiado espaciadas en el tiempo para que sea posible incorporar elementos marcadores.

Las medidas intrínsecas pueden basarse bien en la localización de referencias naturales visibles en ambas imágenes, bien en la segmentación de estructuras características del objeto visualizado, o pueden incorporar el contenido íntegro de las imágenes para aprovechar al máximo la información presente en las mismas.

### Medidas basadas en referencias

Las medidas basadas en referencias naturales son equivalentes a las medidas extrínsecas basadas en elementos marcadores implantables descritas anteriormente, y



**Figura 2.12:** Ejemplo de registro basado en referencias anatómicas. En este caso se han empleado las centroides del esternón, la tráquea y un nódulo en el pulmón para establecer el registro de dos estudios CT de un mismo paciente.

comparten muchas de sus propiedades, como la rapidez y simplicidad de la consiguiente optimización. Su precisión queda sin embargo limitada por la capacidad que se tenga para detectar de forma exacta referencias correspondientes en una y otra imagen.

Las referencias que se utilizan en este tipo de medidas pueden ser tanto elementos anatómicos de interés (el centro de la retina, ramificaciones de vasos sanguíneos, un surco cerebral identificable, etc.), como singularidades geométricas (puntos de máxima curvatura, centros de ventanas con varianza local máxima, intersecciones de líneas, etc.).

En la figura 2.12 se puede ver un ejemplo de registro de cortes axiales CT del tórax basado en la localización del esternón y la tráquea.

Nótese que la detección de referencias no deja de ser una etapa de segmentación de las imágenes, a pesar de no haberse considerado como tal a efectos de esta clasificación. Reservamos el calificativo de segmentación para la extracción de estructuras de mayor complejidad, que requerirán en general de transformaciones y métodos de optimización mucho más sofisticados.

### Medidas basadas en segmentación

Este tipo de medidas se basan en una etapa previa destinada a detectar en ambas imágenes ciertas estructuras de interés sobre las que establecer la correspondencia. En consecuencia, estos métodos están fuertemente vinculados a la aplicación para la que han sido diseñados, siendo difícil su utilización en otro ámbito.

La etapa de segmentación encargada de localizar dichas estructuras de interés es sin duda la parte más crítica del sistema, dado que cualquier error introducido en este punto se propagará inevitablemente al resultado final. Esta etapa puede ser manual, semiautomática, supervisada o estar totalmente automatizada. Si bien la mayoría de sistemas empleados hoy en día en el ámbito de la medicina y la biología requieren de

un cierto grado de intervención por parte del usuario, los nuevos métodos tienden cada vez más hacia la segmentación automática.

Dentro de los métodos que emplean estas medidas, sin duda los más populares son los que se basan en *modelos rígidos* de las estructuras contenidas en las imágenes. Dichos modelos pueden consistir en puntos, curvas, superficies o volúmenes, si bien lo más común es emplear curvas para el caso de imágenes bidimensionales y superficies si se trabaja con datos tridimensionales. Una vez realizada la segmentación el proceso de registro trabaja exclusivamente sobre estructuras obtenidas, sin recurrir en ningún momento al resto de la información contenida en las imágenes.

Nótese que la denominación de modelo rígido se refiere exclusivamente al modelado que se hace de las estructuras segmentadas, y no debe confundirse con el tipo de transformación implicada en el registro, que muy bien puede establecer una correspondencia elástica entre dos estructuras de este tipo.

Una vez segmentadas las estructuras de interés, a partir de ellas se pueden definir distintas medidas de similitud. En la literatura se pueden encontrar desde medidas tan simples como la suma de diferencias absolutas entre los puntos más próximos [3], hasta medidas sumamente complejas que consideran distancias entre árboles o grafos [4] (para el caso de representaciones estructurales), o autómatas [5] (en el caso de representaciones sintácticas).

Los métodos basados en *modelos deformables*, por otra parte, emplean herramientas tales como contornos activos (*snakes*) y superficies activas (*nets*) para hacer evolucionar las estructuras segmentadas en una de las imágenes hacia su estructura correspondiente en la otra. Esto se puede conseguir bien trabajando sobre las imágenes segmentadas, o bien directamente empleando las estructuras obtenidas en una de las imágenes para a la vez segmentar y establecer el registro con su correspondiente en la otra imagen. Dichas herramientas se describen con mayor detalle en la sección 4.3.

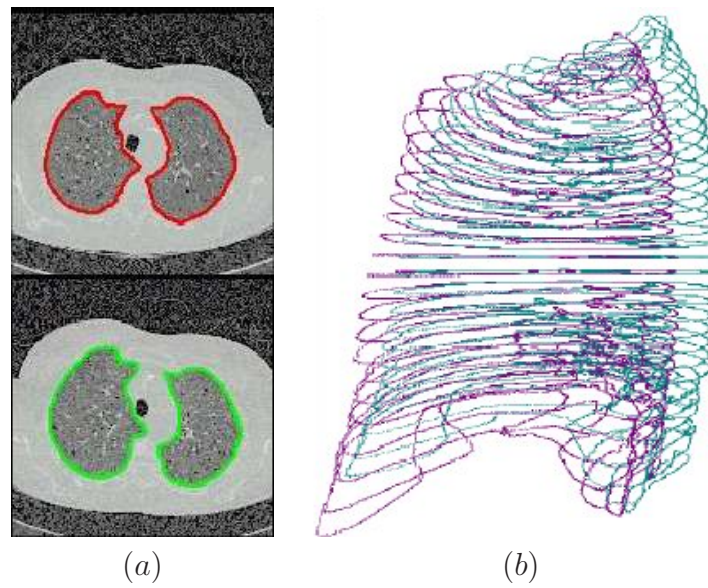
La figura 2.13 muestra un ejemplo de registro de volúmenes CT del tórax basado en el registro de modelos tridimensionales obtenidos de la segmentación de los pulmones en cada corte axial.

## Medidas basadas en intensidades

También conocidas como *medidas basadas en propiedades de los voxels*, este tipo de medidas se caracterizan por emplear el contenido íntegro de las imágenes sin una etapa previa de detección de referencias o de segmentación que reduzca el volumen de datos a manipular. Existen dos grandes grupos dentro de las medidas de este tipo, en función de cómo se aprovechan estos datos.

En primer lugar están los métodos que tratan de reducir la información contenida en las imágenes a un conjunto de valores característicos, tales como sus momentos y ejes principales. Si bien este tipo de métodos está restringido a aplicaciones de registro rígido y los resultados que se obtienen no son demasiado precisos, especialmente para el caso de imágenes multimodales, su uso está muy extendido debido a su fácil implementación, rapidez y al hecho de ser totalmente automáticos. En algunos casos se incluye una etapa





**Figura 2.13:** Ejemplo de registro basado en la segmentación. Contornos de los pulmones segmentados en los distintos cortes de dos estudios CT (a). Esquema de alambres del pulmón izquierdo segmentado en ambos estudios, durante el registro (b).

de presegmentación para mejorar el resultado final.

Por otra parte, las medidas basadas en el contenido íntegro de las imágenes ocupan el extremo opuesto del espectro, consiguiendo gran precisión y robustez frente al ruido y los artefactos, a cambio de un elevado coste computacional. Es en este campo donde se están desarrollando actualmente la mayor parte de trabajos de investigación en el campo del registro de imágenes. Ello se debe al extraordinario aumento de la potencia de cálculo en los últimos años, que ha permitido manejar de forma eficiente los volúmenes cada vez mayores de datos contenidos en una imagen tomográfica.

Se pueden agrupar las distintas medidas de similitud en función de las suposiciones que realizan sobre la naturaleza del problema de registro, más en concreto sobre la relación existente entre los niveles de intensidad que las imágenes asignan a un mismo tipo de tejido. Así, medidas como el error cuadrático suponen una relación lineal entre dichos niveles de intensidad, mientras que las medidas basadas en la correlación suponen tan sólo una relación funcional (esto es, que los tejidos representados por un cierto nivel de gris en una imagen son representados por un único nivel de gris en la otra). En el caso de registro de imágenes multimodales la suposición de una relación funcional resulta todavía demasiado restrictiva, siendo preciso emplear medidas que no precisen más que de una relación estadística entre los niveles de intensidad, como aquellas basadas en la entropía.

La figura 2.14 muestra el aspecto del histograma conjunto de intensidades entre una imagen de referencia y la misma imagen afectada por una transformación lineal, no lineal y no funcional respectivamente de sus niveles de gris.

La figura 2.15 ilustra el comportamiento de algunas medidas de similitud clásicas en función de la relación entre las intensidades de las imágenes. Se aprecia que cuanto más general es una medida a mayor número de casos permite adaptarse, pero es claramente superada por medidas más simples cuando el problema es suficientemente restringido.

Se describen a continuación algunas de las medidas de este tipo que han sido más extensamente empleadas. Esta no es, sin embargo, una relación extensiva, puesto que existen innumerables variaciones de cada una de ellas en función de las necesidades particulares de cada aplicación y hoy en día aún se trabaja para encontrar nuevas medidas o perfeccionar las existentes.

- Suma de diferencias cuadráticas (SSD)

Esta es una de las primeras medidas a las que se recurrió por su simplicidad. Puesto que el error cuadrático medio es óptimo cuando las imágenes difieren tan solo por un ruido gaussiano aditivo, esta medida es adecuada para el registro cuando no hay distorsiones locales.

$$\text{SSD} = \sum_{(i,j) \in \Omega} (I_A - I_B)^2, \quad (2.31)$$

donde  $\Omega$  indica la zona de solapamiento de las imágenes.

En ocasiones esta medida se pondera por el número de pixels en la zona de superposición entre las imágenes, en cuyo caso se denomina "diferencia cuadrática media de intensidades" (MSD).

- Suma de diferencias absolutas de intensidad (SAVD)

Esta medida es una simplificación de la anterior para conseguir una computación más eficiente, si bien conserva la mayor parte de sus propiedades.

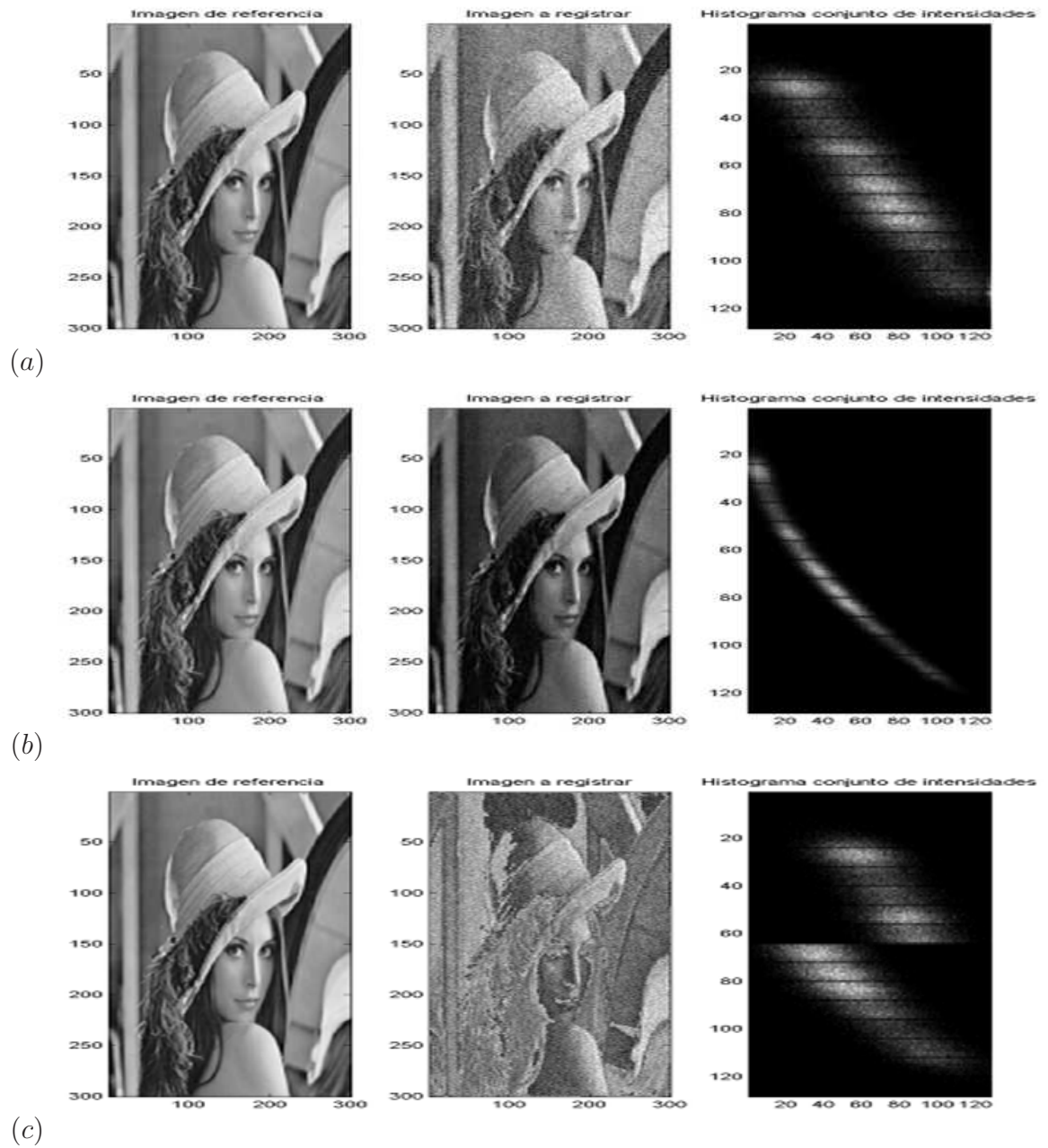
$$\text{SAVD} = \sum_{(i,j) \in \Omega} |I_A - I_B|, \quad (2.32)$$

- Intensidad de pautas (PI)

Esta medida, denominada *pattern intensity* en el artículo original [6], se basa en la suposición de que cuando dos imágenes se hallan correctamente alineadas, el número de estructuras visibles en la imagen diferencia  $I_{dif}$ , es mínimo. Este es el caso cuando se trabaja con imágenes entre cuyas intensidades hay una relación lineal.

Se considera a este fin que un pixel pertenece a una estructura cuando su intensidad es significativamente distinta de la de sus vecinos en un entorno de radio  $r$ . La medida responde así a la siguiente expresión:

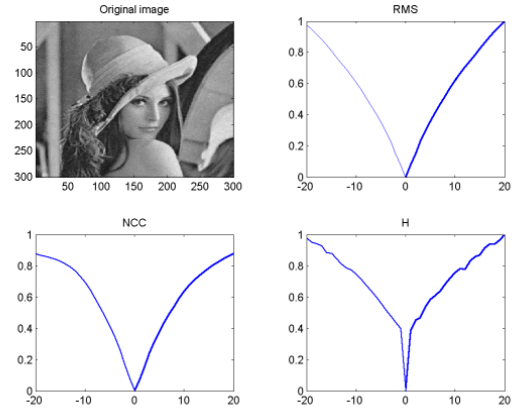




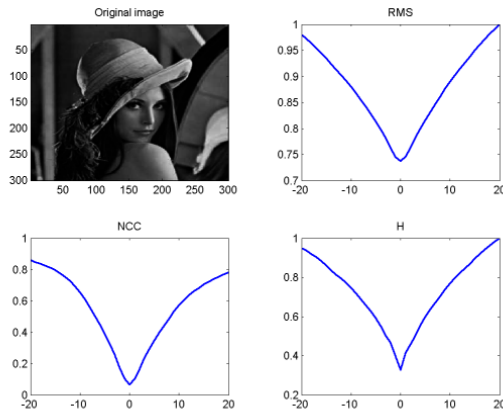
**Figura 2.14:** Ejemplos de histogramas conjuntos de intensidades: Relación funcional lineal (a); relación funcional no lineal (b); relación no funcional (c)



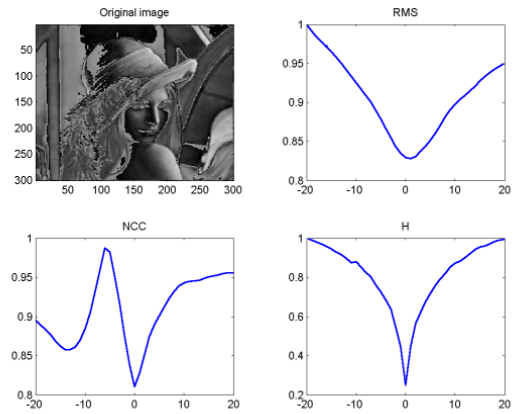
(a)



(b)



(c)



(d)

**Figura 2.15:** Perfil de distintas medidas de similitud en función del desalineamiento en un solo eje. Las medidas empleadas son, en grado de generalidad creciente, el error cuadrático medio (RMS), la correlación cruzada normalizada (NCC) y la entropía de la imagen diferencia (H). Se han empleado estas medidas para evaluar el registro de una imagen (a) con una versión de ella misma afectada por un ruido gaussiano (b), afectada por un operador no lineal (c) y afectada por un operador no funcional (d).

$$PI_{r,\sigma} = \sum_{(i,j) \in \Omega} \sum_{d^2 \leq r^2} \frac{\sigma^2}{\sigma^2 + (I_{dif}(i,j) - I_{dif}(v,w))^2}, \quad (2.33)$$

donde

$$d^2 = (i - v)^2 + (j - w)^2, \quad (v, w) \in \Omega. \quad (2.34)$$

El término constante  $\sigma$  controla la sensibilidad frente a las variaciones de intensidad, permitiendo que la medida se vea poco afectada por pequeñas desviaciones, como las debidas al ruido.

La imagen diferencia se calcula como:

$$I_{dif} = I_A - sI_B, \quad (2.35)$$

donde  $s$  es un factor de escala que puede a su vez precisar de una optimización.

Una ventaja de este método es su naturaleza local, que le permite ser insensible frente a diferencias de baja frecuencia espacial entre las imágenes (este era el caso, en la aplicación para la que fue diseñado este método, de las diferencias introducidas por el desplazamiento del tejido blando, frente a aquellas introducidas por el tejido óseo). Por otra parte, la naturaleza asintótica de esta medida hace que las grandes diferencias de intensidad afecten por igual al resultado, independientemente de su magnitud.

#### ■ Diferencia de gradientes

Con esta medida se trata de aprovechar la robustez del gradiente frente a las deformaciones de estructuras poco contrastadas, como por ejemplo el tejido blando en imágenes de radiografía, a la vez que se evita la sensibilidad de la correlación a las estructuras parásitas.

Se define la medida como:

$$G = \sum_{i,j \in \Omega} \frac{A_v}{A_v + (I_{difV}(i,j))^2} + \sum_{i,j \in \Omega} \frac{A_h}{A_h + (I_{difH}(i,j))^2}, \quad (2.36)$$

donde  $A_v$  y  $A_h$  son constantes que controlan el peso del término asociado al gradiente vertical y horizontal, respectivamente.

La imagen diferencia de gradientes se calcula como:

$$I_{difV}(i,j) = \frac{dI_A}{di} - s \frac{dI_B}{di}, \quad (2.37)$$

$$I_{difH}(i,j) = \frac{dI_A}{dj} - s \frac{dI_B}{dj}, \quad (2.38)$$

Esta medida posee la misma estructura  $1/(1+x^2)$  que la *pattern intensity*, que permite reducir la sensibilidad frente a estructuras parásitas (como por ejemplo un catéter).

- Correlación cruzada normalizada (NCC)

También denominada *Pearson product moment cross-correlation*, esta medida es óptima para el caso de imágenes que difieren tan solo en una transformación lineal de sus intensidades y un ruido gaussiano aditivo.

Se define como:

$$\text{NCC} = \frac{\sum_{(i,j) \in \Omega} (I_A(i,j) - \bar{I}_A)(I_B(i,j) - \bar{I}_B)}{\sqrt{\sum_{(i,j) \in \Omega} (I_A(i,j) - \bar{I}_A)^2} \sqrt{\sum_{(i,j) \in \Omega} (I_B(i,j) - \bar{I}_B)^2}}, \quad (2.39)$$

donde  $\bar{I}_A$  y  $\bar{I}_B$  son los valores de intensidad medios en la zona de superposición de las imágenes  $\Omega$ .

Esta medida es precisa para el caso de imágenes contaminadas con ruido blanco pero resulta poco robusta frente a distorsiones locales. Por otra parte, dado que la contribución de cada pixel depende fuertemente de su intensidad, unos pocos elementos con grandes diferencias de intensidad -tales como los introducidos por la aparición de un catéter en una de las imágenes-, afectarán severamente a la medida. Por último, es preciso tener en cuenta de cara a la optimización que la posición de correcta correspondencia presenta un pico estrecho en el espacio de correlación, cosa que dificulta su búsqueda.

- Correlación de gradientes

Esta medida consiste en calcular separadamente la correlación cruzada normalizada de las imágenes de gradientes horizontales y verticales que se obtienen por diferenciación mediante un filtrado de Sobel de las imágenes originales. El resultado final de la medida es el promedio de dichos valores de correlación.

El hecho de efectuar el registro sobre las imágenes de gradientes permite centrarse en la información contenida en los contornos importantes, reduciendo el impacto de estructuras poco definidas en las imágenes. Por esta razón, sin embargo, resulta especialmente agravada la sensibilidad de la correlación a la aparición de estructuras parásitas como instrumentos quirúrgicos.

- Algoritmo de Woods

Este método, propuesto por primera vez en [7], se basa en la idea de que, cuando las imágenes están correctamente alineadas, el rango de valores de intensidad de la imagen a registrar asociadas con un valor de intensidad de la imagen de referencia debe ser mínimo. Así, se minimiza la suma de las desviaciones típicas de los valores de intensidad de los pixels correspondientes a cada región de la imagen de referencia.

Nótese que esta medida está estrechamente relacionada con la entropía condicional del volumen a registrar, dado el volumen de referencia. Esta fue una de las primeras aproximaciones a las actuales medidas basadas en la teoría de la información.

El principal problema de este método es que precisa de una cierta intervención por parte del usuario, que debe definir las regiones de las imágenes en las que se basará el criterio.

Más adelante Hill et al [8] proponían una adaptación de dicho método para trabajar sobre el histograma conjunto de las imágenes. Si bien esto mejora en cierta medida el comportamiento del criterio, presenta aún el inconveniente de requerir la identificación de regiones de tejido similar en el histograma.

- Momentos de tercer orden del histograma conjunto [9]

Este método, desarrollado al observar que el comportamiento de la representación conjunta de los niveles de gris reflejaba el desalineamiento entre las imágenes, fue el primero en intentar extremar las propiedades de la señal conjunta. Con este fin se emplean los momentos de tercer orden (*skewness*) del histograma conjunto para caracterizar la agregación de los datos y su grado de asimetría. La definición más común de la *skewness*  $\gamma_1$  de una distribución es:

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}, \quad (2.40)$$

donde  $\mu_i$  indica el momento central de orden  $i$  de dicha distribución.

- Entropía del histograma conjunto (H)

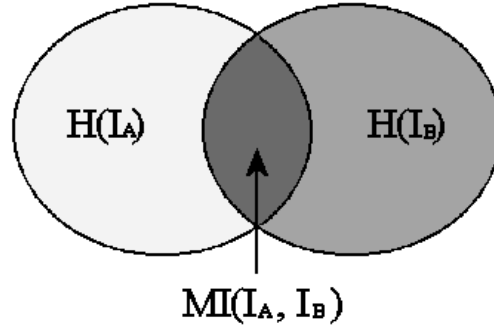
Este criterio propuesto por Collignon [10] y Studholme [11] fue el primer intento de utilizar la entropía como medida de la correspondencia. Dado que la entropía es adecuada para medir la dispersión del histograma conjunto, y dicha dispersión es mínima cuando las imágenes se hallan correctamente alineadas, este tipo de medidas resultan muy adecuadas para registrar imágenes multimodales.

Tanto en esta como en las medidas que se describen a continuación se supone el uso de la entropía de Shannon, por ser la más ampliamente utilizada, si bien existen ciertos estudios que contemplan la utilización de otras formas de entropía [12].

La entropía de Shannon [13] de una señal se puede expresar como:

$$H = - \sum_x p(x) \log_2 [p(x)], \quad (2.41)$$

donde  $p(x)$  es la función de densidad de probabilidad de la variable aleatoria  $x$ . En nuestro caso  $p(x)$  correspondería a la función de densidad de probabilidad conjunta de los niveles de gris de las imágenes.



**Figura 2.16:** Representación de la relación entre la entropía de las imágenes y su información mutua.

Nótese que en esta formulación el peso asignado a cada pixel es independiente de su valor de intensidad. Esto permite que la medida dependa únicamente de las propiedades estadísticas de la función de densidad de probabilidad conjunta de las intensidades, permitiendo emplearla para registrar imágenes entre cuyos niveles de intensidad no existe una relación funcional, a la vez que se consigue mayor robustez frente a errores aberrantes.

- Información mutua (MI)

La Información Mutua, medida basada en la teoría de la información, fue propuesta simultáneamente por Viola et al. [14] y Collignon et al. [10] para el registro de imágenes multimodales cuya relación de intensidades no obedece a una ley funcional. La Información Mutua expresa cuánta información de una imagen  $I_A$  contiene otra imagen  $I_B$ , como se muestra en la figura 2.16, siendo pues máxima cuando dos imágenes equivalentes están correctamente alineadas. Esta medida se expresa como:

$$MI(I_A, I_B) = H(I_A) + H(I_B) - H(I_A, I_B), \quad (2.42)$$

con

$$H(I) = - \sum_{i \in I} p(i) \log p(i) \quad (2.43)$$

$$H(I_A, I_B) = - \sum_{i \in I_A} \sum_{j \in I_B} p(i, j) \log p(i, j), \quad (2.44)$$

donde  $H(I_A)$  y  $H(I_B)$  son las entropías marginales de las imágenes y  $H(I_A, I_B)$  es su entropía conjunta.

Más adelante dentro de este capítulo se detallaran los diversos modos de estimar la función de densidad de probabilidad conjunta  $p(i, j)$  de las imágenes (las funciones de densidad de probabilidad marginales pudiéndose derivar fácilmente de ésta) y su impacto en el resultado del registro. Es importante sin embargo recordar que la calidad de este tipo de medidas de similitud estadísticas está estrechamente relacionada con la cantidad de datos disponibles para realizar dicha estimación, por lo que sus resultados son mucho mejores para el caso de registro de volúmenes tomográficos que cuando se trabaja con imágenes bidimensionales.

- Información mutua normalizada (NMI)

Esta variante de la información mutua fue introducida por Studholme [15] para evitar que las dimensiones de la superposición entre las imágenes afectara a la medida. Se define como:

$$\text{NMI}(I_A, I_B) = \frac{H(I_A) + H(I_B)}{H(I_A, I_B)}, \quad (2.45)$$

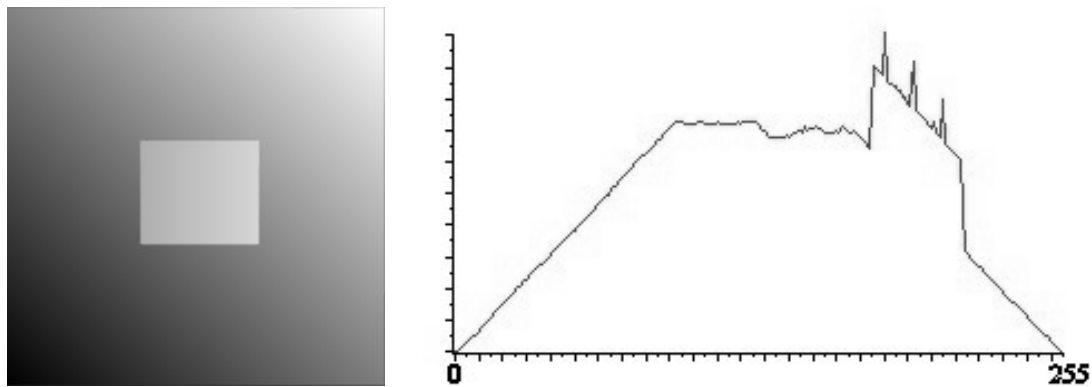
donde, como en el caso anterior,  $H(I_A)$  y  $H(I_B)$  son las entropías marginales de las imágenes y  $H(I_A, I_B)$  es su entropía conjunta.

- Otras medidas

Existen innumerables medidas de similitud cuya descripción está fuera tanto de las posibilidades como de los objetivos de este trabajo. Citamos a continuación algunos de los más importantes que se han encontrado durante el curso de la presente tesis: Numero de cambios de signo de la diferencia de intensidades, coeficiente de correlación, correlación estadística y *matched filters*, coeficiente de correlación en el dominio de Fourier, correlación de fase, relación de correlación, entropía de la imagen diferencia, información mutua generalizada, varianza de cocientes de intensidad, varianza de niveles de gris entre segmentos, criterio del coseno, dispersión del histograma, cambio de signo estocástico, cambio de signo determinista, filtrado de eco cepstral, campo de flujo óptico, distancia entre expansiones de Taylor locales, interpretación como superficie 4D, desviación estándar normalizada,...

#### 2.4.4. Estimación del histograma y la fdp.

Incluimos esta sección dada la importancia que han cobrado en los últimos años los métodos de registro de imágenes basados en medidas estadísticas tales como la entropía y la información mutua. Dichas medidas se basan en la función de densidad de probabilidad de las intensidades de las imágenes, y esta función se estima, directa o indirectamente a partir del histograma de las imágenes. Pueden asimismo calcularse a partir del histograma conjunto medidas tales como la suma de diferencias cuadráticas, la divergencia y el coeficiente de correlación.



**Figura 2.17:** Ejemplo de histograma de intensidades de una imagen.

Citamos a continuación algunas de las técnicas más habituales para aproximar el histograma y la función de densidad de probabilidad, así como su efecto sobre el rendimiento del sistema.

### Histograma e histograma conjunto

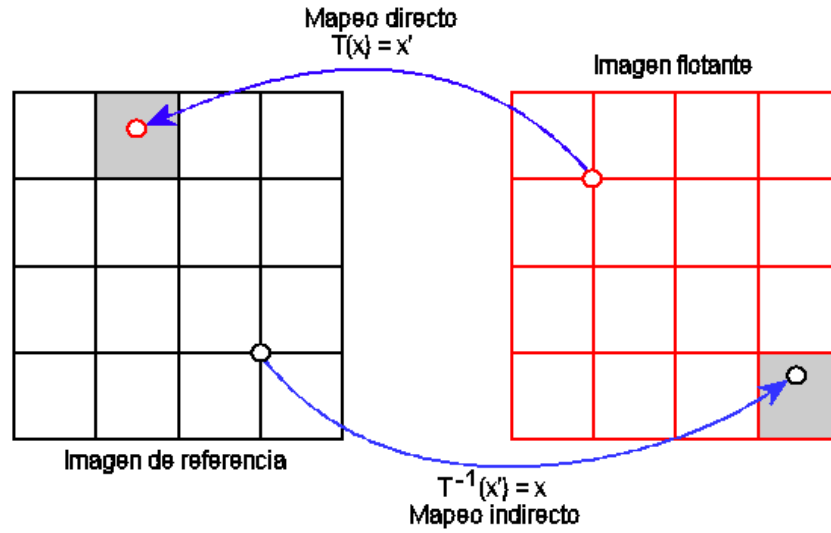
Se denomina *histograma* de una variable aleatoria a la representación que, dada una partición regular del rango de valores que puede tomar, indica el número de ocurrencias dentro de una muestra comprendidas en cada intervalo (fig. 2.17). Se denomina *curva de frecuencia* al gráfico que se obtiene en el caso límite en que el número de elementos de la muestra tiende a infinito y la anchura de los intervalos tiende a cero.

Esta representación es generalizable al caso  $n$ -dimensional, y se habla entonces de histograma conjunto. En nuestro caso nos interesaremos por el histograma conjunto de los pares de valores de intensidad correspondientes en una y otra imagen dada una determinada transformación de coordenadas. El interés de dicho histograma conjunto radica en el hecho de que a partir de su distribución se puede estimar el alineamiento entre las imágenes, hecho del que sacan partido algunas medidas de similitud. El histograma marginal de cada una de las variables implicadas puede calcularse a partir del histograma conjunto por integración a lo largo de cada uno de sus ejes.

Dado que es a partir de este histograma conjunto que en muchos casos se realizará la medida de la bondad del registro para los presentes parámetros de transformación, y dado que por tanto dicho histograma deberá recalcularse en cada actualización de estos parámetros, es evidente la importancia que la rapidez y calidad en su cálculo tendrán sobre el sistema de registro. En efecto, para casi todos de los sistemas de registro basado en la entropía desarrollados actualmente, el cálculo del histograma conjunto representa la mayor parte de la carga computacional total.

Se puede abordar la construcción del histograma conjunto mediante un *método indirecto*, o mediante un *método directo*, actualizando iterativamente sus valores a partir de las imágenes originales.





**Figura 2.18:** Esquema de mapeo directo e indirecto de la transformación. Las zonas sombreadas indican el área donde se realiza la interpolación.

En el método indirecto se genera una imagen intermedia transformada  $I_A^T = I_A \circ T$ , tras lo cual para calcular el histograma conjunto basta con recorrer simultáneamente las imágenes  $I_B$  y  $I_A^T$ . Resulta conceptualmente más simple y su implementación es trivial, sin embargo debe tenerse en cuenta el coste en recursos de almacenar la imagen intermedia, especialmente si se trabaja con grandes volúmenes de datos.

El método directo, por otra parte, emplea la transformación para aplicar uno a uno los voxels de una imagen sobre la otra y actualiza el histograma a medida que lo hace. Este mapeo de una imagen sobre la otra puede ser *directo* o *inverso*, en función de si se emplea la transformación tal como está definida o su inversa, como se ilustra en la figura 2.18.

En principio el mapeo directo es más rápido si la imagen que se proyecta contiene menos muestras que la de referencia, caso habitual dado que se tiende a mantener estática la imagen con mejor resolución. El mapeo indirecto es sin embargo más fácil de implementar si se desea generar una imagen intermedia, puesto que el valor de cada elemento en la nueva imagen se obtiene inmediatamente por interpolación, mientras que mediante mapeo directo sería necesaria una segunda pasada para ponderar todas las contribuciones al mismo.

Una forma de acelerar las transformaciones es aplicarlas de forma iterativa. Por ejemplo, para aplicar una transformación  $T(\vec{x})$  iterativamente por líneas a todos los voxels  $\vec{x} \in I_A$  se haría:

$$T(\vec{x}) = \begin{pmatrix} a & b & c & d \\ e & f & g & k \\ i & j & k & l \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} x' = ax + by + cz + d \\ y' = ex + fy + gz + h \\ z' = ix + jy + kz + l \end{pmatrix} \quad (2.46)$$

$$\begin{cases} x'_{n+1} = x'_n + a & y & x'_0 = by + cz + d \\ y'_{n+1} = y'_n + e & y & y'_0 = fy + gz + h \\ z'_{n+1} = z'_n + i & y & z'_0 = jy + kz + l \end{cases} \quad (2.47)$$

### Efectos de la interpolación

Dado que la transformación empleada para efectuar el registro raramente hará coincidir cada pixel de la imagen transformada exactamente sobre la posición de un pixel en el espacio destino, será necesario efectuar una interpolación, bien para determinar el valor de la imagen de referencia en ese punto, bien para calcular el valor de la imagen transformada sobre la retícula de destino. Dicha interpolación tendrá un impacto significativo en el sistema de registro, comportando siempre un compromiso entre rapidez de cálculo y precisión.

Si bien se ha dedicado el apéndice B a discutir en profundidad el problema general de la interpolación, dada su importancia tanto en el tratamiento como en la visualización de imágenes, en este apartado queremos analizar el caso particular de la interpolación en el cálculo del histograma conjunto de las imágenes. En efecto, hemos visto que el cálculo del histograma se puede realizar de forma directa sin necesidad de calcular una imagen intermedia transformada. Describimos a continuación los métodos de interpolación más comúnmente aplicados, para el caso de imágenes bidimensionales, si bien su generalización a  $n$  dimensiones es trivial.

Sean  $I_A$  y  $I_B$  las dos imágenes a registrar y sea  $p = (\vec{x}, I_A(\vec{x}))$  un elemento de la imagen a registrar, cuya transformada es  $p' = (T(\vec{x}), I_A(\vec{x}))$ . Denominamos  $\vec{n}_k$  con  $k = 0, 3$  a los cuatro elementos de  $I_B$  vecinos a la posición  $T(\vec{x})$  y  $\mu_k$  a las cuatro superficies delimitadas por  $T(\vec{x})$  y  $\vec{n}_k$ , como se muestra en el esquema de la figura 2.19.

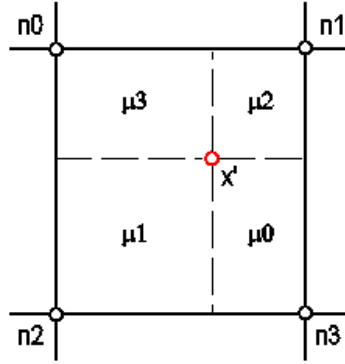
#### ■ Interpolación por vecino más próximo

Esta técnica es la más simple y rápida de las expuestas, ya que sólo precisa del redondeo de las coordenadas del elemento transformado. Como contrapartida, genera importantes discontinuidades que afectan a la medida de registro en forma de *artefactos de interpolación* y que dificultan el funcionamiento de la optimización.

$$\text{Hist}(i, j) = \sum_{\Omega} \delta(i - I_A(\vec{x}), j - I_B(\vec{n}_k)) \quad (2.48)$$

con

$$k = \arg \min_k |T(\vec{x}) - \vec{n}_k| \quad (2.49)$$



**Figura 2.19:** Relación del elemento transformado a interpolar  $x'$  con sus cuatro vecinos  $n_i$ . Los términos  $\mu_i$  indican las áreas delimitadas por la posición de  $x'$  con respecto a la malla.

siendo  $\Omega$  la zona de solapamiento de las imágenes y  $\delta$  una función que toma valor 1 en su origen y 0 en el resto.

- Interpolación lineal

La interpolación lineal reduce significativamente los problemas de discontinuidad que presentaba la interpolación por vecino más próximo, si bien su coste computacional se estima alrededor de un 90 % superior al de ésta.

$$\text{Hist}(i, j) = \sum_{\Omega} \delta \left( i - I_A(\vec{x}), j - \left\lfloor \sum_k \mu_k I_B(\vec{n}_k) \right\rfloor \right) \quad (2.50)$$

donde la función  $\lfloor x \rfloor$  denota la aproximación al entero más cercano.

- Interpolación por volumen parcial

La interpolación por volumen parcial, al contrario que las anteriores, no es una técnica clásica de interpolación de datos sino que fue introducida en [16] para resolver el problema de la construcción del histograma conjunto en particular. Este método minimiza los problemas de discontinuidad, a la vez que comporta un coste computacional tan solo un 67 % superior al de la interpolación por vecino más próximo.

$$\text{Hist}(i, j) = \sum_{\Omega} \sum_k \mu_k \delta(i - I_A(\vec{x}), j - I_B(\vec{n}_k)) \quad (2.51)$$

Se ha mencionado ya el efecto negativo que la discontinuidad en la interpolación tiene sobre la medida de registro. Otra importante fuente de artefactos de interpolación, a tener en cuenta cuando se desee conseguir precisiones subvoxel, es el alineamiento de las retículas de las imágenes. En efecto, cuando las dimensiones de las retículas

de ambas imágenes permiten que, para ciertos parámetros de la transformación, éstas queden alineadas en muchos puntos, dicho alineamiento puede reflejarse en la medida de registro. Así, para una interpolación por volumen parcial, la información mutua de las imágenes presenta un rizado no deseado con mínimos locales en las posiciones de alineamiento. Del mismo modo, la interpolación lineal generará el mismo artefacto si bien las posiciones de alineamiento de la retícula se corresponderán en este caso con máximos locales de la información mutua. Este problema puede solucionarse en parte muestreando de nuevo una de las imágenes, de modo que se minimice la probabilidad de que muchos nodos de la retícula queden alineados al mismo tiempo. Opcionalmente, el empleo de ciertas medidas de la entropía [17] o una estimación robusta de la función de densidad de probabilidad, como las ventanas de Parzen, pueden compensar este problema.

### Estimación de las probabilidades

A menudo no basta con obtener el histograma conjunto de las imágenes, ya que ciertas medidas precisan de una estimación de la función de densidad de probabilidad conjunta de las intensidades. A continuación se describen los tres métodos más comúnmente empleados para realizar dicha estimación.

- Aproximación frecuencial

Se trata del método más simple para obtener una aproximación de la función de densidad de probabilidad, mediante una normalización del histograma frecuencial conjunto de las imágenes para una determinada transformación. Así, dado el histograma conjunto  $Hist(z)$ , la densidad de probabilidad se aproxima como:

$$p(z) = \frac{Hist(z)}{N} \quad (2.52)$$

donde  $z = (z_A, z_B)$  son las intensidades de pares correspondientes de voxels y  $N$  es el número de muestras en la zona de superposición de las imágenes.

- Aproximación Bayesiana

Éste es un caso algo más general y robusta de la aproximación frecuencial anteriormente descrita, ya que introduce un término  $\lambda$  que suaviza la función.

$$p(z) = \frac{Hist(z) + \lambda}{N + \lambda K} \quad (2.53)$$

donde  $K$  indica el número de intervalos del histograma conjunto de las imágenes.

- Ventanas de Parzen

Se trata de una técnica no paramétrica de estimación de la función de densidad de probabilidad de una señal. La densidad de probabilidad estimada se expresa

mediante la superposición de *kernels* centrados en los elementos de un conjunto de muestras  $A$  de la señal  $z$ , en nuestro caso valores de intensidad de parejas de voxels correspondientes. Así, la estimación de densidad por ventanas de Parzen se escribe:

$$p(z) \approx P^*(z) = \frac{1}{N_A} \sum_{z_j \in A} R(z - z_j), \quad (2.54)$$

donde  $N_A$  es el número de elementos de la muestra  $A$  y  $R$  es una función ventana con integral unidad. Las funciones ventana más comúnmente empleadas son la gaussiana y la spline cúbica, por la simplicidad de su derivación.

### Talla del histograma conjunto

Un factor importante a tener en cuenta cuando se desea realizar la estimación de la función de densidad de probabilidad conjunta de las intensidades de dos imágenes es si se dispone de datos suficientes para tal fin. Así, si se deseara estimar la f.d.p. sobre dos imágenes de 8bits se deberían contemplar  $256 \times 256 = 65536$  posibles combinaciones. Es evidente que para que nuestra estimación fuera estadísticamente significativa sería preciso disponer de un volumen de datos muy superior al presente en la mayor parte de aplicaciones.

Para evitar el anterior problema es preciso seleccionar cuidadosamente el número de clases (*bins*) en que se divide cada eje del histograma. La *regla de Sturges* ofrece un criterio para seleccionar el número de clases  $k$  para calcular el histograma de una distribución normal monodimensional:

$$k = 1 + \log_2 n \quad (2.55)$$

donde  $n$  es el tamaño de la muestra. Otros criterios serían la *regla de Scott para la anchura de las clases*:

$$h = 3,5\sigma n^{-\frac{1}{3}} \quad (2.56)$$

y la *regla de Freedman y Diaconis para la anchura de las clases*:

$$h = 2(IQ)n^{-\frac{1}{3}} \quad (2.57)$$

donde  $\sigma$  es la desviación típica de la muestra y  $IQ$  es el rango intercuartil de la muestra.

Si bien para caso de un histograma bidimensional no se dispone de reglas tan directas para la selección del número de clases, en la literatura se pueden encontrar algunas aplicaciones que emplean métodos empíricos con tal fin. En [18], por ejemplo, Thévenaz y Unser ajustan la anchura de las ventanas de Parzen en función del tamaño de la muestra disponible.

## 2.5. Optimización

Bajo este título incluimos la etapa del sistema encargada de determinar los parámetros óptimos de la transformación desde el punto de vista de la medida de similitud seleccionada. Si bien en ciertas aplicaciones esto se podrá conseguir mediante una cálculo directo de dichos parámetros, en la mayoría de casos se deberá recurrir a mecanismos de búsqueda iterativa, con el consiguiente coste computacional.

### 2.5.1. Cálculo directo

A esta primera categoría pertenecen la mayor parte de los métodos de registro rígido basado en elementos de referencia. Dichos métodos se basan en resolver el *problema de Procrustes ortogonal*, [19], es decir, determinar la transformación y rotación que establecen una aplicación entre dos espacios dadas dos nubes de puntos cuyos emparejamientos son conocidos. Se pueden hallar soluciones cerradas de dicho problema empleando descomposición en valores singulares (SVD) [20], descomposición en valores-vectores propios [21] y cuaterniones [22]

### 2.5.2. Cálculo iterativo

Los métodos de optimización se encargan de hallar los valores de un conjunto de variables para los cuales una cierta función objetivo, en nuestro caso la medida de similitud, alcanza un máximo o un mínimo a la vez que se cumplen un conjunto de restricciones (por ejemplo, dos imágenes totalmente desalineadas, para las que sólo se superpone el fondo, pueden presentar un máximo de la medida de similitud, por lo que añadiremos al problema una restricción que impida las traslaciones más allá de un cierto umbral).

En primer lugar describiremos algunos algoritmos de optimización sin restricciones empleados habitualmente en aplicaciones de registro de imágenes, pues los algoritmos de optimización con restricciones son a menudo particularizaciones de estos. Describiremos asimismo ciertos algoritmos que se especializan en el caso particular de los problemas de mínimos cuadrados, dado su conveniencia para ciertas medidas de similitud.

El problema de la optimización sin restricciones trata pues de hallar un minimizador (o maximizador) local de una función  $f(x)$  con valores reales, cuyo parámetro es un vector de variables reales  $x$ . Así pues, el objetivo es hallar un vector  $x^*$  que verifique  $f(x^*) \leq f(x)$  para todo  $x$  en un determinado entorno de  $x^*$ .

Nótese que estamos restringiendo el problema a una búsqueda local del minimizador, es decir, deseamos localizar, a partir de una condición inicial  $x_0$ , el vector  $x^*$  que minimiza  $f(x)$  en una determinada región del espacio de parámetros. En realidad lo que se desearía es realizar una optimización global, esto es, hallar entre todos los posibles vectores de parámetros  $x$  aquél que minimiza  $f(x)$ . Sin embargo éste es un problema mucho más complejo, y no se conoce hoy en día ningún algoritmo

eficiente para realizar ésta tarea. Así pues, el correcto funcionamiento del sistema de registro quedará condicionado a su capacidad para inicializar la optimización en una posición  $x_0$  suficientemente próxima del resultado final. La precisión requerida para esta inicialización dependerá principalmente de la convexidad del criterio de similitud empleado y de la robustez del método de optimización frente a la presencia de mínimos locales en dicho criterio.

Un factor importante a tener en cuenta a la hora de seleccionar un método de optimización es la disponibilidad de un gradiente de la función objetivo. El cálculo del gradiente conlleva en general más tiempo que una simple evaluación de la función, y el cálculo de la Hessiana consume a su vez mucho más tiempo y recursos que el cálculo del gradiente. Por otra parte, los algoritmos que emplean la matriz Hessiana acostumbran a converger en menos iteraciones que aquellos que emplean una estimación de la misma, y aquellos que no emplean ni la matriz Hessiana ni su estimación tienden a ser lentos y poco fiables.

- Método Simplex

También conocido como *ameba*, este algoritmo goza de gran popularidad por su simplicidad y por ser geoméricamente muy intuitivo. Resulta conveniente en aquellas aplicaciones donde el gradiente es difícil de calcular o la función objetivo es ruidosa.

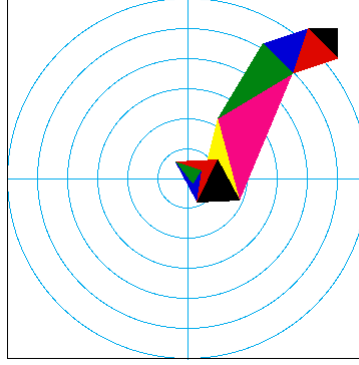
Este algoritmo emplea un *símplex* de  $n + 1$  puntos, donde  $n$  es la dimensión del problema. Así, el *símplex* es un triángulo para problemas bidimensionales y un tetraedro en tres dimensiones. Dicho *símplex* puede efectuar un conjunto de operaciones tales como desplazarse, expandirse, contraerse o distorsionar su forma para intentar minimizar los valores de la función en sus vértices.

Si bien es un método lento y no es adecuado para  $n$  grandes resulta rápido y fácil de implementar, siendo muy utilizado cuando no se dispone del gradiente de la función objetivo. Asimismo, al no ser un método de descenso, puede resultar útil cuando se desee dotar al sistema de una cierta robustez frente a mínimos locales.

La figura 2.20 ilustra el comportamiento de un *símplex* en un espacio de parámetros bidimensional (se trata por tanto de un *símplex* de tres vértices). La función a minimizar, representada por líneas de energía constante, presenta un perfil parabólico, con un mínimo en el origen. Los distintos colores muestran las distintas configuraciones del *simplex* a medida que evoluciona desde su posición inicial en un extremo de la imagen hasta la posición de mínima energía.

- Método de Powell

El método de Powell es en esencia un método de descenso de gradiente conjugado (estos métodos se describen con mayor detalle más adelante), si bien por su simplicidad y la gran popularidad de que ha gozado en el campo del registro de imágenes merece ser descrito a parte.



**Figura 2.20:** Ejemplo de la trayectoria del simplex durante una optimización 2D.

Consiste en efectuar una serie de optimizaciones monodimensionales a lo largo de ciertas direcciones de búsqueda en el espacio de parámetros. Idealmente dichas direcciones deberían ser conjugadas, es decir, la optimización a lo largo de una de ellas no debería empeorar los resultados conseguidos a lo largo de la anterior. El método de Powell permite realizar el cálculo de dichas direcciones sin tener que recurrir al gradiente de la función objetivo.

La optimización monodimensional emplea en general el *método de Brent*, también conocido como el método de van Wijngaarden-Deker-Brent, un algoritmo de búsqueda de raíces mediante *root-bracketing*, bisección de intervalos e interpolación cuadrática inversa. Este método debe converger siempre que la función sea evaluable en el intervalo que contenga la raíz.

Para tres puntos  $x_1$ ,  $x_2$  y  $x_3$  éste método ajusta  $x$  como una función cuadrática de  $y$ , y emplea luego la fórmula de interpolación:

$$x = \frac{[y - f(x_1)][y - f(x_2)]x_3}{[f(x_3) - f(x_1)][f(x_3) - f(x_2)]} + \frac{[y - f(x_2)][y - f(x_3)]x_1}{[f(x_1) - f(x_2)][f(x_1) - f(x_3)]} + \frac{[y - f(x_3)][y - f(x_1)]x_2}{[f(x_2) - f(x_3)][f(x_2) - f(x_1)]}. \quad (2.58)$$

Las posteriores estimaciones de la raíz se consiguen fijando  $y = 0$ , quedando:

$$x = x_2 + \frac{P}{Q}, \quad (2.59)$$

donde



$$P = S[R(R - T)(x_3 - x_2) - (1 - R)(x_2 - x_1)] \quad (2.60)$$

$$Q = (T - 1)(R - 1)(S - 1) \quad (2.61)$$

con

$$R \equiv \frac{f(x_2)}{f(x_3)} \quad (2.62)$$

$$S \equiv \frac{f(x_2)}{f(x_1)} \quad (2.63)$$

$$T \equiv \frac{f(x_1)}{f(x_3)} \quad (2.64)$$

- Métodos de descenso de gradiente

Este tipo de algoritmos, también denominados *steepest descent*, permiten hallar un mínimo local de una función para la que se puede calcular el gradiente:

$$\nabla f(x) = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix} \quad (2.65)$$

El algoritmo parte de un punto  $P_0$ , actualizando iterativamente a la posición  $P_{i+1}$  mediante una minimización a lo largo de la trayectoria que parte de  $P_i$  y sigue la dirección definida por el gradiente local  $-\nabla f(P_i)$ . En el caso de una función monodimensional  $f(x)$ , el algoritmo toma la forma:

$$x_i = x_{i-1} - \epsilon f'(x_{i-1}) \quad (2.66)$$

donde  $x_0$  es el punto de partida y  $\epsilon \ll 0$  permite ajustar el paso de actualización. El proceso se itera hasta verificar un cierto criterio de parada que garantice haber alcanzado una posición estable.

Un inconveniente de estos métodos es que requieren un gran número de iteraciones en el caso de funciones que presentan valles largos y estrechos, en cuyo caso es preferible emplear un método de gradiente conjugado.

- Métodos de gradiente conjugado

Estos algoritmos, al igual que los de descenso de gradiente descritos anteriormente, permite localizar el mínimo local más próximo de una función de  $n$  variables para la que se puede calcular el gradiente. En este caso, sin embargo, en lugar del gradiente local se emplean direcciones conjugadas para efectuar el

descenso. De este modo, en el caso de funciones con mínimos en forma de valle se consigue alcanzar el óptimo con muchas menos iteraciones.

Así, dada la posición  $x_k$  para una iteración y la dirección de búsqueda  $d_k$ , la nueva posición se obtiene mediante la actualización  $x_{k+1} = x_k + \alpha_k d_k$ . Las distintas variantes de este método difieren entre si principalmente en el modo de calcular la nueva dirección de búsqueda. La variante de Fletcher-Reeves genera la dirección conjugada  $d_{k+1}$  mediante:

$$d_{k+1} = -\nabla f(x_{k+1}) + \beta_k d_k, \quad (2.67)$$

con

$$\beta_k = \left( \frac{\|\nabla f(x_{k+1})\|_2}{\|\nabla f(x_k)\|_2} \right)^2. \quad (2.68)$$

Se puede mejorar el rendimiento del método mediante el rearranque periódico, es decir, forzando  $\beta_k = 0$ .

Otra variante es la de Polak-Ribiere, que define  $\beta_k$  como

$$\beta_k = \frac{[\nabla f(x_{k+1}) - \nabla f(x_k)]^T \nabla f(x_{k+1})}{\nabla f(x_k)^T \nabla f(x_k)}. \quad (2.69)$$

Si bien las dos variantes presentadas son equivalentes cuando  $f$  es cuadrática, las pruebas parecen indicar que el método de Polak-Ribiere tiende a ser más eficiente que el de Fletcher-Reeves.

#### ■ Método de Newton

El método de Newton es la base de una gran variedad de algoritmos de optimización que emplean a la vez el vector gradiente y la matriz Hessiana:

$$\nabla^2 f(x) = (\partial_j \partial_i f(x)). \quad (2.70)$$

A pesar del coste computacional que conlleva el cálculo de la matriz Hessiana, en muchos casos el superior rendimiento del algoritmo justifica su uso. Más adelante se presentarán ciertos métodos que logran un cierto compromiso empleando tan solo una estimación de dicha matriz.

El método de Newton se basa en efectuar un modelado de Taylor de la función objetivo entorno a la posición  $x_k$  correspondiente a la presente iteración  $k$  mediante una función cuadrática, definida como:

$$q_k(\hat{x}) = f(x_k) + \nabla f(x_k)^T \hat{x} + \frac{1}{2} s^T \nabla^2 f(x_k) \hat{x}. \quad (2.71)$$

En el método de Newton básico se obtiene la nueva iteración a partir del mínimo del modelo  $q_k$ . Dicho mínimo es único si la matriz Hessiana  $\nabla^2 f(x_k)$  es definida positiva, y se puede obtener resolviendo el sistema lineal:

$$\nabla^2 f(x_k) \hat{x}_k = -\nabla f(x_k). \quad (2.72)$$

De este modo, la nueva iteración será:

$$x_{k+1} = x_k + \hat{x}_k. \quad (2.73)$$

Se garantiza la convergencia cuando la posición inicial está lo suficientemente próxima de un mínimo local  $x^*$  en el que la Hessiana es definida positiva. Una ventaja adicional es que la convergencia es cuadrática, es decir, dada una constante  $\beta$  positiva,

$$\|x_{k+1} - x^*\| \leq \beta \|x_k - x^*\|^2. \quad (2.74)$$

Cabe decir que en la mayoría de problemas el método de Newton debe ser modificado en cierta medida para alcanzar la convergencia.

■ Métodos de aproximación por diferencias

En este caso se trata de reducir el coste computacional sustituyendo la matriz Hessiana por una aproximación razonablemente precisa de la misma. De este modo, se pueden emplear algoritmos que precisen de dicha matriz, alterando tan solo el cálculo de la misma, que en este caso se aproxima mediante diferencias de los valores del gradiente.

Pueden emplearse diferencias hacia delante o diferencias centradas. En el primer caso, se aproxima la columna  $i$ -ésima de la matriz Hessiana como:

$$\frac{\nabla f(x_k + h_i e_i) - \nabla f(x_k)}{h_i} \quad (2.75)$$

donde  $h_i$  es un parámetro que controla la diferencia. El vector  $e_i$  únicamente contiene un 1 en su posición  $i$ -ésima y ceros en el resto.

Si se emplean diferencias centradas, la columna  $i$ -ésima de la matriz Hessiana se aproxima como:

$$\frac{\nabla f(x_k + h_i e_i) - \nabla f(x_k - h_i e_i)}{2h_i} \quad (2.76)$$

Es preciso seleccionar cuidadosamente el valor del parámetro que controla las diferencias,  $h_i$ , ya que si es demasiado pequeño los cálculos se ven afectados por

los errores de redondeo, mientras que si es demasiado grande predominarán los errores de truncado.

La gran mayoría de códigos emplean diferencias hacia delante, ya que en general ofrecen la suficiente precisión, mientras que las diferencias centradas, si bien son más precisas, requieren el doble de evaluaciones del gradiente.

#### ■ Quasi-Newton

De nuevo estos métodos tratan de evitar la necesidad de calcular la matriz Hessiana, sea por su dificultad o por el coste computacional que conlleva. En este caso en lugar de estimar dicha matriz en un punto, se construye iterativamente la estimación a partir de todas las posiciones  $x_k$  recorridas durante la ejecución del algoritmo.

Así, para una iteración  $x_k$ , dada la actual aproximación de la matriz Hessiana  $B_k$ , se genera una nueva dirección de búsqueda  $d_k$  resolviendo el siguiente sistema lineal:

$$B_k d_k = -\nabla f(x_k), \quad (2.77)$$

tras lo cual se efectúa la búsqueda a lo largo de la dirección  $d_k$ , estableciéndose la nueva posición tal que:

$$x_{k+1} = x_k + \alpha_k d_k. \quad (2.78)$$

Hecho esto, se procede a actualizar la nueva estimación de la matriz Hessiana  $B_{k+1}$  a partir de la anterior estimación  $B_k$ . La familia más común de métodos de actualización son las actualizaciones de rango dos de Broyden, que presentan la siguiente estructura:

$$B_{k+1} = B_k - \frac{B_k s_k (B_k s_k - k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} + \phi_k [s_k^T B_k s_k] v_k v_k^T, \quad (2.79)$$

donde  $\phi_k \in [0, 1]$  y

$$v_k = \left[ \frac{y_k}{y_k^T s_k} - \frac{B_k s_k}{s_k^T B_k s_k} \right]. \quad (2.80)$$

Para el caso  $\phi_k = 0$  se consigue la actualización de Broyden-Fletcher-Goldfarb-Shanno (BFGS), que es el método más utilizado en la mayoría de casos. Otra opción es la actualización de Davidon-Fletcher-Powell (DFP), que se obtiene fijando  $\phi_k = 1$ .

La actualización de  $B_k$  se puede realizar bien actualizando la descomposición de Cholesky, bien actualizando la inversa de  $B_k$ . El coste computacional de efectuar

la actualización resolviendo el sistema 2.77 es en ambos casos de orden  $n^2$ . Si bien la actualización de la inversa es más simple, se considera más fiable la actualización de la factorización de Cholesky.

Este tipo de métodos superan ampliamente a los métodos de descenso de gradiente descritos con anterioridad, que también precisan únicamente de la primera derivada. Si bien los métodos quasi-Newton requieren un número ligeramente mayor de operaciones por iteración y algo más de memoria, en la mayoría de casos esto queda compensado por una mejor convergencia.

#### ■ Gauss-Newton

Los problemas de mínimos cuadrados son un caso particular de especial interés en nuestro ámbito. Se definen como:

$$\min r(x) : x \in \mathbb{R}^n, \quad (2.81)$$

siendo  $r(x)$  una función de la siguiente forma:

$$r(x) = \frac{1}{2} \|f(x)\|_2^2 \quad (2.82)$$

donde  $f(x)$  es una función vectorial de residuos  $\mathbb{R}^n \Rightarrow \mathbb{R}^m$ .

Lo que distingue a los problemas de mínimos cuadrados del problema general de la optimización sin restricciones es la estructura de la matriz Hessiana, de  $r(x)$ . Se puede emplear el jacobiano de  $f(x)$ ,  $f'(x) = (\partial_1 f(x), \dots, \partial_n f(x))$  para expresar el gradiente de  $r(x)$  tal que  $\nabla r(x) = f'(x)^T f(x)$ . Asimismo  $f'(x)$  forma parte de la matriz Hessiana

$$\nabla^2 r(x) = f'(x)^T f'(x) + \sum_{i=1}^m f_i(x) \nabla^2 f_i(x). \quad (2.83)$$

Así pues, se tiene que calcular la matriz Jacobiana  $f'(x)$  para obtener el gradiente de  $r(x)$ , y hecho esto ya se dispone del primer término de la matriz Hessiana  $\nabla^2 r(x)$ . Esta propiedad es de la que sacan partido los algoritmos de mínimos cuadrados, ya que en casi todas las aplicaciones prácticas el primer término  $f'(x)^T f'(x)$  de  $\nabla^2 r(x)$  tiene más peso que el segundo. Esto es especialmente cierto cuando los residuos  $f_i(x)$  son pequeños cerca del óptimo. Se considera que los residuos de un problema son pequeños si

$$|f_i(x)| \|\nabla^2 f_i(x)\|, \quad i = 1, 2, \dots, n \quad (2.84)$$

son pequeños en comparación con el menor valor propio de  $f'(x)^T f'(x)$  para toda  $x$  en un entorno de la solución.

Especialmente adecuado en caso de residuos pequeños es el algoritmo de Gauss-Newton. En éste se aproxima la Hessiana por su primer término, y la dirección de búsqueda  $d_k$  se determina mediante el sistema lineal:

$$(f'(x_k)^T f'(x_k)) d_k = -f'(x_k)^T f(x_k). \quad (2.85)$$

■ Método de Levenberg-Marquardt

El método de Levenberg-Marquardt ofrece una alternativa al método de Gauss-Newton para hallar un mínimo local de una función  $r(x)$  compuesta por una suma, de términos cuadráticos.

Este método efectúa la búsqueda en la dirección  $d_k$  que satisface las ecuaciones:

$$(f'(x_k)^T f'(x_k) + \lambda_k I) d_k = -f'(x_k)^T f(x_k), \quad (2.86)$$

donde  $\lambda_k$  es un valor escalar no negativos y  $I$  es la matriz identidad. Cuando  $\lambda_k$  es igual a cero, la dirección  $d_k$  es idéntica a la del método de Gauss-Newton, mientras que a medida que  $\lambda_k$  tiende a infinito  $d_k$  tiende a comportarse como un descenso de gradiente. Así, el término  $\lambda_k$  permite controlar el descenso, aprovechando las robustez de un descenso de gradiente cuando está lejos del mínimo y pasando gradualmente a la eficiencia de un Gauss-Newton a medida que converge.

## 2.6. Modalidades de imagen implicadas

Uno de los principales condicionantes en el diseño de un sistema de registro son las modalidades de las imágenes sobre las que se desea trabajar. En efecto, todas las medidas de similitud basadas en la localización de referencias intrínsecas, así como aquellas basadas en la segmentación de estructuras anatómicas, deberán particularizarse de acuerdo con las características de la modalidad sobre la que se trabaje.

Así, un sistema de registro de imágenes de ultrasonidos deberá ser especialmente robusto frente al ruido impulsional, o si por ejemplo se desea registrar imágenes CT con imágenes PET basándose en una segmentación de los huesos se deberá tener en mente que las primeras muestran la estructura de calcio, mientras que las segundas reflejan tan sólo la actividad metabólica en su interior.

En lo que respecta a las medidas de similitud basadas en el contenido íntegro de las imágenes, el registro de datos de distintas modalidades es la principal causa de que no se pueda suponer una relación funcional entre los niveles de gris de las mismas, con la consiguiente necesidad de emplear fórmulas más generales y de más difícil optimización.

En función de las modalidades implicadas, se consideran los siguientes tipos de problemas de registro

- Registro monomodal

El registro de dos imágenes de la misma modalidad es sin duda el que menos complicación presenta, y en la mayoría de casos existen ya técnicas que resuelven satisfactoriamente este tipo de problemas. Una posible excepción es cuando las imágenes presentan severas deformaciones de naturaleza no rígida, como en el caso de registro de imágenes procedentes de distintos sujetos o en el estudio de la evolución temporal de ciertas patologías.

- Registro multimodal

La complejidad que conlleva el registro de imágenes procedentes de distintas modalidades depende en gran medida de dichas modalidades. Así, el registro de imágenes CT-MRI, ambas modalidades que ofrecen un buen nivel de detalle de la estructura anatómica, es considerablemente menos problemático que el registro de imágenes CT-PET, en cuyo caso se debe hacer corresponder una información de naturaleza anatómica con una información eminentemente funcional.

- Registro de imagen a modelo

En ciertos casos puede interesar efectuar el registro de una imagen médica con un modelo anatómico de la zona visualizada. Este es el caso, por ejemplo, de aquellas aplicaciones en las que se desean cuantificar las desviaciones que una cierta estructura presenta respecto a un modelo ideal de dicha estructura, con el fin de diagnosticar una patología o estudiar su evolución.

- Registro de paciente a imagen

Este es un caso que se da principalmente en aplicaciones de cirugía asistida por computador. Se busca aquí no tanto efectuar el registro de dos imágenes como alinear al propio paciente con un conjunto de datos, en general procedentes de estudios preoperatorios. Dadas las severas restricciones de velocidad y robustez precisas en este tipo de aplicaciones, el registro se basa normalmente en elementos marcadores extrínsecos de algún tipo.

## 2.7. Sujeto visualizado

El sujeto visualizado en cada una de las imágenes a registrar tiene, como es natural, un importante impacto en la dificultad de la aplicación que se considera. En este sentido podemos distinguir tres posibles casos:

- Registro intrasujeto

El registro intrasujeto supone que las imágenes con las que se trabaja provienen todas del mismo individuo, sea procedentes de diversas modalidades de adquisición, sea a partir de datos de una única modalidad tomados bajo distintas condiciones de adquisición.



- Registro intersujeto

En este caso las imágenes a registrar provienen de distintos sujetos. A menudo esto representa una importante complicación del sistema, debido a la gran variabilidad anatómica existente entre distintos individuos.

- Registro de sujeto con atlas

El registro de imágenes con un atlas anatómico es un caso de sumo interés en aquellas aplicaciones donde se desea poner los datos en un determinado marco de referencia. Esto puede ser útil para cuantificar las desviaciones de cada caso respecto a dicha referencia. Asimismo, en algunos casos el registro con un atlas se puede emplear como alternativa a la segmentación de una imagen.

## 2.8. Objeto visualizado

En general las técnicas de registro de imágenes son sumamente restrictivas en su campo de aplicación, estando estrechamente ligadas a la estructura o región anatómicas para las que han sido diseñadas. Esto es debido a las importantes diferencias en las propiedades de las distintas regiones anatómicas, tanto desde el punto de vista de las deformaciones a las que se ven sometidas como por la forma en que se reflejan en las distintas modalidades de imágenes.

Todos estos aspectos deben ser cuidadosamente considerados antes de emprender el diseño de un sistema de registro, de modo que se éste se ciña lo mejor posible a las particularidades del objeto de estudio.

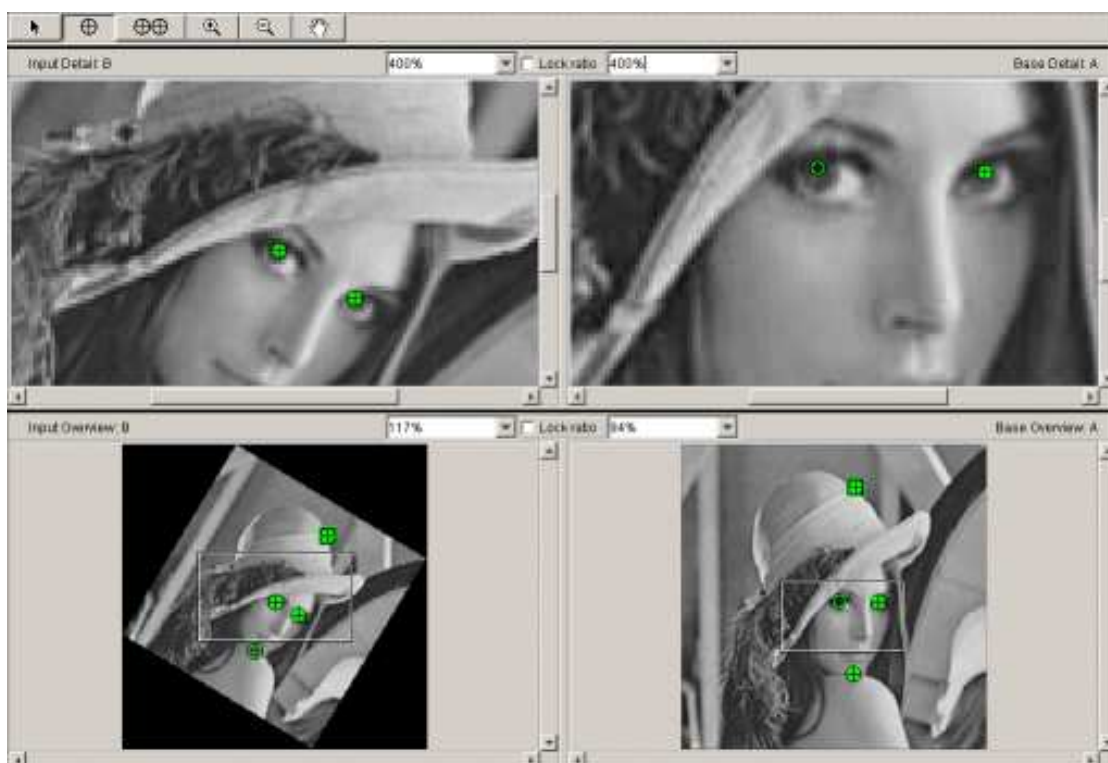
## 2.9. Necesidad de interacción

Existe un compromiso importante en lo referente al grado de interacción del usuario con un sistema de registro de imágenes. Si bien es cierto que un sistema completamente automatizado sería lo ideal, esta automatización acostumbra a implicar un alto en velocidad y robustez. Por contra, aunque la participación de un usuario experimentado en la inicialización, supervisión y validación facilita enormemente la puesta en correspondencia, exigir demasiada participación puede hacer que un sistema resulte poco práctico o que se prefiera continuar realizando un registro manual de las imágenes.

Tradicionalmente se consideran tres categorías en función del grado de interacción de un sistema:

- Interactivo

En este caso el usuario establece manualmente el registro de las imágenes, asistido por herramientas informáticas que muestran gráfica o numéricamente el alineamiento conseguido. En ocasiones el sistema ofrece una primera aproximación al registro, dejando que el usuario realice el ajuste final.



**Figura 2.21:** Ejemplo de interfaz de registro interactivo. Las marcas verdes indican los puntos de referencia introducidos por el usuario que se emplearán para establecer el registro de las imágenes.

Si bien en muchos casos este tipo de sistemas son preferidos por aquellos usuarios que prefieren mantener el control sobre los resultados, no pueden en general conseguirse las precisiones que ofrece un método automático. Por otra parte, esta clase de herramientas interactivas resultan difíciles de emplear con volúmenes tridimensionales de datos, especialmente si las transformaciones implicadas no son rígidas.

La figura 2.21 muestra una aplicación típica de registro interactivo de imágenes, donde el usuario debe indicar un conjunto de puntos de referencia comunes en ambas imágenes, cuyo número varía en función del tipo de transformación deseada.

- Semiautomático

Denominamos semiautomáticos a todos aquellos sistemas que, si bien se basan en una búsqueda automática de la transformación óptima, requieren la intervención del usuario para inicializar dicha búsqueda, supervisar su evolución y/o validar los resultados conseguidos.

Si bien los mayores esfuerzos de investigación están hoy en día dirigidos a

conseguir sistemas de registro totalmente automatizados, en la práctica son los sistemas semiautomáticos los que gozan de mayor aceptación, especialmente en el campo de la medicina.

- Automático

Como su nombre indica, estos sistemas no requieren en ningún momento la intervención del usuario. Cabe decir que son pocos los sistemas de este tipo que se pueden encontrar en el ámbito médico en que se centra esta tesis, puesto que la supervisión de los resultados por parte de un especialista es un requisito difícilmente eludible.

## Capítulo 3

# Diagnosis en oncología de tórax y abdomen

El eje principal y objetivo último de la tesis es el desarrollo de nuevas técnicas de registro elástico de imágenes para su aplicación en un entorno médico concreto, la oncología de la región torácica y abdominal.

En este capítulo se demostrará el interés de emplear las técnicas de registro para reforzar el diagnóstico oncológico en las regiones torácica y abdominal, afectadas ambas por severas deformaciones debidas tanto a la actividad metabólica del paciente como al protocolo de adquisición de las imágenes. Este interés deriva de la información complementaria que aportan las modalidades CT y PET. En efecto, las imágenes CT muestran con detalle las distintas estructuras anatómicas de la zona visualizada, mientras que un estudio PET brinda información sobre la actividad metabólica asociada a dichas estructuras.

Se revisarán los distintos métodos descritos en la literatura que han abordado el registro de tórax y abdomen, evaluándose sus pros y contras. Se justificará asimismo la incapacidad de dichos métodos para resolver de forma completamente satisfactoria el problema en cuestión, debida principalmente a la necesidad de establecer un compromiso entre precisión y robustez de la convergencia, probándose por tanto la necesidad de una aportación original que eluda la necesidad de este compromiso.

Acto seguido, se estudiarán las distintas particularidades de la aplicación de registro que se va a abordar, razonando cómo afectan al método de registro que se desea diseñar.

Hecho esto, se propondrá razonadamente un sistema completo de registro elástico de imágenes multimodales que, adaptándose a las necesidades particulares de la aplicación, supere las limitaciones halladas en los métodos existentes. Se analizará el comportamiento de dicho sistema con casos clínicos reales, describiéndose los principales inconvenientes detectados, relacionados con la velocidad y robustez de la convergencia.

Finalmente, se presentarán las distintas técnicas de inicialización que se han desarrollado a lo largo de esta Tesis para compensar los inconvenientes arriba mencionados, comparando sus distintas prestaciones para concluir con una propuesta concreta.

### 3.1. Motivación

En los últimos años se ha producido una significativa proliferación de los sistemas de diagnóstico por la imagen basados en la tomografía de emisión de positrones (PET) en aplicaciones de oncología clínica. Dicha técnica permite apreciar las modificaciones metabólicas causadas por un cáncer, posibilitando la detección y evaluación de la extensión de tumores normalmente no visualizables por los exámenes morfológicos clásicos, como los basados en rayos X.

Las imágenes PET empleadas en oncología emplean como radiofármaco la 18-fluorodesoxiglucosa (FDG), una molécula que se metaboliza de forma análoga a la glucosa. De este modo, el examen *PET de emisión* muestra el nivel de absorción de glucosa por parte de los distintos tejidos, siendo este parámetro un claro indicador de la actividad tumoral.

Sin embargo, la resolución de estas imágenes, con voxels normalmente isotrópicos de varios milímetros de lado, queda muy lejos de las resoluciones submilimétricas ofrecidas por otras modalidades como la MRI o la CT. Debido a esta limitada resolución y a la naturaleza ruidosa de las imágenes PET, la información anatómica presente en estos estudios es confusa, imprecisa y de difícil interpretación.

Dado que no es posible en PET visualizar suficientes estructuras de referencia, la incorporación de datos adicionales que permitan una buena localización anatómica de la información ofrecida por la imagen es primordial para el análisis diagnóstico de numerosos tipos de cáncer.

Las imágenes CT muestran con detalle las distintas estructuras anatómicas presentes en la región torácica y abdominal. Presentan una excelente resolución en el plano axial (inferior al milímetro) que contrasta con las distancias entre cortes sucesivos (eje  $z$ ), que oscilan entre cinco y diez veces ese valor. Esta anisotropía, debida al tipo de adquisición por barrido, tiene un fuerte impacto en los algoritmos aplicados a este tipo de imágenes, como se verá más adelante en el capítulo dedicado a la segmentación de estructuras anatómicas.

En esta modalidad se detecta fácilmente el perfil del paciente, así como de los pulmones, bronquios, tráquea y tejido óseo. Los órganos blandos, como los riñones y el hígado, pueden visualizarse también, si bien con un contraste significativamente menor. Los tumores pueden ser más o menos visibles en función de su naturaleza, siendo sumamente dificultosa su localización, si bien una vez localizado es posible determinar con precisión su emplazamiento respecto a la anatomía circundante.

Las figuras 3.1 y 3.2 muestran sendos ejemplos de adquisiciones CT y PET típicas de las regiones torácica y abdominal. En este caso la representación de los volúmenes de datos se realiza por un mosaico de cortes axiales, muy comúnmente empleado por el personal médico para el diagnóstico mediante este tipo de imágenes. Nótese que ambos volúmenes de datos, si bien provenientes de un mismo paciente, no coinciden exactamente ni en la resolución ni en la talla de la región visualizada. Ello dificulta considerablemente la apreciación de las correspondencias entre las estructuras visualizadas en una y otra adquisición, complicando la localización precisa de las

patologías detectadas.

Dado que la información aportada por dichas modalidades es fundamental en aplicaciones de diagnóstico oncológica torácica (cáncer pulmonar) y abdominal (cáncer colo-rectal), es evidente el sumo interés del desarrollo de un método de registro aplicable al análisis multimodal de datos anatómico-funcionales de tales regiones.

El presente trabajo de tesis aborda por consiguiente el estudio y diseño de un método de registro de volúmenes CT-PET de la zona torácica y abdominal. El principal obstáculo que se presenta al desarrollo de dicho método se debe al carácter móvil de los órganos en dichas zonas. Dicha movilidad, junto con las condiciones de adquisición de las imágenes, impiden el empleo de cualquier solución de registro rígido entre ambas modalidades, obligando a recurrir a métodos, mucho más complejos, de registro elástico.

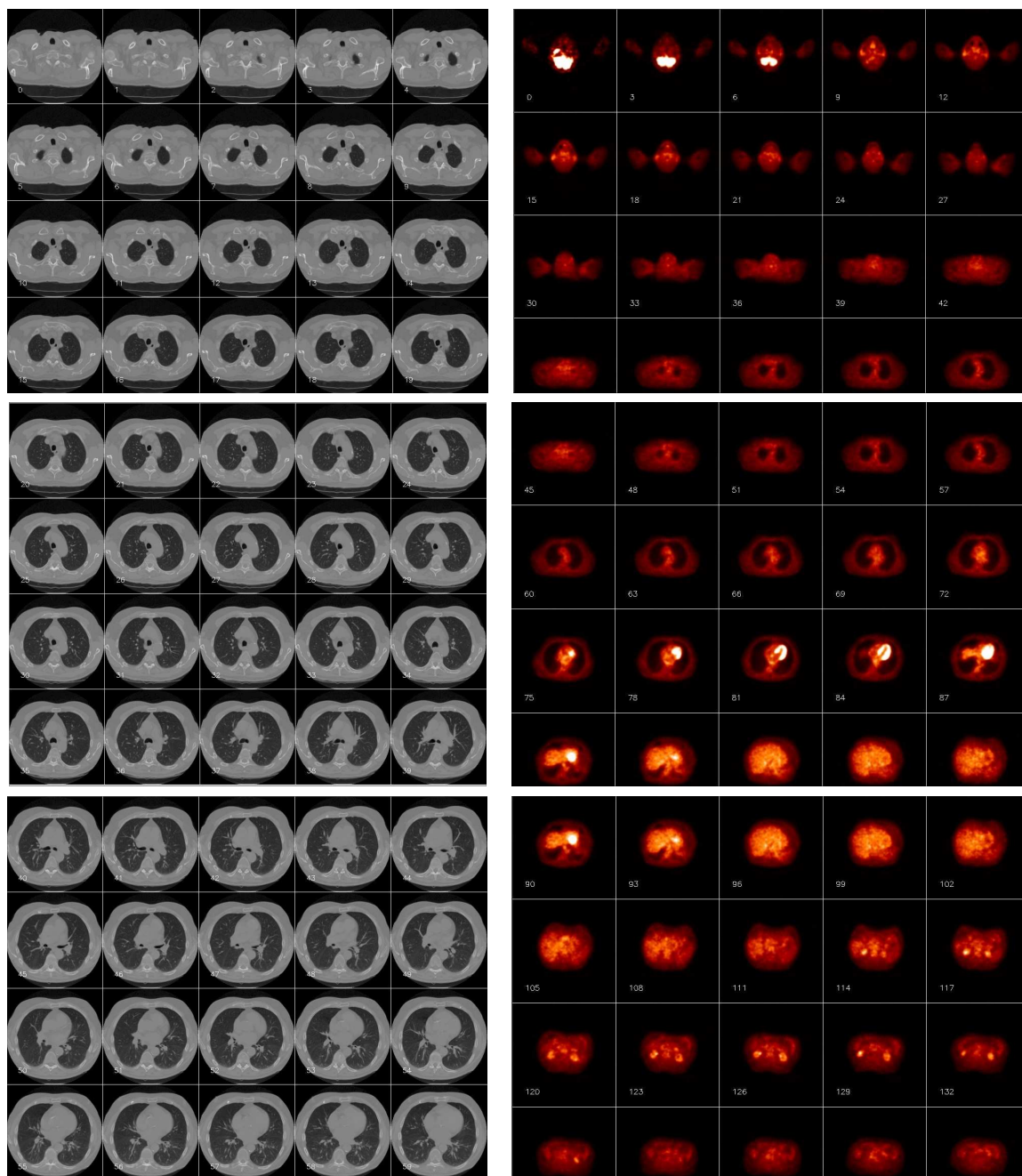
Tal como se detallará en la siguiente sección, dedicada al repaso de las distintas técnicas con las que se ha tratado de abordar anteriormente el problema que nos ocupa, precisamente a causa de la severidad de las deformaciones elásticas que se dan en las zonas torácica y abdominal no se ha conseguido hasta la fecha proponer un método que resuelva el problema de su registro de forma satisfactoria. Así, ni las aplicaciones basadas en transformaciones lineales alcanzan la precisión deseada en sus resultados, ni se ha conseguido que las aplicaciones basadas en transformaciones de mayor orden converjan de forma suficientemente robusta.

En parte ello se debe a un error de concepción por el cual, hasta muy recientemente, se consideraba que un método de registro podía basarse, de forma excluyente, bien en la segmentación de los datos, bien el uso directo de los mismos. La originalidad del método aquí propuesto consiste en lograr integrar en un mismo sistema ambas técnicas, de forma que se garantiza en todo momento a la vez una correcta convergencia y una buena precisión final.

Este es, por lo que hemos podido saber, uno de los primeros métodos publicados que combinan eficazmente ambos tipos de información, y el único que logra independizar completamente la precisión del resultado final de la calidad de la segmentación. Nótese la significativa diferencia que esto comporta frente a aquellos métodos que incorporan en un mismo criterio de similitud la información de intensidades y las referencias anatómicas, especialmente en un campo como es el tratamiento de imágenes médicas, donde la realización de una segmentación precisa de los datos es, en el mejor de los casos, compleja e incierta.

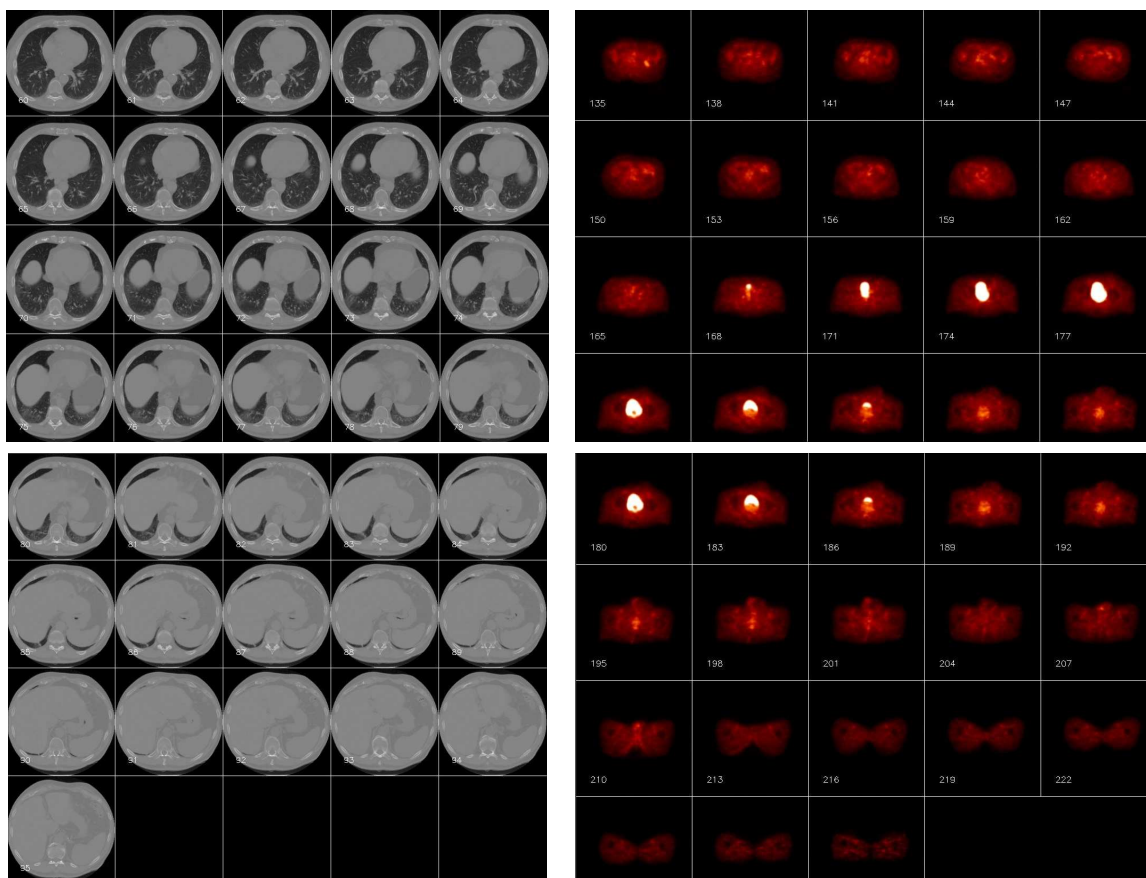
## 3.2. Antecedentes

A lo largo de este trabajo de tesis se ha efectuado un estudio extensivo del campo de la registro de imágenes médicas, del que ya se ha realizado un resumen en capítulos anteriores. Dado que el eje del trabajo de investigación ha sido el desarrollo de una metodología de registro de imágenes CT-PET para aplicaciones de diagnóstico y tratamiento en oncología de tórax y abdomen, se ha dedicado una especial atención al estado del arte en dicho ámbito, siguiéndose con interés los avances presentados por



**Figura 3.1:** Representación por secuencia de cortes axiales de un volumen CT (izquierda) y PET (derecha) correspondientes a las zonas torácica y abdominal. Cortes del 0 al 59 y del 0 al 132 respectivamente.





**Figura 3.2:** Representación por secuencia de cortes axiales de un volumen CT (izquierda) y PET (derecha) correspondientes a las zonas torácica y abdominal. Cortes del 60 al 95 y del 135 al 222 respectivamente.



otros grupos de investigación

Si bien durante los últimos años se ha apreciado una significativa proliferación de las técnicas de registro elástico, en contextos tales como el cerebro [23] [24] [25] [26], pecho [27] o corazón [28], dichas técnicas no son, en general, apropiadas para el registro de zonas con deformaciones locales tan severas y anatomía tan variable como son el tórax y el abdomen. Se sabe de muy pocos trabajos que hayan abordado el registro de dichas regiones, especialmente para el caso multimodal. A continuación se presenta una relación de las principales contribuciones en este sentido de las que se tiene noticia hasta esta fecha. Antes, sin embargo, es preciso mencionar una excepción importante a lo dicho. Se trata de las aplicaciones de registro de imágenes mamográficas. Aunque las particularidades de este tipo de aplicaciones hacen que la mayoría de métodos diseñados para ellas no sean trasladables a nuestro problema, existe sin embargo una cierta afinidad a nivel de la importancia de las deformaciones a compensar. Tanto es así que, como se verá más adelante, el esquema de deformación propuesto para nuestro sistema de registro se basa en el método desarrollado por Rueckert [27] para el registro de imágenes mamográficas.

En primer lugar cabe mencionar que algunos métodos consiguen realizar el registro de imágenes anatómicas y funcionales de tórax y abdomen mediante transformaciones lineales. Éste es el caso del programa *MIRIT*, de la *Katholieke Universiteit Leuven*, que emplea un algoritmo automático basado en la maximización de la información mutua para establecer un registro afín de las imágenes [29]. Éste método fue empleado con éxito para mejorar la localización de metástasis de nodos linfáticos, aunque, si bien las pruebas de evaluación indicaron claramente que la precisión en la lectura de las imágenes PET se incrementaba al combinarlas con imágenes CT, el registro automático de ambas aportaba tan sólo un beneficio marginal frente a la correlación visual de las mismas. Con toda probabilidad esto se debe a la incapacidad de la transformación afín de compensar las fuertes deformaciones locales debidas al reposicionamiento del paciente entre una y otra adquisición y a la respiración.

Existen asimismo varios estudios que realizan el registro anatometabólico de estructuras torácicas o abdominales a partir de un conjunto de referencias. Así, Caldwell et al [30] emplean una técnica de registro rígido basada en referencias para preparar una radioterapia de cáncer de pulmón *non small cell lung cancer (NSCLC)*, permitiendo así reducir significativamente el volumen estimado del cáncer *gross tumor value GTV*. Sato [31] propone una metodología basada en splines de Cauchy-Navier para efectuar un registro elástico del tórax a partir de un conjunto referencias. Mahmoud [32] realiza una comparación de tres métodos interactivos para el registro de imágenes CT-SPECT de la zona pélvica/abdominal, dos de ellos basados en referencias naturales y un tercero basado en la segmentación de estructuras anatómicas. En los tres casos se requieren de 10 a 30 minutos por parte de un usuario experimentado y con un buen conocimiento de la anatomía para localizar las referencias necesarias. Un problema de este tipo de enfoque es el gran número de referencias de las que es preciso disponer en ambas modalidades para poder compensar incluso las deformaciones más elementales, quedando además el resultado final limitado por la precisión que se consigue en su

localización.

Dentro de los métodos basados en la segmentación de las estructuras a registrar, las principales aplicaciones se centran en órganos grandes tales como los pulmones y el hígado. Fleming et al. [33] proponen una transformación esférica para proyectar imágenes de la distribución de radionúcleos inhalados en forma de aerosol sobre un modelo estándar de los pulmones. Chung et al. [34] emplean un alineamiento de los ejes principales seguido de una adaptación de superficies para el registro de imágenes CT-SPECT del hígado. Masutani et al. [35] combinan una representación modal de las transformaciones con Free Form Deformations para registrar imágenes del hígado en una aplicación de cirugía guiada por computador. Como en el caso de las aplicaciones de registro basadas en referencias, el obstáculo principal al que se enfrentan este tipo de métodos es la propia segmentación de las estructuras anatómicas. En efecto, esta es una tarea compleja que se ve enormemente agravada si se debe garantizar que dicha segmentación sea equivalente en imágenes de modalidades distintas.

Son pocos los métodos que hayan abordado el registro de este tipo de imágenes basándose en medidas que empleen íntegramente la información de intensidades, en parte debido al gran volumen de datos a manipular. Tai et al. [36] recurren a un polinomio de segundo grado para efectuar una corrección de las intensidades de las imágenes, para así poder utilizar un criterio de error cuadrático en el registro de imágenes CT-PET del tórax. La deformación se consigue por interpolación de desplazamientos obtenidos para un conjunto de subbloques de la imagen a registrar. El problema de este método es que todavía se supone una relación funcional entre las intensidades, lo que no siempre es cierto. Una solución elegante para el registro de imágenes torácicas multimodales es la propuesta por Mattes [37], que emplea FFD's controladas mediante el criterio de la información mutua. La estimación del histograma se realiza mediante ventanas de Parzen, y el proceso de optimización utiliza una estrategia multiresolución. El principal inconveniente de este método es que la optimización es propensa a caer en mínimos locales, por lo que no siempre se consiguen compensar las deformaciones locales más importantes, como las que se encuentran en la zona del diafragma a causa de la respiración.

Por último, es preciso mencionar una opción que ha estado cobrando una importancia creciente en el campo de la oncología médica. Se trata de la adquisición simultánea de imágenes PET y CT mediante un tomógrafo combinado [38]. Hoy en día son ya varias las empresas de dispositivos médicos que ofrecen este tipo de máquinas, siendo incluso posible equipar un tomógrafo PET o CT existente con su contrapartida, de cara a poder realizar estudios combinados. Las adquisiciones de este tipo resuelven eficazmente el problema del registro rígido de las imágenes, así como las deformaciones elásticas debidas al reposicionamiento del paciente, quedando sin embargo todavía sin compensar los artefactos metabólicos debidos a los distintos tiempos de adquisición. Hany et al. [39] informan de una mejora en la precisión en el diagnóstico del cáncer de pulmón (NSCLC) empleando una máquina de este tipo. El principal inconveniente de este tipo de solución, es, sin lugar a dudas, su precio. En efecto, el coste de adquirir una máquina combinada, que en modo alguno puede suplir la presencia de un escáner

CT y otro PET trabajando paralelamente, a menudo en departamentos distintos de un mismo centro, hace pensar que su introducción al ámbito del diagnóstico en oncología, de tener éxito, se producirá de forma muy paulatina y dispersa.

En conclusión, en los últimos años se ha podido apreciar un interés creciente por el registro elástico de imágenes, con una proliferación de las aplicaciones que emplean este tipo de técnicas, si bien han sido pocos los trabajos consagrados al registro de las zonas torácica y abdominal, especialmente, como requiere el caso que nos ocupa, de forma automática y sobre imágenes multimodales. Adicionalmente, la mayor parte de los métodos aplicados sobre estas regiones dependen fuertemente de una segmentación de la anatomía, a menudo restringiéndose al registro de un único órgano, siendo inadecuadas para su uso sobre zonas más extensas y de mayor variabilidad. Es cierto que se han publicado unos pocos trabajos, citados arriba, que reúnen las condiciones que se precisan para su aplicación en diagnosis oncológica, pero ninguno de ellos es, en nuestra opinión, capaz de ofrecer un buen compromiso entre capacidad de compensar deformaciones severas y precisión del resultado final.

La principal aportación del presente trabajo de tesis ha sido el desarrollo de un método de registro elástico de imágenes multimodales capaz de superar dicha limitación, como respaldan los resultados que se detallan en el capítulo 5.

### 3.3. Condicionantes del método

Considérense pues los requisitos del método de registro que se desea desarrollar, en función de las particularidades de la aplicación. Es importante establecer estos de forma clara y precisa, previo a realizar cualquier propuesta, de cara a ajustarse lo mejor posible a las necesidades de cada sistema. A dicho fin resulta sumamente útil emplear los criterios propuestos por Van den Elsen, presentados en el anterior capítulo.

A continuación se detallan, para cada uno de dichos criterios, las condiciones particulares de nuestra aplicación, sobre las que se tendrá que construir el método de registro:

#### Dimensionalidad de las imágenes

Se trabajará con volúmenes tridimensionales de datos, provenientes de estudios PET y CT de la zona torácica y abdominal, como las que se muestran en la figura 3.3. Sus dimensiones típicas serán:

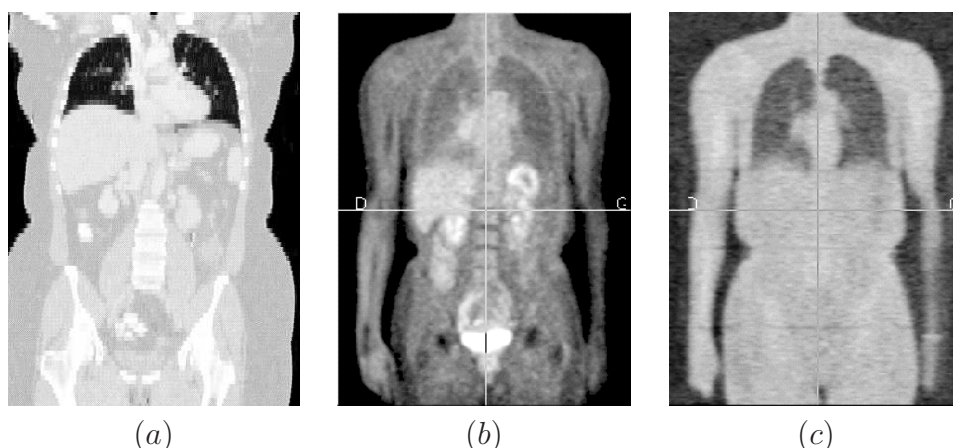
- *CT*

Resolución: de (256x256x40) a (512x512x65)voxels

Dimensiones de los voxels: de (0.8x0.8x4.0) a (1.5x1.5x8.0)mm

- *PET*

Resolución: de (128x128x63) a (256x256x171)voxels



**Figura 3.3:** Modalidades implicadas en el registro: Corte coronal de un volumen CT (a); corte coronal de un volumen PET de emisión (b); corte coronal de un volumen PET de transmisión (c).

Dimensiones de los voxels: de (3.9x3.9x4.0) a (4.3x4.3x4.3)mm

En todos los casos la dinámica de las imágenes se limitará a *8 bits*, si bien en algunos estudios se dispone de datos de *16 bits*. En cualquier caso los distintos métodos presentados en este trabajo de tesis son fácilmente generalizables para aprovechar dicha información adicional.

### Tipo de transformación

Las principales diferencias de posición entre las imágenes CT y PET se pueden atribuir a tres causas:

En primer lugar, en la gran mayoría de casos la adquisición de las imágenes se realizará en emplazamientos distintos, mediando un cierto tiempo entre ambos estudios. Será pues imposible reproducir en ambos casos la posición del paciente y su alineamiento con la máquina.

Por otra parte, los protocolos de adquisición pueden diferir entre las distintas modalidades. Un ejemplo que se ha encontrado con cierta frecuencia es la posición de los brazos del paciente, levantados para las imágenes CT y pegados al cuerpo para las PET. Estas discrepancias de posición comportan distintas deformaciones de las estructuras anatómicas internas que será preciso compensar, siendo éstas especialmente severas dada la naturaleza elástica de los órganos en las regiones torácica y abdominal.

Finalmente, debido a los distintos tiempos de adquisición (unos treinta segundos para las imágenes CT y unos treinta minutos para las imágenes PET) es inevitable que la actividad metabólica produzca deformaciones de las estructuras anatómicas que se desea estudiar. El ejemplo más claro de esto, de gran importancia en el estudio de tumores pulmonares, es la posición del diafragma. En efecto, la adquisición de

la imagen CT puede realizarse mientras el paciente bloquea la respiración, sea en un estado de máxima inspiración, sea en una posición media. Por contra, el tiempo de adquisición de las imágenes PET exige que ésta se realice mientras el paciente respira con normalidad, por lo que la escena final reflejará un promediado del conjunto de la actividad respiratoria. Estas deformaciones locales son extremadamente severas, pudiendo llegar a comportar discrepancias de *más de 5cm* entre una y otra imagen.

Por todo lo expuesto, parece evidente que para efectuar el registro de estas imágenes será preciso recurrir a una *transformación no lineal* con capacidad para compensar deformaciones locales muy severas. Se trata, por otra parte, de un modelo de deformación que ya ha sido empleado con éxito en otras aplicaciones de registro de imágenes médicas, y cuya eficiencia le está ganando una merecida popularidad en este ámbito.

El gran volumen de datos a manejar y los requisitos de rapidez de esta aplicación (sería deseable lograr un tiempo de registro inferior al tiempo de adquisición, para no introducir un "cuello de botella" en el proceso de diagnóstico) descartan la posibilidad de usar deformaciones basadas en modelos físicos, por ser éstas demasiado costosas, así como aquellas transformaciones compuestas por funciones cuyo soporte no es local (véase el apartado 2.3.2).

Se ha optado pues por una transformaciones mediante *free form deformations (FFD)* basada en *B-splines*, ya que su separabilidad y soporte local permiten una implementación eficiente a la vez que ofrecen una buena adaptabilidad a deformaciones locales severas.

## Medida de similitud

Dada la naturaleza interna de las estructuras que se desea estudiar y a la naturaleza local de las deformaciones, resulta imposible recurrir a ningún tipo de elemento extrínseco de referencia.

La detección de referencias anatómicas no es tampoco una opción viable, debido por una parte a la gran cantidad de las mismas que se precisaría para efectuar un registro elástico, y por otra parte a la gran diferencia entre las modalidades con las que se debe trabajar, que dificulta en extremo la detección de una misma referencia en una y otra imagen.

En lo que respecta a las medidas de similitud basadas en la segmentación, se ha descartado su uso por razones de precisión y robustez. En efecto, la calidad de este tipo de medidas se ve limitada por la calidad de la segmentación que se consiga. Esto presenta dos problemas importantes en la aplicación que nos ocupa.

Por un lado, si bien es posible localizar estructuras correspondientes en ambas modalidades (piel, esqueleto, pulmones, riñones, . . .), y la segmentación de la mayoría no presenta excesivos problemas en la CT, hablar de una segmentación automática, robusta y precisa de dichas estructuras en las imágenes PET es, a nuestro entender, poco realista.

Por otra parte, está el problema de establecer un registro a partir de estructuras

segmentadas en modalidades distintas, segmentación para la que se habrán debido definir ciertos umbrales y establecer criterios cuya equivalencia en ambas modalidades no es posible garantizar.

Por todo lo anterior, se ha optado por una medida basada en el contenido íntegro de las imágenes. Dada la relación no funcional existente entre los niveles de gris de las modalidades a registrar, la medida más apropiada es la *información mutua normalizada* (*NMI*).

### Procedimiento de optimización

El cálculo del gradiente de la información mutua normalizada es un proceso complejo, sumamente lento y cuyos resultados dependen de un conjunto de suposiciones realizadas sobre las imágenes. Por esta razón se ha optado por aplicar un método de optimización que emplee una estimación local del mismo.

Así pues, la optimización puede realizarse mediante un método de *steepest descent*, actualizándose simultáneamente el conjunto de los puntos de control de la FFD. Para evitar alteraciones de la topología de la imagen, debe introducirse un término de regularización que impida el cruce de dicha malla de puntos de control

### Modalidades de imagen implicadas

Se trata de una aplicación de registro multimodal entre imágenes de *tomografía computarizada por rayos X (CT)* y *tomografía de emisión de positrones (PET)*.

En los volúmenes de datos CT el contraste entre los distintos tejidos y el aire es excelente, por lo que se detecta fácilmente el perfil del paciente, así como de los pulmones, bronquios y tráquea. El tejido óseo también presenta un buen contraste, por lo que es posible visualizar con facilidad la columna vertebral y la caja torácica, si bien la reconstrucción de esta última presenta ciertos problemas debidos a la mala resolución entre cortes. Los órganos blandos, como los riñones y el hígado, pueden visualizarse también, si bien con un contraste significativamente menor. Esto, junto con la mencionada mala resolución del eje  $z$ , dificulta en ciertos casos la determinación de ciertos umbrales, como el que separa la parte superior del hígado del corazón. Los tumores pueden ser más o menos visibles en función de su naturaleza, siendo sumamente dificultosa su localización, si bien una vez localizado es posible determinar con precisión su emplazamiento respecto a la anatomía circundante.

Por contra, los volúmenes de datos PET ofrecen poca información anatómica, puesto que las intensidades van asociadas al nivel de actividad metabólica del tejido, más concretamente al consumo de glucosa. En estos estudios se puede localizar con suma facilidad cualquier tumor, así como la vejiga y ciertos ganglios. El perfil del paciente, así como los pulmones, riñones y en cierta medida el esqueleto y el hígado, son también visibles aunque de forma ruidosa y poco contrastada.

Una posibilidad poco empleada de la que se saca partido en este estudio es emplear también las imágenes *PET de transmisión*. Estas imágenes del coeficiente de atenuación a los rayos gamma de los tejidos son un resultado intermedio de la adquisición PET, que



las emplea para efectuar correcciones sobre la imagen de atenuación. En ellas se puede apreciar, de forma bien contrastada, tanto el perfil del paciente como los pulmones, si bien la resolución raramente permite apreciar la tráquea o los bronquios.

### Sujeto visualizado

Dado que lo que se desea es combinar información de un mismo paciente proveniente de distintas modalidades se trata de una aplicación de registro intrasujeto. No cabe por tanto esperar variabilidad anatómica más allá de las deformaciones descritas anteriormente.

### Objeto visualizado

Este estudio se centra en el registro de las regiones torácica y abdominal para aplicaciones de diagnosis y tratamiento en oncología.

Las estructuras anatómicas más significativas que visibles en los estudios CT y PET con los que se trabajará son: la caja torácica, columna vertebral, pulmones, corazón, hígado, riñones, vejiga. La figura 3.4 muestra algunas de estas estructuras en sendos volúmenes CT y PET.

Los tipos de cáncer a cuya diagnosis se destina esta aplicación son principalmente el cáncer de pulmón, de esófago, gastrointestinal, de riñón, de hígado y de páncreas.

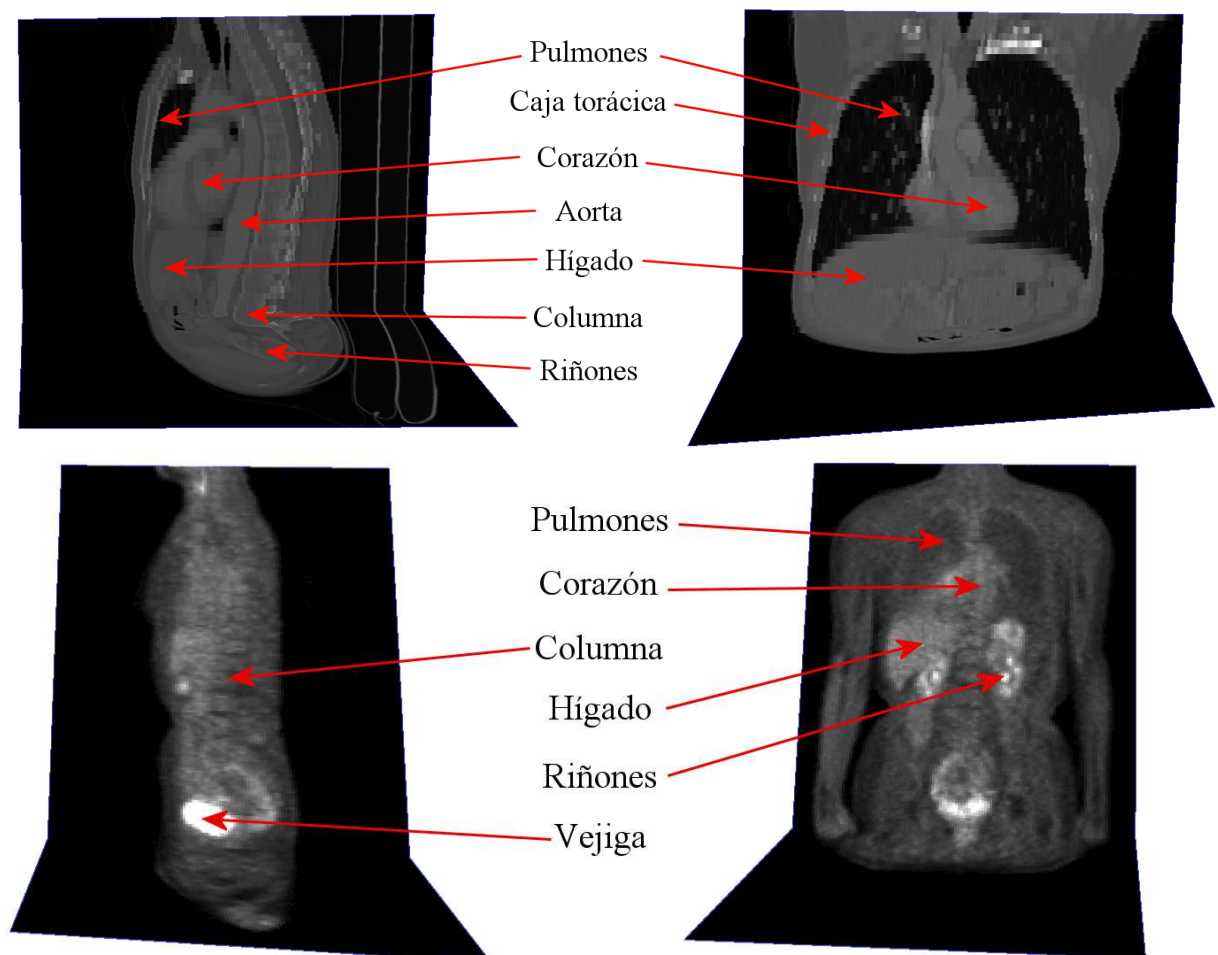
### Necesidad de interacción

Al tratarse de una aplicación de diagnosis a partir de un conjunto de imágenes cuya adquisición se realiza de forma separada, mediando un tiempo considerable, no se precisa que el registro se efectúe en tiempo real. De acuerdo con el personal médico especializado, se ha establecido que el tiempo máximo del que dispone la aplicación para realizar el registro de dos volúmenes de datos es el periodo necesario para realizar una nueva adquisición, de forma que esta etapa no represente un cuello de botella en el proceso de diagnosis.

Si bien el sistema trabaja de forma *totalmente automática*, la propia definición de la aplicación implica una *supervisión de los resultados* por parte del oncólogo especialista que deberá realizar el diagnóstico a partir de la imagen fusionada.

## 3.4. Método propuesto

Cuando, durante el transcurso de esta Tesis, se planteó por primera vez la posibilidad de abordar el registro de volúmenes PET-CT del tórax, no éramos aún conscientes de hasta que punto la gravedad de las deformaciones elásticas en dicha zona complicaba el problema. Por consiguiente, los primeros intentos de establecer el registro se llevaron a cabo empleando dos paquetes de software, el AIR y el ANIMAL, conocidos por ofrecer buenos resultados en el alineamiento de imágenes cerebrales multimodales.



**Figura 3.4:** Estructuras anatómicas típicas visualizadas en un volumen CT y PET respectivamente.



El AIR (de *Automated Image Registration*) consiste en realidad un conjunto de herramientas de software desarrolladas en la Universidad de California (UCLA), inicialmente para el registro de volúmenes PET monomodales, pero posteriormente generalizado para su uso con imágenes PET-MRI y PET-CT [7]. Por desgracia las transformaciones lineales que es posible realizar con el AIR, si bien conseguían alcanzar un registro global satisfactorio de las imágenes, distaban mucho de resultar suficientes para compensar las deformaciones debidas a la respiración.

El ANIMAL (de *Automatic Nonlinear Image Matching and Anatomical Labeling*) es un método destinado al registro de volúmenes cerebrales MRI con un volumen de referencia de esta misma modalidad, con el fin de segmentar ciertas estructuras anatómicas [40]. El interés de esta técnica para nuestros objetivos consistía en su capacidad de compensar deformaciones no lineales entre las imágenes. Sin embargo este método se reveló excesivamente específico para imágenes cerebrales, no siendo apto para trabajar con las deformaciones más severas presentes en nuestra aplicación.

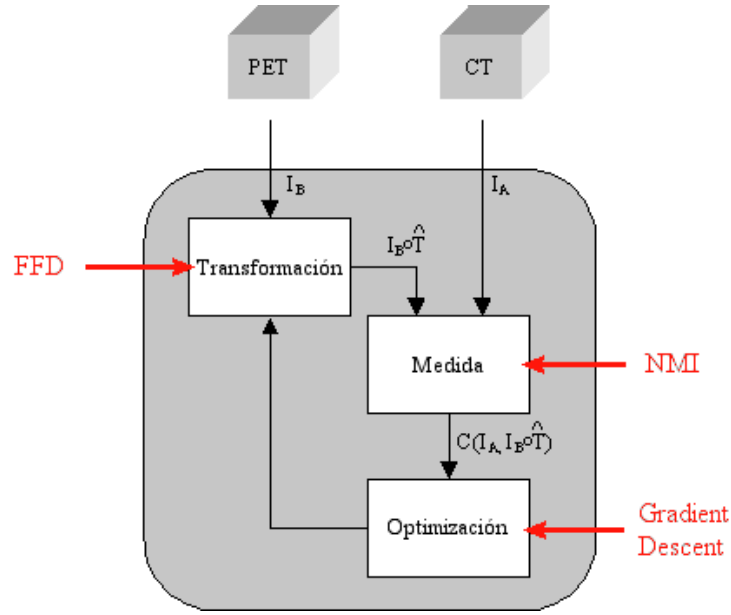
Se realizaron finalmente ciertas pruebas para adaptar al registro elástico algunas herramientas de registro multimodal basadas en la información mutua, que habíamos desarrollado con anterioridad para ciertas pruebas de registro rígido. Con este fin se optó por realizar un registro por bloques de las imágenes, interpolándose los resultados del mismo para simular las deformaciones elásticas que se deseaba compensar. Si bien se alcanzaron algunos resultados prometedores, esta técnica carecía de elegancia en el modelado de la transformación elástica, y presentaba serios problemas cuando se reducía excesivamente el tamaño de los bloques.

A medida que se trabajaba en anteriores pruebas, se nos fue revelando toda la problemática asociada a la aplicación, tal y como se ha resumido en las particularidades del problema detalladas en 3.3. Es a partir de este conocimiento que se ha trazado el esquema básico de registro elástico que presentamos a continuación. Este diseño, esquematizado en la figura 3.5, se ha basado en el método empleado por Rueckert et al. [27] para el registro elástico de imágenes mamográficas MRI.

Como se puede ver en la figura, la transformación elástica se ha modelado mediante *free form deformations (FFD)*. Esta técnica permite deformar la imagen a través de una malla regular de *puntos de control*, las traslaciones de los cuales son interpoladas para conseguir una transformación continua  $C^2$ . Esta interpolación puede expresarse en forma de producto tensorial de B-Splines cúbicas monodimensionales. Sea  $\Phi$  una malla tridimensional uniforme de  $n_x \times n_y \times n_z$  puntos de control  $\phi_{i,j,k}$  con una separación  $\delta$  entre ellos, donde  $-1 \leq i < n_x - 1, -1 \leq j < n_y - 1, -1 \leq k < n_z - 1$ . Así, el desplazamiento de cada voxel  $x,y,z$  de la imagen se calcula como:

$$T_{elast}(x, y, z) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 \theta_l(u) \theta_m(v) \theta_n(w) \phi_{i+l, j+m, k+n} \quad (3.1)$$

donde,  $i = \lfloor \frac{x}{n_x} \rfloor - 1$ ,  $j = \lfloor \frac{y}{n_y} \rfloor - 1$  y  $k = \lfloor \frac{z}{n_z} \rfloor - 1$  indican el punto de control que afecta al voxel  $(x, y, z)$ , y  $u = \frac{x}{n_x} - \lfloor \frac{x}{n_x} \rfloor$ ,  $v = \frac{y}{n_y} - \lfloor \frac{y}{n_y} \rfloor$ ,  $w = \frac{z}{n_z} - \lfloor \frac{z}{n_z} \rfloor$  son las posiciones relativas del voxel respecto a dicho punto ( $\lfloor \cdot \rfloor$  indica un redondeo al entero inferior más



**Figura 3.5:** Esquema del método de registro propuesto, como particularización del que se mostraba en la figura 2.1.

próximo). Los polinomios  $\theta_0$  a  $\theta_3$  constituyen una B-Spline monodimensional definida por tramos como sigue:

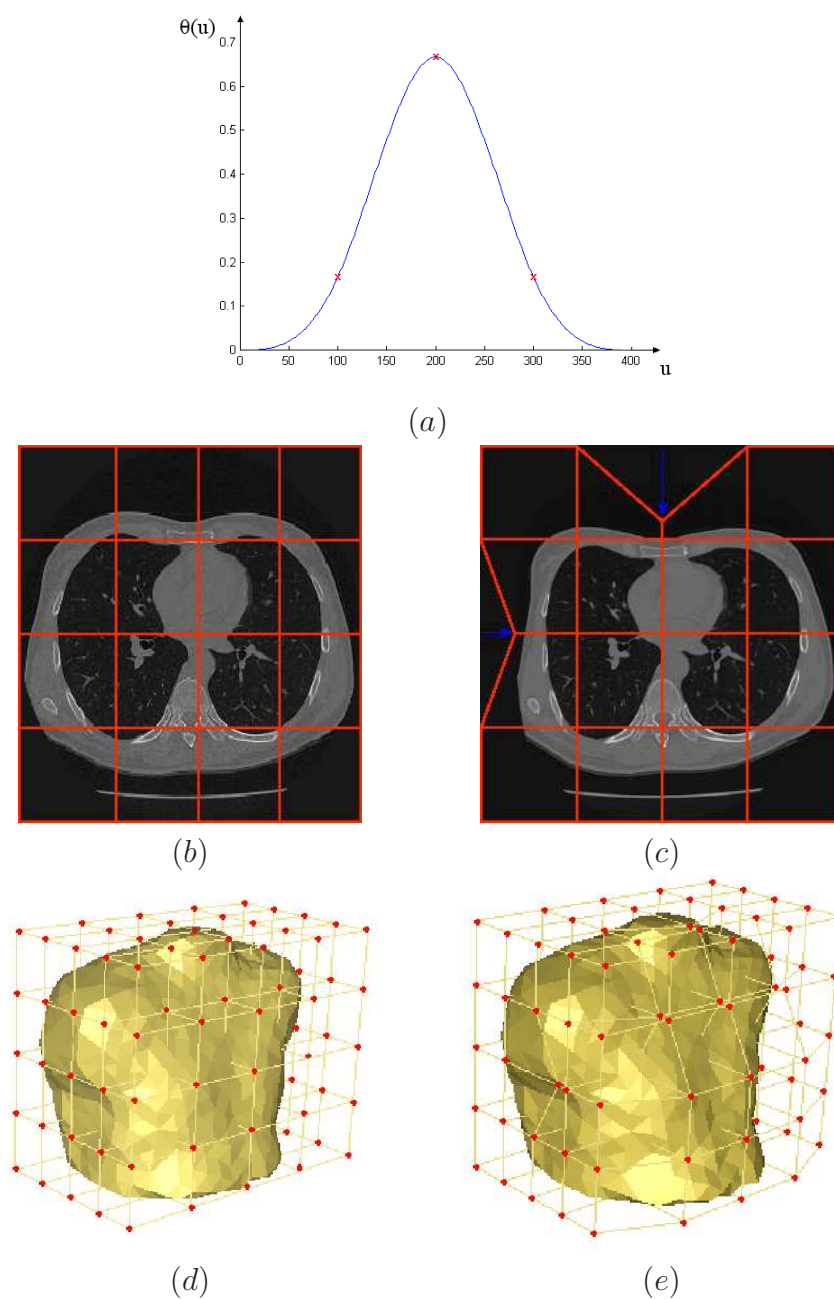
$$\begin{aligned}
 \theta_0(u) &= (1 - u)^3/6 \\
 \theta_1(u) &= (3u^3 - 6u^2 + 4)/6 \\
 \theta_2(u) &= (-3u^3 + 3u^2 + 3u + 1)/6 \\
 \theta_3(u) &= u^3/6.
 \end{aligned} \tag{3.2}$$

Este tipo de transformación, gracias a su separabilidad y al hecho de emplear funciones con soporte local, puede implementarse de forma sumamente eficiente mediante una tabla. Adicionalmente, dado que el desplazamiento de un punto de control afecta tan sólo un volumen limitado a su entorno, el proceso de optimización se simplifica considerablemente.

La figura 3.6 ilustra el funcionamiento de una deformación FFD. Nótese como el desplazamiento de cada pixel (o voxel) corresponde a una interpolación B-Spline de los desplazamientos de los puntos de control entorno a éste.

La medida de similitud que se emplea como criterio de la bondad del registro es la *información mutua normalizada (NMI)* [15], una modificación de la *información mutua (MI)* [16] que permite independizar ésta de la cantidad de solapamiento existente entre las imágenes.

El hecho de trabajar con imágenes cuyas distribuciones de intensidad no guardan una relación funcional entre sí obliga a recurrir a este tipo de medidas de la dependencia estadística de los datos, a pesar de su elevado coste computacional y la considerable



**Figura 3.6:** Funcionamiento de las FFD: B-Spline cúbica definida por tramos (a); ejemplo de deformación FFD 2D de un corte axial de un volumen CT torácico (b)(c); deformación FFD 3D de todo el volumen CT torácico, representado para mayor claridad mediante un renderizado de la piel del paciente (d)(e).

dificultad que comporta su optimización debido a la presencia de numerosos mínimos locales.

La información mutua normalizada se expresa como:

$$NMI(I_A, I_B) = \frac{H(I_A) + H(I_B)}{H(I_A, I_B)} \quad (3.3)$$

siendo  $H(I_A)$  y  $H(I_B)$  las entropías marginales de las imágenes y  $H(I_A, I_B)$  su entropía conjunta. La estimación de la función de distribución de probabilidad conjunta de las intensidades de las imágenes  $p(i, j)$  que es precisa para el cálculo de las entropías se calcula por aproximación frecuencial.

Como proceso de optimización se emplea un descenso de gradiente iterativo simultáneamente para todos los puntos de control de la FFD. El gradiente se consigue mediante una estimación local en un entorno de cada punto de control.

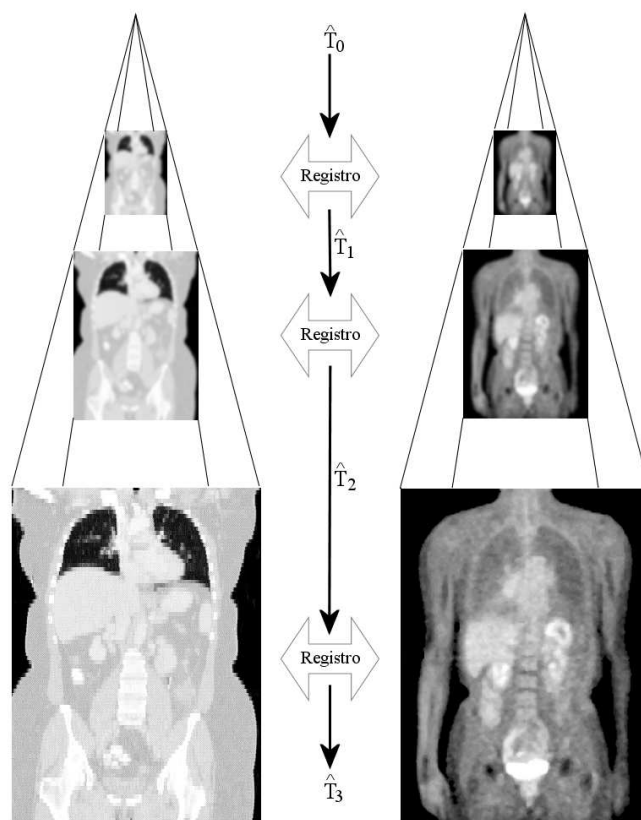
Con el fin de ampliar el rango de convergencia de la aplicación se ha integrado todo el proceso dentro de una estructura multiresolución. Así, la optimización se inicia sobre una versión filtrada y submuestreada de las imágenes, empleándose las transformaciones obtenidas como inicialización para una sucesión de etapas de resolución creciente. Esta técnica, muy comúnmente empleada en las aplicaciones de registro basadas en intensidades, se ilustra en la figura 3.7. Nótese como, a partir de una estimación inicial  $\hat{T}_0$  de la transformación entre las imágenes, se efectúa sucesivamente el registro de versiones de las mismas de resolución creciente, propagándose los resultados de cada etapa como inicialización de la siguiente.

El principal inconveniente que presenta este método es la excesiva flexibilidad de la FFD, que a menudo queda atrapada en mínimos locales de la medida de similitud. Un modo de evitar este problema es inicializar la optimización lo suficientemente cerca del resultado final. Sin embargo la presencia de fuertes deformaciones de carácter local impide, en general, conseguir tal inicialización mediante ninguno de los métodos empleados habitualmente.

En este trabajo de tesis se propone emplear un método de registro basado en la segmentación de estructuras anatómicas para alcanzar dicha inicialización. De este modo se cumplen dos objetivos: Por una parte se garantiza en gran medida la convergencia del sistema hacia el estado de correcto registro de las imágenes, robusteciéndolo frente a la presencia de mínimos locales. Por otra parte, el hecho de iniciar la optimización en la proximidad del resultado final reduce enormemente el número de iteraciones que se precisarán para alcanzarlo, disminuyendo por consiguiente la cantidad de costosas evaluaciones de la NMI que se realizarán.

### 3.5. Inicialización mediante referencias

Tal como se ha visto en el apartado anterior, el principal problema al que se enfrenta el sistema de registro propuesto es la dificultad de garantizar una correcta convergencia hacia la solución deseada. Sin embargo, es un hecho bien conocido dentro del ámbito



**Figura 3.7:** Esquema conceptual de un registro multiresolución.

de la optimización que el mejor modo de asegurar una buena convergencia del sistema es mediante una adecuada inicialización del mismo.

No obstante, la aplicación de este sencillo principio al problema que nos ocupa no resultó en absoluto trivial. Por supuesto, la mayoría de los métodos de registro existentes empleaban algún tipo de inicialización, siendo los más comunes la inicialización directa por parte del usuario y el alineamiento de centros de gravedad y momentos de inercia. Pero, cómo conseguir una inicialización de la transformación elástica tridimensional? Por su complejidad era ésta una tarea que no se podía confiar al usuario, y por otra parte nuestro objetivo era un método de registro totalmente automático.

Poco a poco se hizo pues evidente que la inicialización de nuestro método de registro requería, precisamente, de otro registro. Inicialmente se intentó resolver esto basando la inicialización en un registro afín de las imágenes, pero por desgracia este tipo de transformación resultó ser insuficiente para inicializar adecuadamente las regiones de mayor deformación. Así pues, se optó finalmente por desarrollar un esquema totalmente nuevo de inicialización, que constituyera en sí mismo un método de registro elástico basado en la información anatómica disponible en las imágenes.

Por consiguiente, el objetivo de esta etapa es sacar el máximo partido posible a la información anatómica presente en las imágenes, para de este modo restringir la búsqueda de la transformación óptima que se llevará a cabo en la siguiente etapa. En consecuencia, la efectividad de la presente fase estará fuertemente condicionada por la cantidad de información anatómica que se pueda extraer simultáneamente de ambos volúmenes de datos.

En aras de la claridad, nos referiremos a esta etapa como *inicialización*, dado que su objetivo es establecer una estimación inicial de los parámetros de la transformación, con el fin de facilitar la convergencia del registro basado en la información mutua. Que este término no lleve a confusión: las técnicas que se emplearan constituyen por sí mismas un sistema de registro elástico, y distan mucho del simple alineamiento de momentos de inercia tan comúnmente empleado por otros métodos.

Es importante no confundir esta etapa con el procedimiento que siguen los métodos de registro basados exclusivamente en la segmentación de las imágenes, puesto que, en nuestro caso, los errores introducidos en este punto por la segmentación no se propagarán al resultado final, ya que la posterior etapa de registro basada en las intensidades se encargará de corregirlos.

Por lo tanto, no se trata aquí de abordar la formidable tarea de implementar un sistema capaz de realizar de forma totalmente automática una buena segmentación de los datos CT y PET, sino que bastará con alcanzar una aproximación razonablemente precisa de aquellas estructuras visibles en ambos volúmenes. En consecuencia, antes que la precisión se debe priorizar la rapidez y, por encima de todo la fiabilidad del sistema.

A lo largo del presente trabajo de tesis se han desarrollado dos métodos que emplean enfoques distintos para establecer una inicialización a partir del registro de imágenes segmentadas.

El primero de ellos parte de un registro rígido entre las superficies de los órganos,

empleando para ello un método conocido como es el *iterative closest point (ICP)*, para luego estimar la deformación elástica entre estas mediante un método de proyección.

El segundo método se ha desarrollado para dar respuesta a algunos de los problemas que el anterior presenta mediante una mejor integración dentro del sistema global de registro.

Si bien este último diseño ha demostrado ser superior en muchos sentidos a la inicialización basada en el ICP, se ha decidido presentar a continuación ambos métodos, dado que la solución final seleccionada surge en gran medida como respuesta a un conjunto de limitaciones, descubiertas durante el desarrollo del primer método, y que es interesante revisar.

### 3.5.1. Método 1: Proyección de superficies.

- Paso 1 - Localización de estructuras anatómicas

Ambos métodos de inicialización trabajan sobre aquellas estructuras anatómicas que puedan segmentarse en ambas imágenes. en el caso que nos ocupa las principales estructuras de interés son los pulmones, hígado y riñones.

Dada la importancia del proceso de segmentación, se ha dedicado a ella un capítulo a parte 4, donde se describen con detalle las distintas técnicas empleadas para localizar y extraer las diferentes estructuras anatómicas de interés.

El resultado de esta etapa son dos volúmenes de datos binarios, de dimensiones idénticas a las CT y PET originales, donde se representan los distintos órganos localizados como regiones conexas y sin agujeros.

- Paso 2 - Extracción de la región deseada

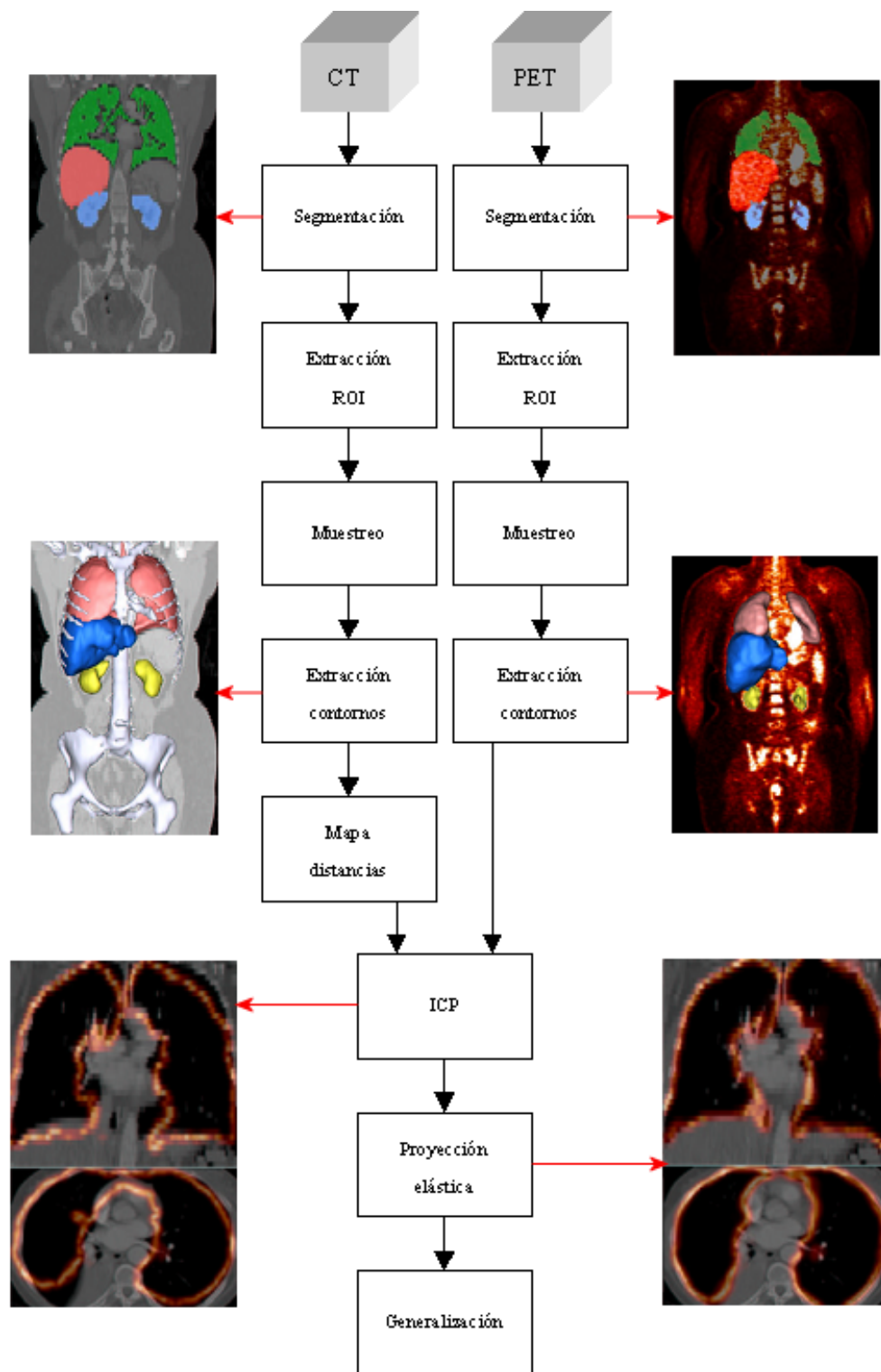
Una vez se tienen las distintas estructuras anatómicas a registrar, es posible restringir el volumen de trabajo a una caja englobante ajustada a dichas estructuras.

Esta operación cumple una doble función. Por una parte se reduce considerablemente el tiempo de cálculo requerido por las posteriores etapas, y por otra se consigue independizar el análisis de la escala de las imágenes. En efecto, el hecho de ajustar los volúmenes a unos órganos cuyas dimensiones son aproximadamente las mismas en ambas imágenes, equivale implícitamente a realizar una corrección de escala, cosa que nos permite en adelante eliminar tres de los grados de libertad del problema y efectuar un primer registro rígido mediante una técnica relativamente simple como es el ICP.

- Paso 3 - Muestreo de los volúmenes a resolución mínima

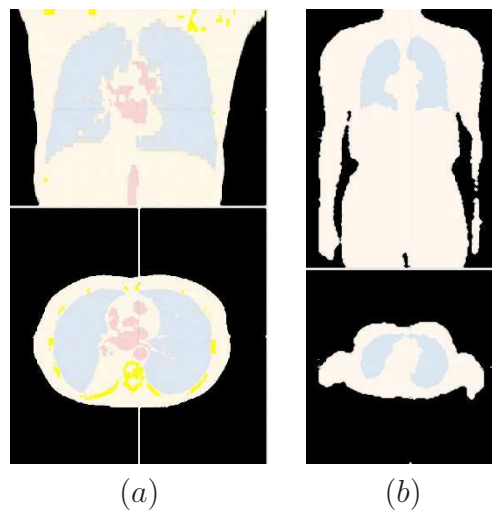
Dado que la calidad del registro que se puede conseguir queda limitado principalmente por la imagen de menor resolución, para simplificar y acelerar aún más los siguientes pasos se reduce la resolución espacial de las imágenes. Con tal fin, se modifica para cada eje la frecuencia de muestreo del volumen





**Figura 3.8:** Esquema del primer método de inicialización propuesto. Es preciso mencionar que la segmentación del hígado que se muestra en las figuras se ha conseguido gracias a los métodos desarrollados por O.Camara [41] y O.Colliot [42] para sus respectivos trabajos de tesis.





**Figura 3.9:** Ejemplo de estructuras anatómicas segmentadas para el registro en oncología de tórax.

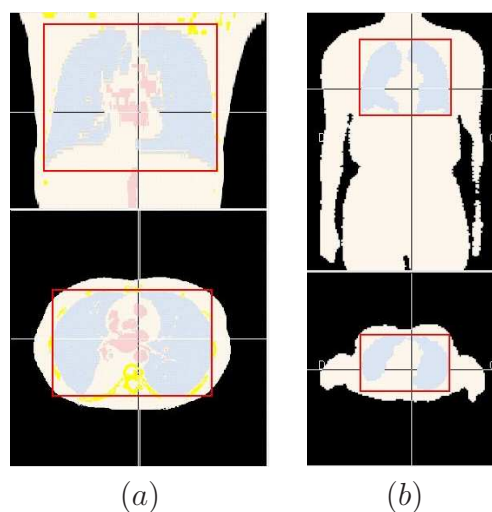
con mayor resolución, de modo que su número de voxels coincida con el de la resolución menor. Mediante esta reducción se consigue prescindir de gran cantidad de información de la que de cualquier modo se hubiera podido sacar poco o ningún partido, al no tener contrapartida en el otro volumen.

Una importante ventaja de efectuar este muestreo es que permite en adelante trabajar de forma independiente de las dimensiones de los voxels, suponiendo que son iguales y unitarios en ambos volúmenes. Así, la mayoría de cálculos realizados a partir de este punto pueden realizarse en enteros, mejorando por consiguiente su rapidez.

Es importante tener en cuenta que esta operación se basa en la suposición de que la región anatómica visualizada en ambas imágenes es equivalente, y por tanto comparte las mismas dimensiones. Si bien esto no es exactamente cierto, debido a las deformaciones que sufren los órganos entre una y otra adquisición, el error derivado de esta imprecisión es a todos efectos negligible, y será compensado por las posteriores etapas de registro elástico.

#### ■ Paso 4 - Extracción de contornos

Dado que el algoritmo ICP que se utiliza más adelante trabaja sobre superficies, es preciso antes de emplearlo realizar una extracción de los contornos de las estructuras anatómicas a registrar. Ello se consigue mediante una simple operación de barrido en la que se consulta una vecindad de 6 voxels entorno a cada elemento de la estructura, marcándose éste como contorno si cualquiera de ellos no pertenece a la misma (se han realizado pruebas empleando vecindades de 18 y 26 voxels, sin que se aprecien mejoras significativas en los resultados).



**Figura 3.10:** Regiones seleccionadas a partir de las estructuras anatómicas localizadas en la etapa anterior.

■ Paso 5 - Creación de un mapa de distancias

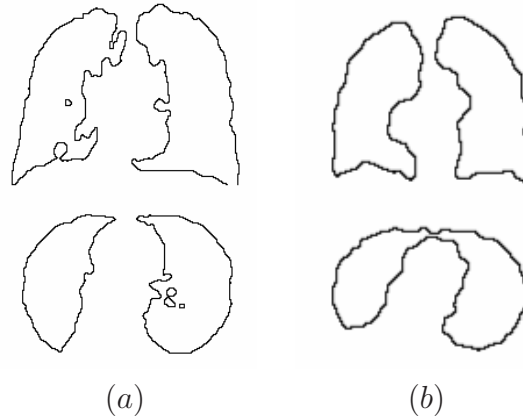
Finalmente, para acelerar la evolución del algoritmo ICP, se calcula un mapa de distancias a las estructuras dentro de la imagen que se tomará como referencia para el registro.

Puesto que el cálculo exacto de la distancia euclídea es relativamente elevado, se emplea una *chamfer distance transform (CDT)* para su aproximación. Se trata de un tipo de algoritmos rápidos desarrollados por Rosenfeld y Pfaltz [43] y posteriormente generalizados por [44] que emplean tan sólo dos barridos del volumen para calcular diversos tipos de transformaciones distancia.

Sea un conjunto  $F$  de los elementos de una imagen binaria tales que  $f(x, y, x) = 1$ , su transformada de distancias  $D$  es una imagen de niveles de gris en la que los valores de cada pixel representan su distancia al elemento más próximo de  $F$ . En general se prefiere la norma euclídea como medida de dicha distancia, ya que ofrece mediciones isotrópicas.

La CDT se basa en la propagación de las distancias locales en una vecindad, generalmente de  $3 \times 3 \times 3$  o  $5 \times 5 \times 5$  voxels para realizar el cálculo de la transformación global, con un error que depende del tamaño de dicha vecindad y de la elección de distancias locales. La distancia desde cualquier punto del volumen al conjunto  $F$  se calcula como la suma de las distancias locales que forman el camino más corto desde  $F$  hasta éste.

Así, para el caso de una propagación por pasos en una vecindad de  $3 \times 3$  voxels, si definimos como  $a$  la distancia entre voxels que difieren en una coordenada,  $b$  entre aquellos que difieren en dos coordenadas y  $c$  entre los que difieren en tres coordenadas, se denomina *chamfer metric*  $(a;b;c)$  a la métrica generada. Por



**Figura 3.11:** Extracción de los contornos de las estructuras a registrar de la imagen CT (a) y PET (b).

tanto, para el caso  $(a;b;c)=(1;1;1)$  se obtiene la distancia *chessboard*, mientras que para  $(a;b;c)=(1;2;3)$  se tiene la distancia *city block*.

Si bien se puede emplear este método con distancias locales no enteras, como es el caso de la aproximación de la distancia Euclídea  $(a;b;c)=(1;1.4;1.7)$ , en la práctica se emplean distancias enteras y se indica  $(a;b;c)=k$  donde  $k$  es el factor de normalización que será preciso aplicar al resultado para obtener los verdaderos valores de distancia. De este modo, se ha empleado la métrica  $(3;4;5)=3$  para estimar las distancias euclídeas a las estructuras contenidas en la imagen de referencia.

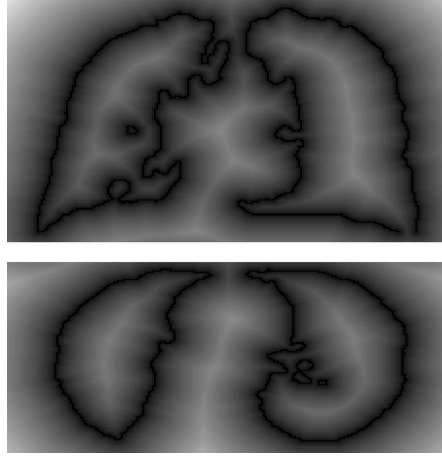
#### ■ Paso 6 - Iterative closest point

El *iterative closest point*, introducido por Besl y McKay [45] es un algoritmo de registro de datos tridimensionales a un modelo que ha sido extensamente empleado en la literatura. Su popularidad se debe, principalmente, al hecho de que siempre converge monótonamente hacia el mínimo local más próximo, con una tasa de convergencia rápida en las primeras iteraciones. Adicionalmente, este algoritmo permite ser empleado sobre múltiples representaciones de los datos tridimensionales, como puntos, segmentos, curvas, mallados triangulares y otras superficies.

Sean  $S$  y  $S'$  dos superficies, existen dos modos de abordar su registro: O bien se determina para cada punto  $P$  de  $S$  su punto correspondiente  $P'$  en la superficie  $S'$  mediante una búsqueda bajo ciertas restricciones en una determinada región de  $S'$ , o bien se emplea una estimación iterativa de dichas correspondencias.

Ésta última es la aproximación empleada por el ICP. El algoritmo itera los siguientes pasos hasta verificar un determinado criterio de parada:

- Sea  $S'$  la superficie que se toma como referencia estática, y  $S$  otra superficie



**Figura 3.12:** Transformación distancia del contorno de referencia.

que se desea registrar con la primera. A cada elemento  $P_i \in S$  le asignamos un elemento correspondiente  $P'_j \in S'$  que minimice  $d^2(P_i - P'_j)$ . Para acelerar este cálculo, se puede emplear una transformación de distancia precalculada del volumen que contiene  $S'$ .

- Se calculan los parámetros de la transformación rígida  $T$  que minimizan la distancia

$$D = \sum_i d^2(T(P_i) - P'_j), \quad (3.4)$$

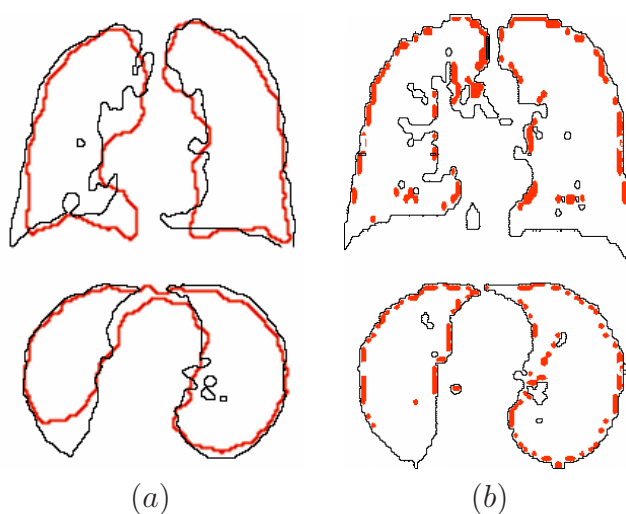
aplicándose dicha transformación a la superficie  $S$ .

Como se ha dicho, este esquema converge monótonamente hacia la posición de registro más próxima.

#### ■ Paso 7 - Proyección no rígida de la superficie

Una vez establecido el registro rígido de las superficies pertenecientes a las distintas estructuras anatómicas localizadas, se aprovecha el propio mapa de distancias calculado anteriormente para proyectar los elementos  $P$  de la superficie a registrar  $S$  sobre sus correspondientes más próximos pertenecientes a la superficie de referencia. De este modo, se obtiene un campo de deformación elástica que mapea ambas estructuras entre si.

Es preciso realizar una cierta regularización de dicho campo, para evitar posibles discontinuidades. Con tal fin, se promedia cada vector con la media de sus vecinos (nótese que hablamos de vecindad en el contexto del volumen de datos, puesto que no se han establecido relaciones de vecindad sobre la propia superficie).



**Figura 3.13:** Registro rígido de los contornos mediante ICP (a) y de proyección elástica de los mismos (b).

- Paso 8 Generalización de la deformación al volumen

Dado que el campo de deformación que se ha obtenido en el paso anterior está restringido a los elementos situados sobre la superficie de los órganos, es necesario propagarlo al conjunto del volumen. Con tal fin, se aplica un proceso de difusión iterativa que fuerza una condición de laplaciano nulo sobre el campo vectorial de deformación.

### Ventajas e inconvenientes del método

El método descrito se ha aplicado con éxito a la inicialización de un registro elástico de imágenes torácicas. Los resultados obtenidos indicaban una clara mejora frente a los mecanismos de inicialización utilizados habitualmente en aplicaciones médicas.

Entre las principales ventajas de este método, se pueden citar:

- Mayor calidad del registro.

Si bien la evaluación de los resultados de un registro elástico es un problema complejo, que se contempla en el capítulo 5, en este caso la mejora con respecto a una inicialización rígida es lo suficientemente significativa como para poder apreciarse mediante una simple inspección visual (sobre todo en las regiones de mayor movilidad, como el diafragma). Las figuras 3.14 a 3.16 muestran sendos cortes axiales y coronales de volúmenes CT y PET, registrados por un lado con una inicialización rígida interactiva y por otro lado con el método aquí propuesto. Se constata una mejora significativa en la correspondencia de las estructuras en este último caso.

- Mejora de la robustez.

Dado que la etapa de inicialización elástica corrige en gran medida las deformaciones más severas entre las imágenes, se reduce significativamente la probabilidad de que el proceso de optimización converja hacia una solución incorrecta. En todas las pruebas realizadas con el método aquí descrito se ha alcanzado una solución de registro aceptable, sin que en ningún caso se apreciara un alineamiento de estructuras no correspondientes.

La robustez del sistema queda pues principalmente determinada por la robustez de la etapa de segmentación. En el capítulo 4 se describen las distintas técnicas empleadas para localizar las distintas estructuras anatómicas, así como para la detección y corrección de los errores que se pudieran cometer en dicha etapa.

- Reducción del tiempo de cómputo.

El hecho de proporcionar una estimación inicial de la transformación elástica permite reducir en gran medida el número de ciclos de optimización necesarios para alcanzar la solución de registro. Dado que cada uno de dichos ciclos implica el cálculo de una medida de similitud basada en la totalidad de la información de intensidades de las imágenes, como es el caso de la NMI, el coste computacional total del proceso disminuye significativamente.

No obstante la considerable mejora que comporta la estrategia de inicialización descrita, es preciso llamar la atención sobre ciertos inconvenientes que esta conlleva, tales como:

- Inexactitud de la proyección elástica.

La proyección de los contornos de las estructuras segmentadas en una y otra modalidades, que representa un método simple y eficiente de estimar la deformación elástica entre estas, resulta excesivamente sensible a las diferencias entre dichas estructuras. Un claro ejemplo de ello son los errores introducidos en el registro de los pulmones, debido a que los bronquios son visibles en el volumen CT y no en el PET.

En cierto modo este error es una consecuencia previsible de la falta de un término que mantenga la cohesión de la estructura durante la deformación.

- Lentitud de la difusión.

A pesar de la rapidez conseguida para la etapa rígida de la inicialización, así como para el cálculo de la deformación elástica entre las distintas superficies, el proceso de generalización de dicha deformación al conjunto del volumen de datos resulta sumamente lento.

Se han ensayado diversas técnicas destinadas a acelerar el proceso, tales como la inicialización de la difusión, el procesado por cortes y la multiresolución, sin que ninguna de ellas haya brindado resultados enteramente satisfactorios.

- Integración en el sistema.

En efecto, la integración del método de inicialización aquí propuesto dentro del marco del sistema de registro descrito en la sección 3.4 resulta ineficiente y poco elegante.

Así, en la inicialización de la transformación basada en FFD's mediante el campo de deformación denso calculado se pierde información y, en consecuencia, parte de la precisión conseguida en el registro de las estructuras segmentadas.

Debido a los inconvenientes arriba expuestos, se ha optado por desarrollar un método alternativo, descrito a continuación, que permita integrar de forma más eficiente la etapa de registro de estructuras anatómicas dentro del sistema de registro basado en intensidades.

Este segundo método trata de compatibilizar los resultados de la inicialización con la estructura del bucle principal de registro. Para ello, proponemos construir una etapa de inicialización basada en el mismo esquema de deformación FFD que emplea el registro, pero sustituyendo la información mutua por un criterio de error cuadrático (RMS).

### 3.5.2. Método 2. Deformación RMS-FFD

Debido a las limitaciones halladas en el método basado en el ICP, se planteó la posibilidad de integrar esta primera etapa de registro, basada en las imágenes segmentadas, dentro de la misma estructura de deformación FFD que en un principio se reservaba para la fase final de registro, basada en intensidades.

Esto implica definir un determinado criterio que permita ajustar la malla de puntos de control que rige la deformación. La elección de dicho criterio es directa en el presente caso, puesto que se trata de registrar imágenes segmentadas la relación entre cuyas distribuciones de intensidades es perfectamente lineal. Por consiguiente, se emplea como medida de similitud la integración sobre todo el volumen del *error cuadrático medio (RMS)* entre las intensidades de los voxels correspondientes.

Así pues, el nuevo método se puede resumir en los siguientes pasos:

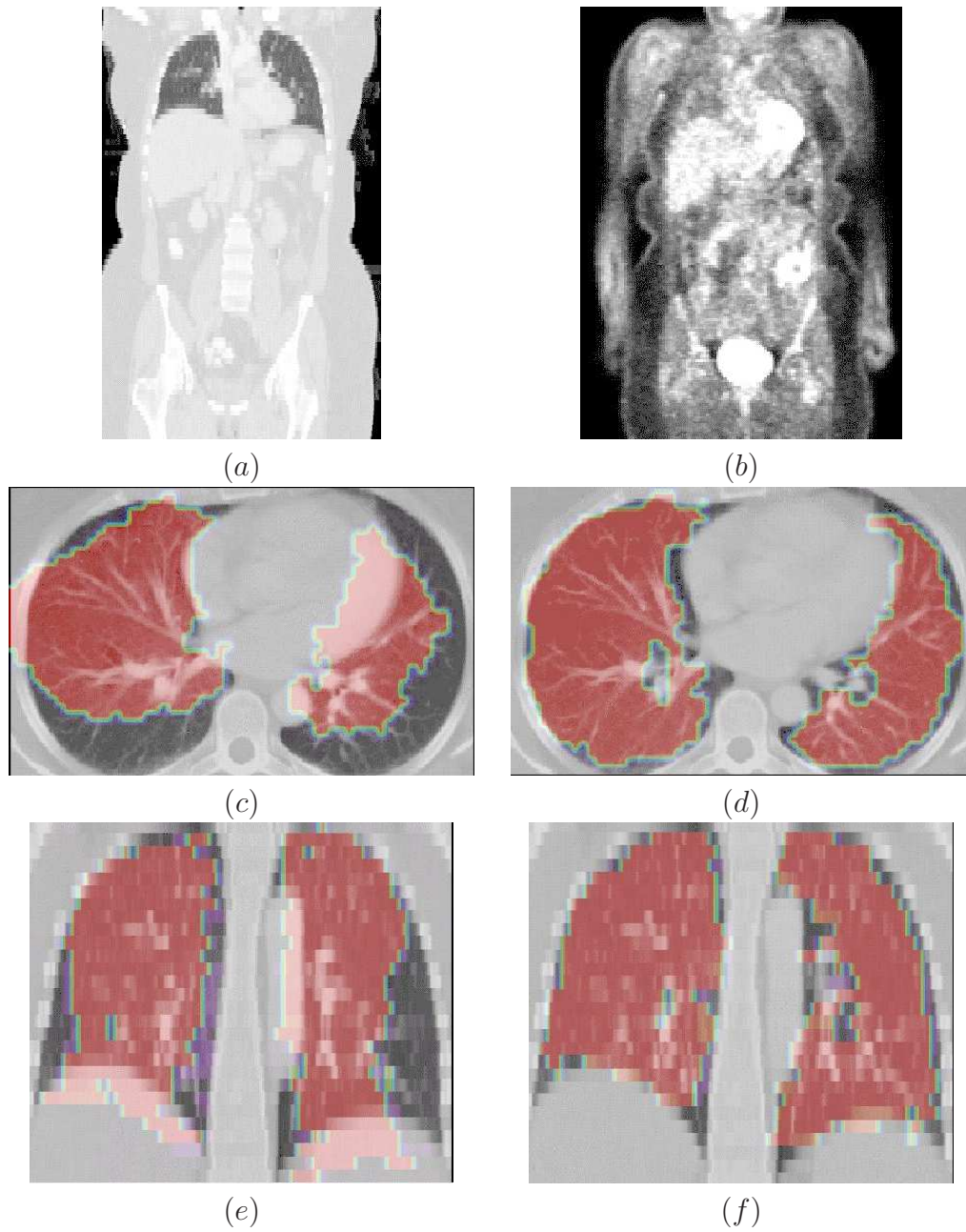
- Paso 1 - Localización de referencias anatómicas

Tal como se ha descrito en la sección 3.5.1, esta etapa permite obtener dos volúmenes de datos binarios, de dimensiones idénticas a las CT y PET originales, donde se representan los distintos órganos localizados como regiones conexas y sin agujeros.

En el capítulo 4, se detallan las distintas técnicas empleadas para localizar y extraer las estructuras anatómicas de interés (piel, esqueleto, pulmones, riñones e hígado).

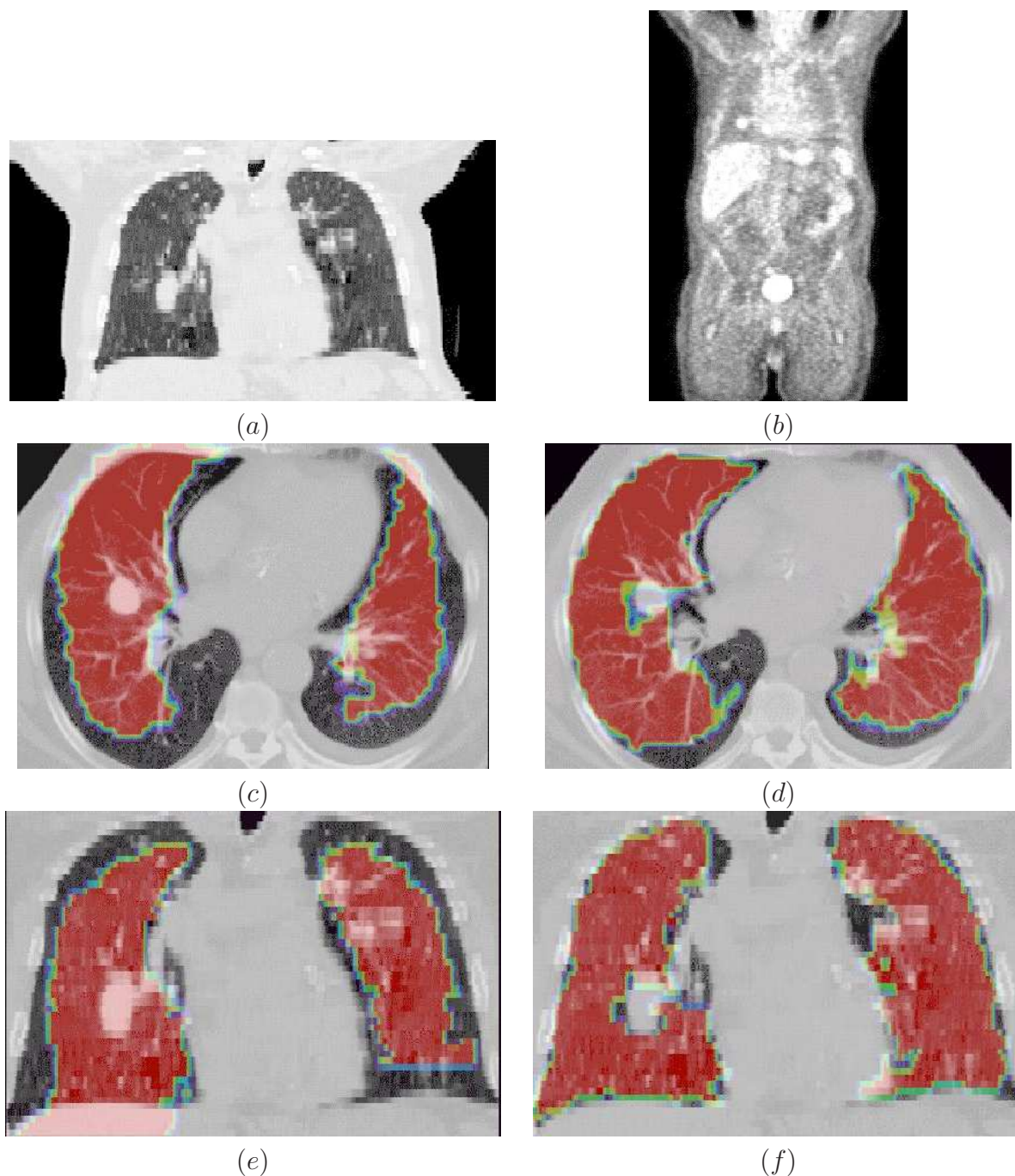
- Paso 2 - Extracción de la región deseada



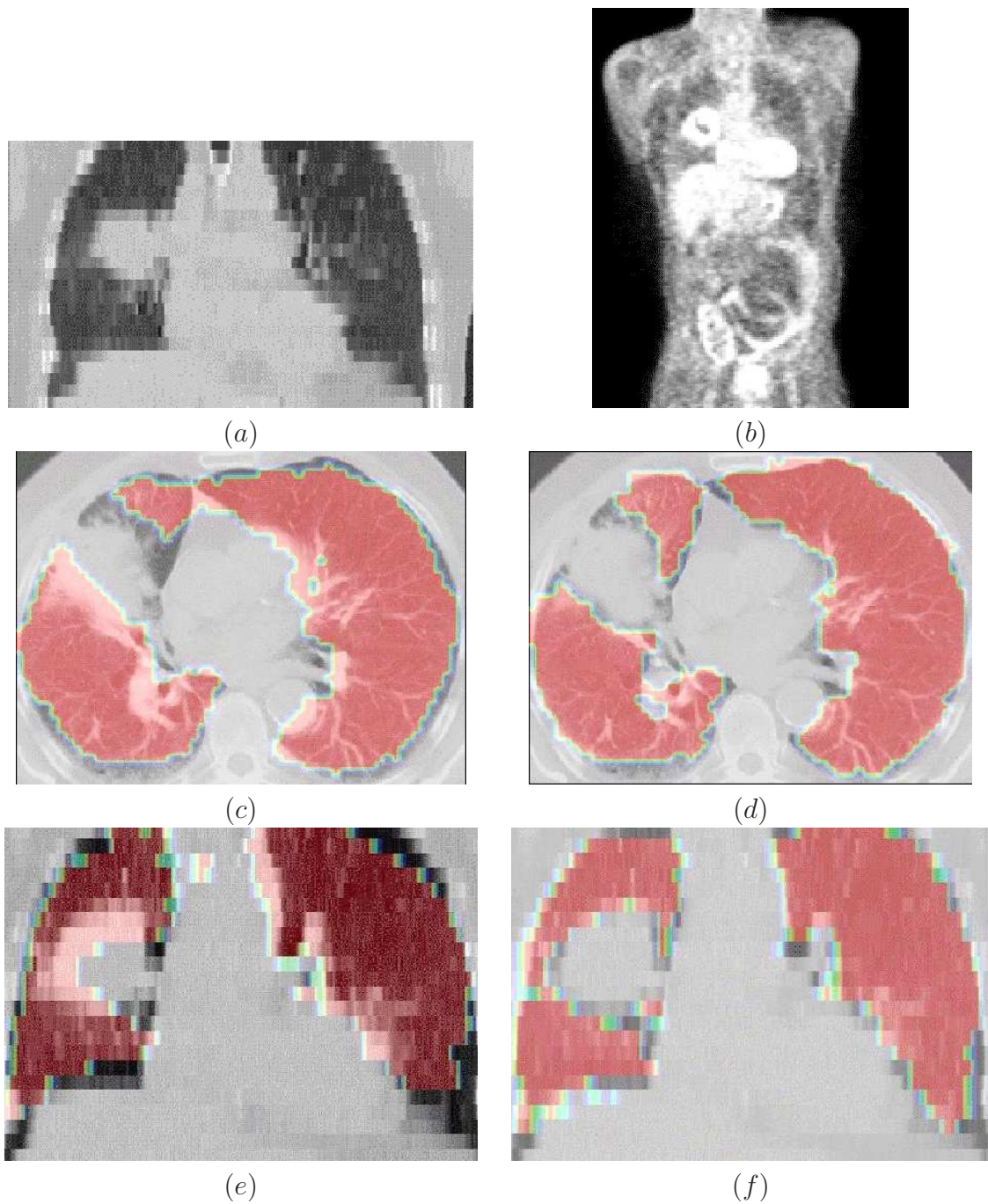


**Figura 3.14:** Resultados del registro empleando la inicialización elástica basada en ICP (Paciente 1): Corte coronal de los volúmenes CT (a) y PET (b) originales. Superposición de los pulmones segmentados a partir de la imagen PET registrada sobre cortes axiales y coroneles de la imagen CT, con inicialización rígida manual (c)(e) y con el método propuesto (d)(f).





**Figura 3.15:** Resultados del registro empleando la inicialización elástica basada en ICP (Paciente 2): Corte coronal de los volúmenes CT (a) y PET (b) originales. Superposición de los pulmones segmentados a partir de la imagen PET registrada sobre cortes axiales y coroneles de la imagen CT, con inicialización rígida manual (c)(e) y con el método propuesto (d)(f). Nótese como la técnica de inicialización elástica consigue un correcto registro del tumor situado en el pulmón derecho.



**Figura 3.16:** Resultados del registro empleando la inicialización elástica basada en ICP (Paciente 3): Corte coronal de los volúmenes CT *(a)* y PET *(b)* originales. Superposición de los pulmones segmentados a partir de la imagen PET registrada sobre cortes axiales y coronales de la imagen CT, con inicialización rígida manual *(c)(e)* y con el método propuesto *(d)(f)*. Nótese como la técnica de inicialización elástica consigue un correcto registro de la masa tumoral necrosada situado en el pulmón derecho.

Al igual que en el método anterior, una vez se tienen las distintas estructuras anatómicas a registrar se restringe el volumen de trabajo a una caja englobante ajustada a las mismas.

Esto permite reducir el coste computacional del resto del proceso, a la vez que se compensan las diferencias de escala entre los volúmenes.

- Paso 3 - Muestreo de los volúmenes a resolución mínima

Se reduce la resolución espacial de los volúmenes, de modo que su número de voxels en cada dimensión coincida con el de resolución menor.

Como se ha descrito para el anterior método, esta medida posibilita que en adelante se trabaje de forma independiente de las dimensiones de los voxels.

- Paso 4 - Optimización del RMS

La optimización de los parámetros de la transformación FFD se consigue mediante la aplicación de un descenso iterativo de gradiente. La obtención de la dirección del gradiente se realiza mediante una estimación local por diferencias finitas.

Se ha incorporado un término de regularización constituido por un conjunto de fuerzas muelle que evitan la intersección de los nodos de la malla, cosa que provocaría alteraciones de la topología del volumen.

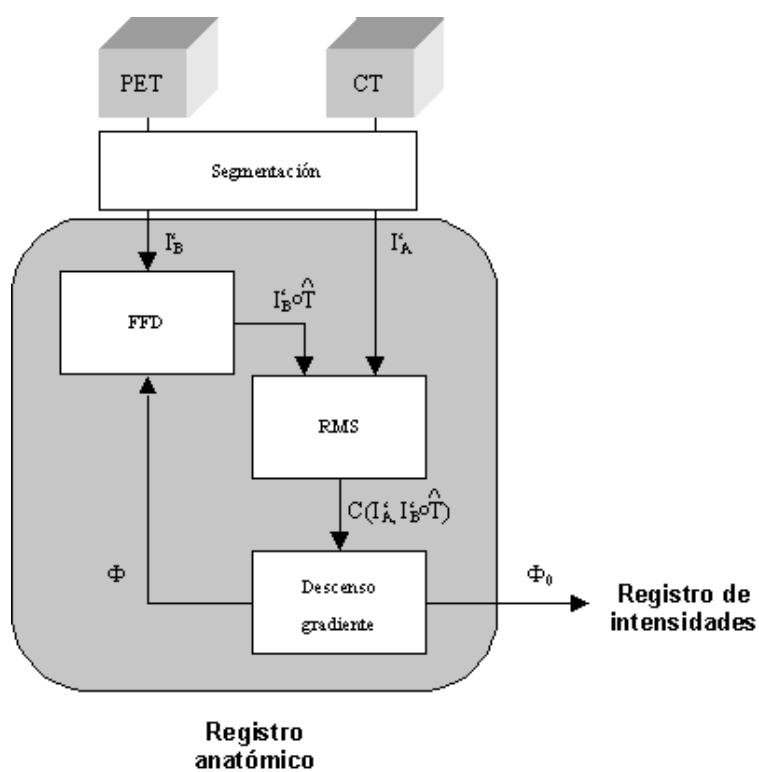
### Ventajas e inconvenientes del método

Es preciso señalar que la utilización de este criterio, que toma en cuenta la totalidad del volumen de datos, incrementa el coste computacional y dificulta la convergencia, problemas que podrían evitarse de emplear un método de registro más simple, como el citado ICP. Sin embargo, la estructura propuesta para esta fase permite su incorporación como una etapa adicional del proceso de multiresolución empleado para el registro basado en intensidades, simplificando enormemente la transferencia de resultados entre estas dos etapas, y en consecuencia eliminando por completo los errores y retardos que esto comportaba en el método anterior. Por otra parte, dada la simplicidad estructural de las imágenes segmentadas, en ningún caso se han experimentado problemas importantes de convergencia, siendo el sistema en su conjunto aproximadamente igual de rápido que el basado en ICP y muy superior en robustez.

### 3.5.3. Resultados

Por bien que el capítulo 5 se consagra a la validación del método de registro propuesto, en este punto se presentan algunos resultados que justifican la elección del esquema de inicialización que se empleará en adelante.

Con tal fin, se han empleado dos medidas distintas de la calidad del registro entre las estructuras anatómicas segmentadas: la fracción de solapamiento entre las mismas



**Figura 3.17:** Esquema del método de inicialización definitivo. Nótese la similitud con el método basado en las intensidades.

y la información mutua normalizada (NMI) entre las imágenes originales afectadas por la deformación calculada.

A continuación se presentan los resultados de dichas medidas para la inicialización mediante ICP y RMS de diversos pacientes, incluyéndose asimismo como referencia el resultado para un registro rígido ordinario constituido por una traslación y un factor de escala.

Se puede apreciar cómo el ICP no aporta demasiado con respecto a una inicialización rígida, mientras que los resultados del RMS, por bien que significativamente más costoso en tiempo de cálculo, justifican sobradamente su elección como método de inicialización.

La figura 3.18 muestra un claro ejemplo de la gran calidad del registro obtenido por éste método.

### 3.6. Conclusión

A lo largo del presente capítulo se ha visto el interés de un método de registro elástico de imágenes PET-CT en aplicaciones de oncología de tórax y abdomen. Se han estudiado las necesidades y condiciones particulares de esta aplicación, y se ha presentado el actual estado del arte.

Teniendo en cuenta todo lo anterior, se ha propuesto una metodología original que permite registrar de forma totalmente automática imágenes tomográficas CT y PET de dichas regiones anatómicas, apta para compensar deformaciones severas tales como las debidas al reposicionamiento del paciente o a su actividad metabólica.

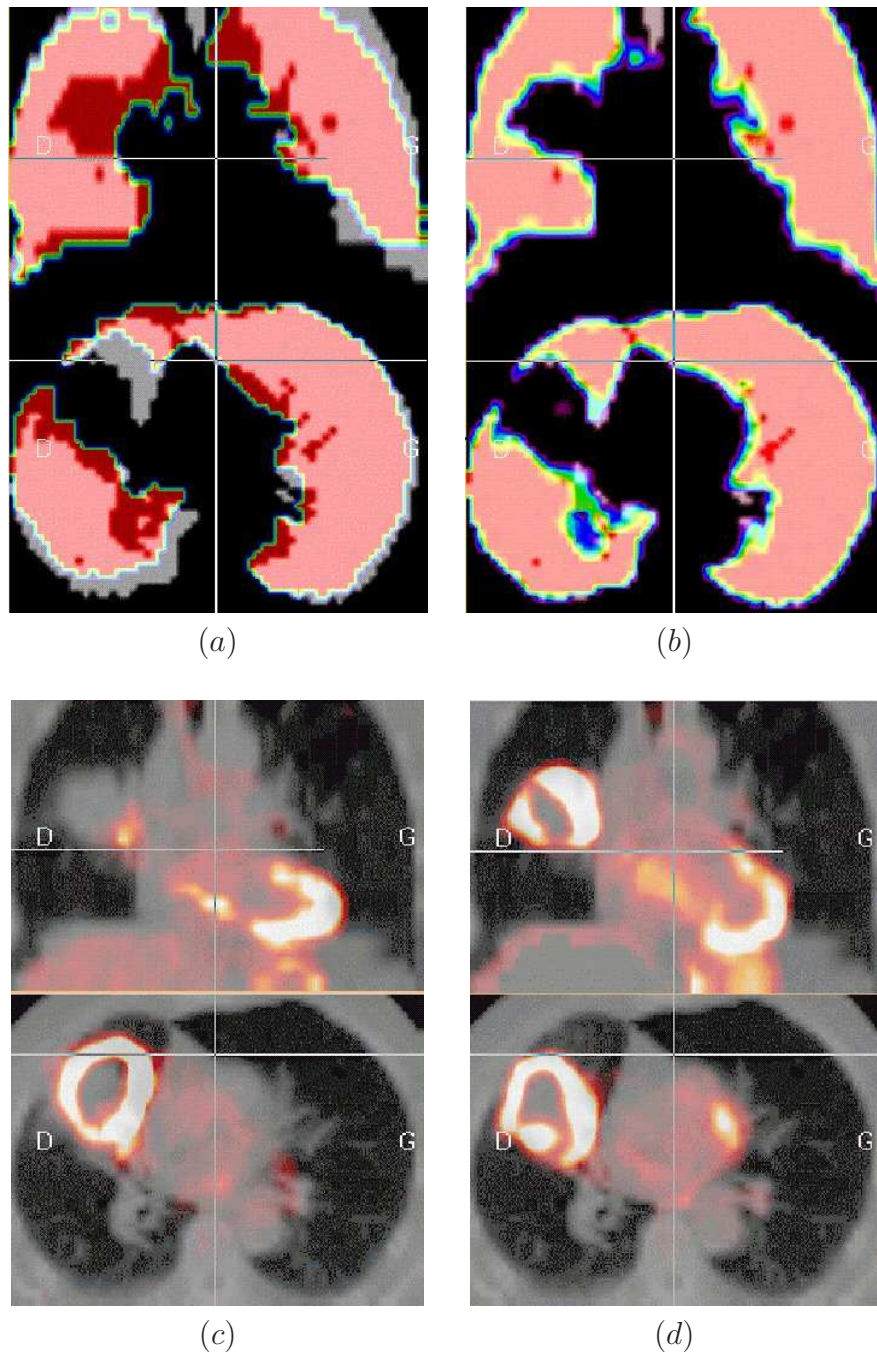
Gracias a una novedosa combinación de registro basado en segmentación y basado en intensidades, se ha conseguido compensar las limitaciones típicas de ambos enfoques, superándose las prestaciones del resto de métodos que han sido empleados en este tipo de aplicación.

Por otra parte, nada en el esquema propuesto impide su uso en otros ámbitos de aplicación, siempre y cuando sea posible identificar y segmentar un conjunto de

**Cuadro 3.1:** Resultados del registro de las estructuras (Solapamiento)

<b>Solapamiento</b>	<b>Rígida</b>	<b>ICP</b>	<b>RMS</b>
Paciente 1	0.68	0.73	0.91
Paciente 2	0.57	0.59	0.85
Paciente 3	0.68	0.71	0.89
Paciente 4	0.78	0.73	0.93
Paciente 5	0.57	0.75	0.89
Paciente 6	0.50	0.54	0.90
Paciente 7	0.68	0.74	0.90
Paciente 8	0.65	0.69	0.89
Paciente 9	0.78	0.80	0.94
Paciente 10	0.65	0.58	0.92





**Figura 3.18:** Resultados del registro empleando la inicialización elástica basada en FFD-RMS: Cortes coronal y axial de los volúmenes segmentados, mostrando la estimación inicial rígida del registro obtenida mediante ICP (a) y mediante el método propuesto (b). Cortes coronal y axial del volumen final, registrado y fusionado, empleando una inicialización rígida (c) y el método propuesto (d).

**Cuadro 3.2:** Resultados del registro de las estructuras (NMI)

NMI	Rígida	ICP	RMS
Paciente 1	0.104	0.106	0.140
Paciente 2	0.147	0.151	0.204
Paciente 3	0.158	0.172	0.258
Paciente 4	0.190	0.176	0.298
Paciente 5	0.248	0.254	0.318
Paciente 6	0.213	0.289	0.340
Paciente 7	0.076	0.081	0.253
Paciente 8	0.310	0.352	0.354
Paciente 9	0.142	0.146	0.162
Paciente 10	0.181	0.179	0.241

estructuras anatómicas suficiente para realizar una inicialización efectiva. Hasta la fecha no tenemos noticia de que se haya empleado una estructura similar en ninguna otra aplicación de registro de imágenes médicas.

En el capítulo siguiente se discute en profundidad el problema de la segmentación de estructuras anatómicas en las zonas torácica y abdominal, proponiéndose un método de clasificación jerárquica capaz de extraer de forma robusta las referencias necesarias para la etapa de inicialización del registro.

## Capítulo 4

# Segmentación de estructuras anatómicas

Este capítulo se centra en la etapa de extracción automática de referencias anatómicas, necesaria para la inicialización y restricción del sistema de registro.

En primer lugar se justificarán las necesidades particulares de nuestra aplicación, que influyen de forma significativa en el tipo de técnicas empleadas en la segmentación. Hecho esto, se describirán extensivamente los distintos algoritmos que componen el método propuesto.

Una primera parte se ocupará de las técnicas basadas en el umbralizado automático de las imágenes y en la morfología matemática, presentándose el esquema de extracción jerárquica de estructuras anatómicas que se ha desarrollado.

En una segunda parte, se describirán las distintas técnicas basadas en modelos paramétricos deformables que se han implementado para ofrecer mayor robustez al sistema, detallándose los resultados obtenidos con las mismas y proponiéndose posibles modificaciones de los mismos para mejorar su rendimiento.

### 4.1. Objetivo de la segmentación

Tal como se justificaba en el capítulo 3, es preciso realizar la segmentación de ciertas estructuras anatómicas para realizar una primera etapa de registro, que posteriormente será refinada mediante una nueva etapa, esta vez basada en las intensidades de las imágenes.

Gracias a esta estructura, no es necesario realizar una segmentación precisa de los datos, dado que los errores que se cometan afectaran a la inicialización, y no necesariamente al resultado final.

Por otra parte, el hecho de emplear los órganos segmentados como base de un registro de las imágenes impone ciertas restricciones a tener en cuenta. Así, únicamente serán de utilidad aquellas estructuras que puedan segmentarse en ambas imágenes, y donde la distinta naturaleza de las modalidades de adquisición empleadas no comporte diferencias excesivas. Esto último se debe al hecho de que el registro de las imágenes



segmentadas trataría de compensar tales diferencias, introduciendo deformaciones erróneas a menudo severas.

Dada la considerable dificultad de segmentar las imágenes CT y, especialmente, las PET, se ha optado por abordar esta tarea mediante un procedimiento jerárquico. De este modo, para la localización de una estructura en una imagen, se empleará información obtenida a partir de otras imágenes previamente segmentadas. Esta información se utilizará en la forma de regiones de interés (ROI) que restrinjan el volumen de búsqueda. Esta idea ya ha sido usada en [46] para la segmentación de imágenes MRI del cerebro.

En el caso que nos ocupa, los órganos que se emplearán para dirigir el registro son los pulmones, riñones e hígado. Se localizarán también durante las primeras etapas del proceso de segmentación la piel y el esqueleto, si bien carecen de interés para el registro en sí mismo, la primera por ofrecer poca información sobre las deformaciones experimentadas por los órganos internos, y el segundo por no poder garantizarse una segmentación equivalente en ambas modalidades.

La figura 4.1 muestra un ejemplo típico de las distintas estructuras segmentadas de un volumen CT, y el orden jerárquico en que se procederá a su segmentación.

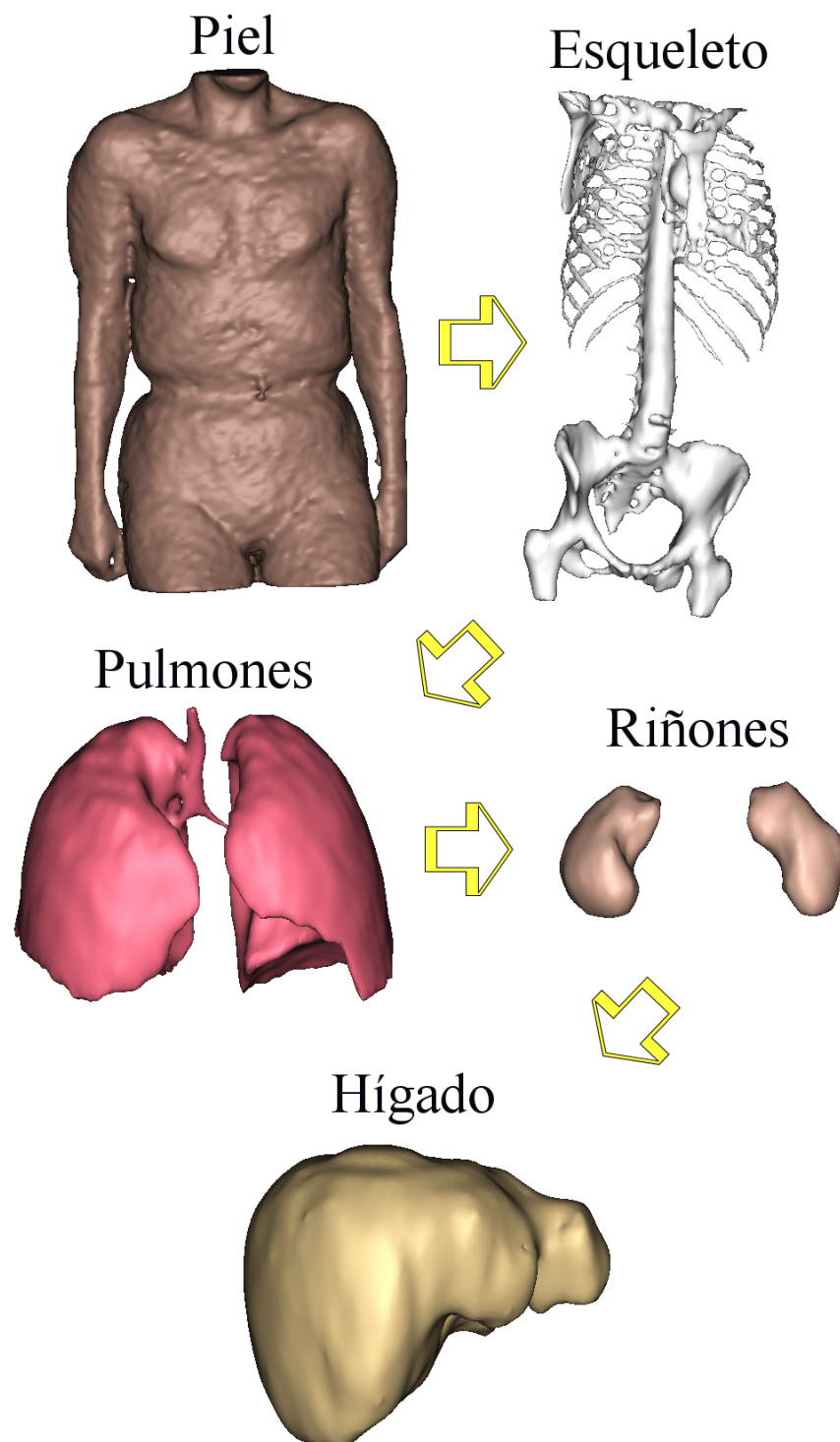
La elección del procedimiento concreto empleado para la segmentación de las anteriores estructuras anatómicas ha sido el resultado de una extensa observación de las técnicas utilizadas por otros grupos, así como a numerosas pruebas con los distintos métodos a nuestro alcance.

Más en concreto, el estudio de las herramientas más comúnmente halladas en la literatura para la segmentación de imágenes médicas reveló una tendencia a emplear técnicas con un fuerte contenido de información a priori, destinada tanto a restringir los resultados como a suplir las posibles lagunas en los datos. Dos técnicas en particular parecían especialmente adecuadas al caso que nos ocupaba: los modelos deformables (o *snakes*) y los conjuntos de niveles (o *level sets*).

Tras algunas pruebas se optó por los primeros, ya que la capacidad de adaptarse a la topología del problema que es característica de los conjuntos de niveles no resultaba realmente necesaria, y no justificaba su mayor coste computacional. Por otra parte, la estructura mallada de los modelos deformables nos permitía introducir fácilmente la información anatómica a priori que deseábamos, necesaria para garantizar los buenos resultados del mismo.

Las primeras pruebas con modelos deformables confirmaron la utilidad de los mismos para nuestra aplicación, siempre y cuando fuera posible inicializarlos en un entorno cercano de la solución final. No obstante, nuestros intercambios con el personal médico especializado nos revelaron que, dada la gran variabilidad de los órganos entre unos pacientes y otros, no es posible realizar dicha inicialización empleando un modelo preestablecido de cada estructura.

Así pues, se decidió recurrir a técnicas de tratamiento de imágenes de bajo nivel para lograr dicha primera aproximación a los órganos de interés, basándose su elección en criterios de velocidad y robustez (y, naturalmente, la necesidad de poder ejecutarse de forma totalmente automática, adaptándose a las condiciones de las imágenes de



**Figura 4.1:** Esquema del orden seguido por el método de segmentación jerárquica de las estructuras anatómicas. Las distintas estructuras mostradas (piel, esqueleto, pulmones, riñones e hígado [41] [42]) corresponden a un renderizado del resultado de la segmentación de un volumen CT.

entrada).

A continuación se presenta un esquema de los distintos pasos seguidos por el proceso de segmentación de los órganos, detallándose en apartados posteriores el funcionamiento de cada una de las etapas.

La extracción de cada estructura se realiza en dos pasos consecutivos:

- En primer lugar se emplean un conjunto de operaciones de umbralizado automático mediante el método de las *k-medias*, acompañadas de ciertas operaciones de morfología matemática para garantizar la topología del resultado.

Tras cada operación, se comprueba que características tales como las dimensiones, volumen y compacidad del objeto obtenido están dentro del rango esperado para la estructura anatómica que deseamos obtener. De no ser así, se usan estos datos para redefinir los parámetros del umbralizado y se repite la operación.

La figura 4.2 ilustra este proceso para la primera etapa de la segmentación jerárquica del volumen CT, que corresponde a la segmentación del cuerpo del paciente. En ella se aprecia como, tras un conjunto de operaciones de preprocesado, se somete al volumen a un umbralizado automático que posteriormente se refina mediante una apertura. A partir de estos resultados se obtiene el fondo de la imagen y, por extensión, el cuerpo del paciente. Una verificación de las dimensiones y posición de la estructura segmentada permite detectar posibles errores y repetir automáticamente el proceso modificando los parámetros pertinentes.

- Hecho esto, se procede a una etapa de refinamiento del objeto segmentado mediante un *modelo deformable 3D*. Esto responde no sólo a la necesidad de introducir restricciones en la segmentación de aquellas estructuras más complejas o menos contrastadas, como es el caso del hígado, sino también para eliminar en la medida de lo posible las diferencias de segmentación debidas a la distinta naturaleza de las imágenes. Un ejemplo muy claro de esto último puede encontrarse en la segmentación de los pulmones, donde los bronquios pueden apreciarse con sumo detalle en la imagen CT mientras que apenas se aprecian en la PET.

En los siguientes apartados se describen con mayor detalle los distintos métodos empleados en la segmentación por umbralizado, así como en la posterior regularización de los resultados mediante un modelo deformable.

## 4.2. Segmentación por umbralizado

Si bien el umbralizado es una operación sumamente sencilla y rápida, a menudo no basta por si sola para conseguir una buena segmentación. Sin embargo, combinando un método de umbralizado automático, como el de las *k-medias*, con operaciones básicas

de morfología matemática se ha conseguido obtener aproximaciones razonables de las estructuras que se desea segmentar.

Dichas aproximaciones han demostrado ser suficientes para el propósito de inicializar el registro, especialmente si previamente se efectúa una regularización mediante un modelo deformable.

Cabe mencionar que, si bien la precisión no es una de las prioridades de esta etapa, si lo es la robustez. Este problema se ha solucionado añadiendo rutinas de verificación de los resultados tras cada paso de la segmentación, que se encargan de ajustar los distintos parámetros del proceso de modo que la estructura obtenida cumpla un conjunto de requisitos en cuanto a sus dimensiones, volumen, etc.

La figura 4.2 muestra el procedimiento que se emplea para segmentar la primera y más simple de las estructuras anatómicas necesarias para la inicialización del registro: el perfil del paciente. Nótese que no se incluye una etapa de regularización del contorno. Esto se debe a que la piel no es una referencia suficientemente fiable para ser utilizada en el registro, por lo que se utiliza únicamente de forma indirecta, para facilitar la segmentación del resto de estructuras (para lo que no se precisa de un contorno regularizado).

A continuación se describen el método de las k-medias y las distintas operaciones morfológicas empleadas en la etapa de segmentación por umbralizado.

### 4.2.1. El método de las k-medias

El método de las k-medias, aparecido en [47], es un caso particular de un método de clasificación automática mucho más general: las nubes dinámicas de Diday [48]

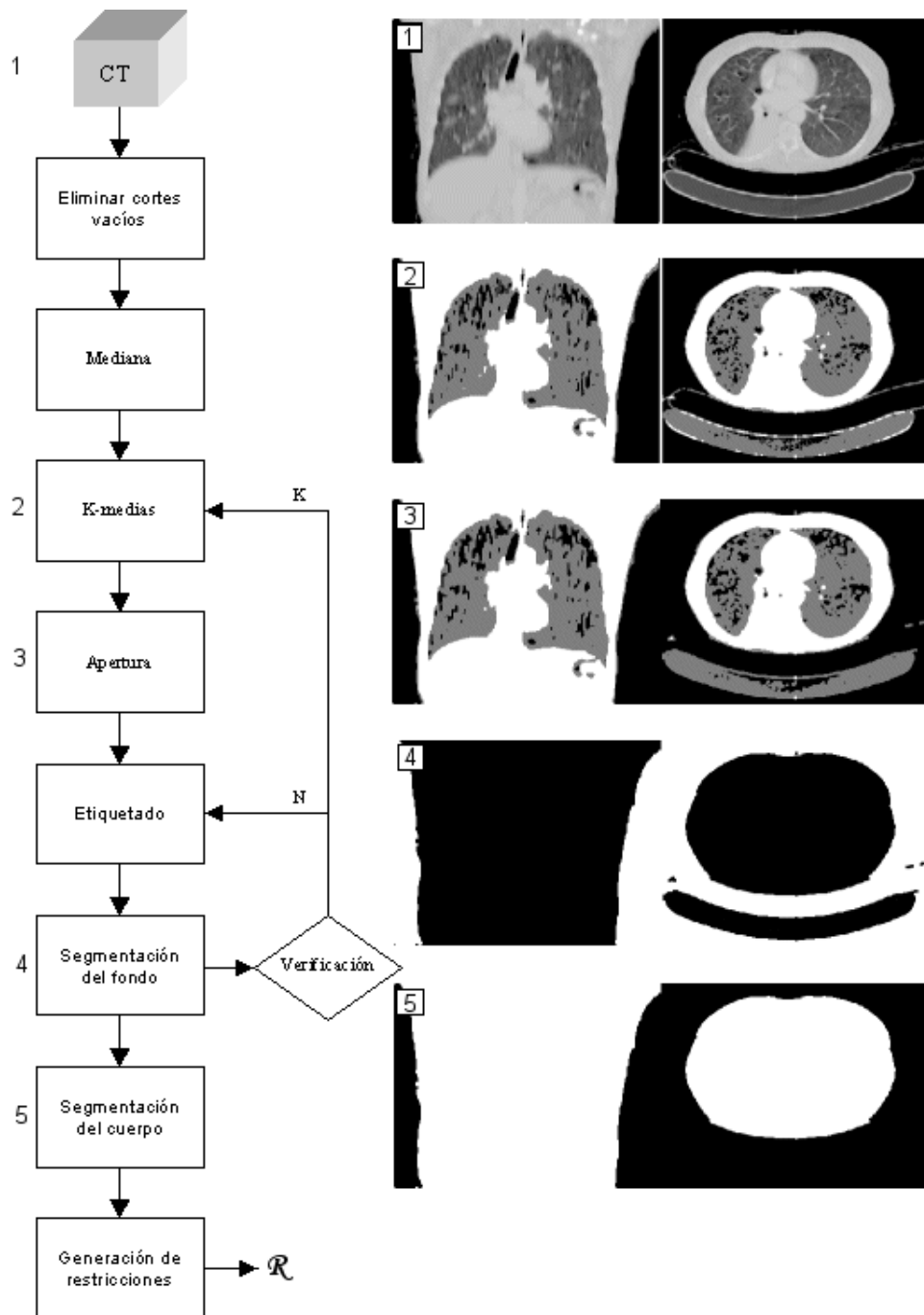
#### Nubes dinámicas

El método de las nubes dinámicas consiste en considerar que las clases formadas por individuos tienen un comportamiento de nubes agrupadas entorno a núcleos. Consideremos la medida  $f(x, w_i)$  de disparidad entre el individuo  $x$  y la clase  $w_i$ ; dicha medida, con valores en  $\mathbb{R}^+$ , se calcula a partir del núcleo de  $w_i$ . Nos proponemos clasificar la imagen de tal modo que la suma global para toda ella de las disparidades entre los individuos y su correspondiente clase sea lo menor posible. La función a minimizar se escribe pues:

$$g_f(\Omega) = \sum_{w_i \in \Omega} \sum_{x_k \in w_i} f(x_k; w_i) \quad (4.1)$$

Para minimizar la función objetivo  $g$ , el algoritmo más frecuentemente utilizado es el siguiente:

1.  $it = 0$
2. seleccionar una clasificación inicial de  $X : w_1^{(it)} .. w_c^{(it)}$



**Figura 4.2:** Ejemplo de una etapa de la segmentación. En este caso, se parte del volumen CT para conseguir delimitar el cuerpo del paciente. Nótese como, de no verificar los resultados ciertas condiciones, se reinicia la segmentación ajustando el número de clases ( $K$ ) o el número de componentes conexos ( $N$ ) a seleccionar.

3.  $it = it + 1$
4. actualizar cada clase según:

$$w_i^{(it)} = \left\{ x_k \in X, f(x_k; w_i^{(it-1)}) \leq f(x_k; w_j^{(it-1)}) \forall j \neq i \right\} \quad (4.2)$$

5. si  $\exists i, w_i^{it} \neq w_i^{it-1}$  entonces volver al paso 3

Así, en una iteración se asigna cada individuo a la clase cuya representación sea la más próxima a las características del mismo. Las clases así redefinidas van cambiando dinámicamente sus características a medida que su forma evoluciona. Tras la clasificación, los elementos  $E$  en la frontera entre las clases de índices  $i$  y  $j$  verifican:

$$f(x, w_i) = f(x, w_j) \quad (4.3)$$

En el método de las nubes dinámicas las características de cada clase dependen del núcleo que describe esa clase;  $f$  es pues la función de disimilitud entre un individuo y el núcleo de una clase. Diday propone diversos ejemplos de medidas  $f$  que garantizan la convergencia del algoritmo de descenso descrito arriba hacia un mínimo local. El ejemplo más simple reduce el núcleo de una clase a un punto de  $E$ , centro de gravedad de la clase.

$$v_i = \frac{1}{\text{card} w_i} \sum_{x_k \in w_i} x_k \quad (4.4)$$

La medida puede construirse mediante una función distancia  $d$  tal que  $f(x, w_i) = d(x, w_i)$ . Esta versión del método de las nubes dinámicas se denomina método de las k-medias. Por extensión, una clase  $w_i$  puede representarse mediante un núcleo constituido por un conjunto de varios puntos  $\{v_i^1..v_i^L\}$ . Una forma de determinarlo y minimizarlo es:

$$\sum_{x_k \in w_i} \inf_{l=1..L} d(x_k, v_i^l). \quad (4.5)$$

Luego, la nueva medida es, por ejemplo:

$$f(x, w_i) = \frac{1}{L} \sum_{l=1..L} c_l d(x, v_i^l), \quad (4.6)$$

con  $c_l$  el coste asociado al  $l$ -ésimo punto del nodo; o también:

$$f(x, w_i) = \frac{1}{L} \sum_{l=1..L} d^2(x, v_i^l). \quad (4.7)$$

La descripción mediante núcleos repartidos entre varios puntos permite adaptarse mejor a clases de formas más complejas. Un núcleo construido entorno al eje principal de inercia  $\Delta i$  de la clase  $w_i$  podría generar la medida:

$$f(x, w_i) = \inf_{x' \in \Delta_i} d(x, x'). \quad (4.8)$$

### Limitaciones

Este método tiene tres particularidades a destacar.

Primeramente, este algoritmo funciona para un número de clases establecido; no resuelve pues la cuestión de cuántas clases hay presentes en la imagen.

En segundo lugar, se trata de un algoritmo de descenso; converge pues hacia un mínimo local. Si se desea evitar la atracción de un mínimo local demasiado alejado de la solución óptima, se debe poner en práctica un estrategia de búsqueda global.

Una primera solución, empírica, consiste en lanzar repetidamente el algoritmo con distintas clasificaciones iniciales. Se conserva pues, bien el resultado que minimiza el valor final de la función objetivo, bien, si la función objetivo no se evalúa, aquel que aparece más repetidamente, evidenciando así la existencia de unas clases estables. Una segunda solución consiste en perturbar regularmente las clases mediante matrices de dispersión.

Finalmente, la medida  $f$  elegida debe garantizar la convergencia del método de optimización.

### Puesta en práctica de las k-medias

Estando los individuos a clasificar descritos por vectores en el espacio de características, generalmente se definen distancias métricas en este último. Seleccionado una métrica para cada una de las clases, la distancia de un individuo  $x_k$  al centro  $v_i$  de la clase  $w_i$  se escribe pues:

$$d(x, w_i) = d_i(x, v_i) \quad (4.9)$$

Se utiliza frecuentemente una de las tres distancias siguientes:

$$d_i = L_1 \quad (4.10)$$

$$d_i = L_2 \quad (4.11)$$

$$d_i^2(x, y) = (x \Leftrightarrow y)^t M_i (x \Leftrightarrow y) \quad (4.12)$$

Con la tercera distancia, la elección de la matriz  $M_i$  condiciona el aspecto morfológico de la clase  $w_i$ , influenciando por ello mismo la forma de las otras clases:

- Si  $M_i = I^p$  (matriz de identidad de dimensión  $p$ ), obtenemos de nuevo la distancia euclídea  $L_2$ .
- Si  $M_i = \sum_i^{-1}$  (matriz de covarianza centrada de un conjunto  $A_i$ ), obtenemos la distancia de Mahalanobis que tiene en cuenta la dispersión, variable en función de la clase, de los individuos entorno de su centro de clase (la separación entre las clases en  $E$  continúan siendo porciones de hiperplanos).

Esta tercera forma es a priori la distancia más representativa de los datos, y por tanto la que da mejores resultados. Desafortunadamente, los cálculos de la matriz de covarianza para cada clase y de la distancia a cada clase para todos los individuos deben repetirse en cada iteración, cosa que comporta tiempos de cálculo largos para una única convergencia.

Como veremos a continuación, utilizamos el algoritmo de las k-medias de forma intensiva, por lo que el coste de la utilización de esta distancia resulta prohibitivo. Así pues, hemos renunciado a esta medida limitándonos a las dos primeras, que consiguen sin embargo resultados excelentes.

### Algoritmo

El algoritmo de descenso empleado en las k-medias es una adaptación del descrito anteriormente:

1.  $it = 0$
2. seleccionar los centros iniciales  $\{v_1^{(it)} .. v_c^{(it)}\}$
3.  $it = it + 1$
4. actualizar cada clase siguiendo

$$w_i^{(it)} = \left\{ x_k \in X, d(x_k, v_i^{(it-1)}) \leq d(x_k, v_j^{(it-1)}) \forall j \neq i \right\} \quad (4.13)$$

5. si una clase está vacía volver al paso 2 y realizar una nueva selección, si no, actualizar los centros siguiendo:

$$v_i^{(it)} = \frac{1}{\text{card} w_i^{(it)}} \sum_{x_k \in w_i^{(it)}} x_k \quad (4.14)$$

6. si  $\exists i, w_i^{(it)} \neq w_i^{(it-1)}$

### Variantes

Existen muchas variantes de este algoritmo. Un ejemplo es una, más lenta, que reestima el centro de una clase cada vez que un nuevo individuo le es asignado o que uno de los individuos pertenecientes a esta clase es asignado a otra.

Una alternativa es el algoritmo ISODATA [49], *Iteratif Self-Organizing Data Analysis Techniques(A)*. Al final de una iteración se procede a un examen de la clasificación obtenida y se modifica si es preciso. Por ejemplo, si una clase contiene demasiado pocos elementos, estos últimos se reasignan a otras clases y esta desaparece; si por el contrario contiene demasiados, será dividida en dos nuevas clases; si dos clases están demasiado próximas, serán fusionadas.



El comportamiento de ISODATA se puede regular mediante siete parámetros que permiten definir reglas de supresión, división y reagrupamiento de clases. Desafortunadamente, la necesidad de tales ajustes hace que su uso repetitivo sea delicado, particularmente si los conjuntos de datos a clasificar son muy heterogéneos.

Para terminar, nótese que cuanto más alejado de el centro de una clase está un individuo mayor es su contribución al cálculo de la función objetivo. O si dicho individuo está aproximadamente a una distancia equivalente del centro de su clase y el de otra clase, su pertenencia en la clasificación es incierta. Con objeto de que no influya en la evolución de las clases, este tipo de individuo puede omitirse en el cálculo del núcleo de las mismas. Esta ultima versión del método de las k-medias está sin embargo sesgada ya que, aunque estos individuos son poco numerosos, su contribución puede eventualmente ser importante debido a su distancia a las clases; olvidarlos modifica mucho el comportamiento de la convergencia y vuelve dudosos sus resultados.

### 4.2.2. Morfología matemática

En este apartado describimos muy brevemente las distintas operaciones morfológicas binarias que se han aplicado durante la segmentación.

Utilizaremos para ello una notación de conjuntos, donde la imagen está constituida por el conjunto de los elementos del fondo y el conjunto de los elementos de los objetos.

Un elemento estructurante es un subconjunto de puntos en  $\mathbb{Z}^2$  (es así para nuestro uso discreto, genéricamente son puntos en el espacio Euclídeo de dimensión  $N$ :  $\mathbb{E}^N$ ), cuya representación en el plano tiene cierta forma y tamaño. El elemento estructurante se concibe como un simple parámetro de forma para los filtros morfológicos. La idea básica es probar la imagen con un elemento estructurante y cuantificar el modo en que está contenido dentro de la imagen.

En una determinada ubicación dentro de la imagen pueden pasar dos cosas: que el elemento estructurante esté contenido o que no lo esté. Marcando las ubicaciones en que está contenido obtenemos información estructural de la imagen. Esta depende de la forma y del tamaño del elemento estructurante. La característica de estar contenido depende de la relación de subconjunto. En adelante, llamaremos  $A$  a la imagen de entrada y  $B$  al elemento estructurante.

Las operaciones morfológicas básicas que se han empleado son:

- Dilatación.

La dilatación es una transformación morfológica que combina dos conjuntos usando la suma vectorial de elementos de un conjunto.

$$A \oplus B = \{a + b : a \in A, b \in B\} \quad (4.15)$$

Es decir, el conjunto de todos los posibles vectores suma de pares de elementos, uno perteneciente a  $A$  y el otro a  $B$ .

- Erosión.

Es la transformación que combina dos conjuntos usando la resta vectorial de elementos de conjuntos. Si  $A$  y  $B$  son conjuntos en el espacio euclídeo de dimensión  $N$ , entonces la erosión de  $A$  por  $B$  es

$$A \ominus B = \{x \in \mathbb{E}^N / x + b \in A, \forall b \in B\} \quad (4.16)$$

Es decir, son los puntos  $x$  para los cuales la traslación de  $B$  por  $x$  está contenida en  $A$ .

- Apertura.

Es el resultado de aplicar consecutivamente una erosión y una dilatación sobre la misma imagen.

$$A \circ B = A_B = (A \ominus B) \oplus B. \quad (4.17)$$

- Cierre.

Es el resultado de aplicar consecutivamente una dilatación y una erosión sobre la misma imagen.

$$A \bullet B = A^B = (A \oplus B) \ominus B. \quad (4.18)$$

Otra operación de morfología matemática, esta vez no binaria, que se ha empleado en esta aplicación, es el filtro de mediana. Se ha utilizado un filtrado reiterado de mediana como etapa de preprocesado, para mejorar el contraste y reducir el ruido de las imágenes antes de realizar el umbralizado.

Sea  $u : \mathbb{Z}^2 \rightarrow \mathbb{R}$  una imagen, y sea un elemento estructurante definido como un conjunto  $B \subset \mathbb{Z}^2$ . Se asocia una medida de Lebesgue a  $\mathbb{Z}^2$  que escribimos como  $|\cdot|$ . Si se define, para  $x \in \mathbb{R}^2$  y  $\lambda \in \mathbb{R}$  los conjuntos

$$I_\lambda(x) = \{y \in B, u(x + y) < \lambda\} \quad (4.19)$$

$$S_\lambda(x) = \{y \in B, u(x + y) > \lambda\} \quad (4.20)$$

entonces, la mediana de  $u$  con respecto a  $B$  es la función:

$$\text{med}_B u : x \mapsto \left\{ \lambda, |I_\lambda(x)| \leq \frac{|B|}{2} \text{ y } |S_\lambda(x)| \geq \frac{|B|}{2} \right\} \quad (4.21)$$

### 4.3. Modelos deformables

Los modelos deformables [50] son curvas o superficies definidas en el dominio de una imagen que pueden moverse bajo la influencia de fuerzas internas provenientes del propio modelo y de fuerzas externas calculadas a partir de los datos de la imagen.

Las fuerzas internas y externas se definen de forma que el modelo se adaptara a la frontera de un objeto o a otras características deseadas de la imagen. Los modelos deformables son ampliamente utilizados en muchas aplicaciones, incluyendo detección de contornos, modelización de formas, segmentación y seguimiento.

Hoy en día existen en la literatura dos tipos generales de modelos deformables: modelos deformables paramétricos y modelos deformables geométricos. En este capítulo nos centraremos en los modelos deformables paramétricos, que generan curvas o superficies parametrizadas en el dominio de la imagen y los hacen evolucionar hacia las características deseadas, en general contornos.

Típicamente, los modelos son atraídos hacia los contornos por fuerzas potenciales, definidas como el gradiente negativo de una función de potencial. Dichas fuerzas potenciales, junto a ciertas fuerzas adicionales, como las fuerzas de presión, comprenden las fuerzas externas. Hay asimismo fuerzas internas diseñadas para mantener el modelo unido (fuerzas de elasticidad) y evitar que se curve demasiado (fuerzas de torsión).

Los modelos deformables paramétricos presentan dos dificultades principales:

En primer lugar, el modelo inicial debe, generalmente, estar próximo al verdadero contorno o es probable que converja hacia un resultado erróneo. Se han propuesto muchos métodos para superar este problema, incluyendo métodos multiresolución, fuerzas de presión y potenciales de distancia. La idea básica es incrementar el rango de captura de los campos de fuerzas externas y guiar el modelo hacia el contorno deseado.

El segundo problema es que los modelos deformables tienen dificultades para evolucionar hacia el interior de concavidades del contorno. No existe una solución satisfactoria a esto, aunque se han propuesto fuerzas de presión, puntos de control, adaptabilidad del dominio, atracciones direccionales y el uso de campos solenoidales.

La mayor parte de los métodos propuestos para solucionar estos problemas, sin embargo, resuelven solo un problema al tiempo que crean nuevas dificultades. Por ejemplo, los métodos multiresolución se centran en mejorar el rango de captura, pero especificar cómo el modelo deformable debe moverse a través de las distintas resoluciones resulta problemático. Otro ejemplo es el de las fuerzas de presión, que pueden forzar un modelo hacia el interior de las concavidades del contorno, pero no pueden ser demasiado fuertes o el modelo rebasará los bordes débiles. Las fuerzas de presión deben asimismo inicializarse para actuar hacia el exterior o hacia el interior, condición que requiere una cuidadosa inicialización.

Los campos de fuerzas generados mediante *Gradient Vector Flow* proporcionan una buena solución a ambos problemas, pero a cambio de un elevado costo computacional, a menudo mayor que el de la propia evolución del modelo.

### 4.3.1. Modelos deformables paramétricos 2D: Contornos activos

Un contorno activo (o *Snake*) es una curva parametrizada  $X$  de la cual se da su posición en el espacio en función de su abscisa curvilínea normalizada por su longitud:

$$X(s) = [x(s), y(s)]^t \quad \text{donde} \quad s \in [0, 1]. \quad (4.22)$$

A este contorno se asocia una energía compuesta por un término de *energía interna* relacionado con el comportamiento físico del contorno, y un término de *energía externa* que evalúa la fidelidad a los datos que se desean modelar.

$$E(X) = E_{int} + E_{ext}(X) \quad (4.23)$$

#### Término de energía interna

En general la energía interna del modelo se expresa de la forma siguiente:

$$E_{int} = \int_0^1 \left( \alpha(s) \left| \frac{dX(s)}{ds} \right|^2 + \beta(s) \left| \frac{d^2 X(s)}{ds^2} \right|^2 \right) ds, \quad (4.24)$$

donde  $\alpha$  y  $\beta$  son respectivamente el coeficiente de elasticidad y de rigidez del modelo. Estos coeficientes permiten ajustar el comportamiento del modelo, de forma que un valor relativamente elevado de  $\alpha$  tiende a acortar el snake y eliminar bucles, mientras que un valor elevado de  $\beta$  penaliza las zonas de gran curvatura, alisando el contorno.

#### Término de energía externa

La energía externa representa la información proveniente de la imagen sobre la que se está trabajando, por lo que toma valores mínimos entorno a los elementos de interés, como pueden ser los contornos.

Dada una imagen de nivel de gris  $I(x, y)$  (que consideraremos continua en función de las coordenadas  $(x, y)$ ), las expresiones más comunes para que el término de energía externa haga converger el modelo hacia los contornos más pronunciados serían:

$$E_{ext}^1(x, y) = -|\nabla I(x, y)|^2 \quad (4.25)$$

$$E_{ext}^2(x, y) = -|\nabla (G_\sigma(x, y) * I(x, y))|^2 \quad (4.26)$$

donde  $G_\sigma(x, y)$  es una función gaussiana bidimensional de desviación típica  $\sigma$ .

Es evidente que cuanto mayor sea  $\sigma$ , más borrosas e imprecisas serán las fronteras que se obtengan. Sin embargo dicha imprecisión es a menudo necesaria para que el efecto de atracción hacia los contornos sea apreciable a cierta distancia de los mismos, ensanchándose de este modo el rango de captura del modelo.

## Resolución

El problema consiste pues en hallar la curva que minimiza el funcional de energía:

$$E(X) = \int_0^1 \left( \alpha(s) \left| \frac{dX(s)}{ds} \right|^2 + \beta(s) \left| \frac{d^2X(s)}{ds^2} \right|^2 \right) + E_{ext}(X(s)) ds \quad (4.27)$$

Esto equivale a resolver la ecuación de Euler-Lagrange [51]:

$$\alpha(s) \frac{d^2X(s)}{ds^2} - \beta(s) \frac{d^4X(s)}{ds^4} - \nabla E_{ext}(X(s)) = 0 \quad (4.28)$$

Otra posible interpretación de 4.28 sería una ecuación de equilibrio de fuerzas

$$F_{int} + F_{ext} = 0 \quad (4.29)$$

en la que  $F_{int} = \alpha(s)X''(s) - \beta(s)X''''(s)$  y  $F_{ext} = -\nabla E_{ext}$ . La fuerza interna  $F_{int}$  penaliza el estiramiento y la torsión, mientras la fuerza potencial externa  $F_{ext}$  atrae el modelo hacia el contorno de la imagen deseado.

Para hallar una solución de 4.28, se vuelve dinámico el modelo, incorporando el tiempo  $t$  como variable de la función  $X(s, t)$ . Así, la derivada parcial de  $X$  respecto al tiempo se iguala al primer término de la ecuación como sigue

$$\frac{dX(s, t)}{dt} = \alpha(s) \frac{d^2X(s, t)}{ds^2} - \beta(s) \frac{d^4X(s, t)}{ds^4} - \nabla E_{ext}(X(s)) \quad (4.30)$$

A medida que la solución  $X(s, t)$  se estabiliza, su derivada respecto al tiempo desaparece, con lo que se llega a la deseada solución de 4.28. Esto puede interpretarse también como un descenso de gradiente destinado a resolver 4.27.

La solución de 4.30 puede obtenerse discretizando la ecuación y resolviendo iterativamente el sistema obtenido, como se muestra a continuación.

## Implementación numérica

Si bien la resolución de la ecuación diferencial que rige el comportamiento del modelo deformable puede realizarse por otras vías (elementos finitos, Fourier), aquí nos centraremos en el método de las diferencias finitas en el tiempo, por su mayor simplicidad y rapidez.

Así, se parte de un modelo del contorno caracterizado por un vector  $X$  de  $n$  muestras, con  $n$  constante. Dicho vector contiene en el caso más simple las coordenadas de las muestras, si bien podría incluir asimismo la discretización de otras magnitudes físicas del modelo, como masa, elasticidad, rigidez, etc., que en este caso consideraremos constantes. Entonces, para un paso de muestreo espacial  $\kappa$ , se escribe  $X = (x_i)$ , con  $x_i = (x(i\kappa), y(i\kappa))$ .

Supondremos ahora que la energía total del modelo es la suma de las energías asociadas a cada una de las muestras del mismo, y que éstas se pueden descomponer en



**Figura 4.3:** Pulmón segmentado mediante un snake 2D a partir de un corte axial del volumen CT. Se aprecia como el modelo ha localizado correctamente los contornos del pulmón, sin verse afectado por la presencia de los bronquios. Nótese asimismo como la energía interna del modelo ha limitado la curvatura del mismo, por ejemplo en la zona próxima al esternón.

un término de energía externa (asociada a la imagen) y otro de energía interna (debida a las fuerzas de tensión y torsión a las que está sometido ese punto del contorno),

$$E_{total} = \sum_{i=1}^n E_{int}(i) + E_{ext}(i). \quad (4.31)$$

La energía externa puede obtenerse a partir de una imagen de energías precalculada. Nótese que, dado que la posición de cada elemento del contorno no está discretizada y puede tomar cualquier valor, su correspondiente energía debe obtenerse por interpolación de los valores para las posiciones discretas más próximas. Esta interpolación es sumamente importante, pues permitirá que el contorno converja suavemente hacia la solución en vez de oscilar entre dos valores discretos de energía.

Obtenemos asimismo una expresión sencilla de la energía interna del modelo a partir de los valores de sus muestras,

$$E_{int}(i) = \alpha \frac{|x_i - x_{i-1}|^2}{2\kappa^2} + \beta \frac{|x_{i-1} - 2x_i + x_{i+1}|^2}{2\kappa^4}. \quad (4.32)$$

En consecuencia, la solución de la ecuación de Euler-Lagrange 4.28 equivale a resolver la ecuación matricial

$$AX + F_{ext}(X) = 0, \quad (4.33)$$

donde  $A$  es la matriz pentadiagonal  $(n \times n)$ :

$$\begin{vmatrix}
2\alpha + 6\beta & -\alpha - 4\beta & \beta & 0 & 0 & \dots \\
-\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \beta & 0 & \\
\beta & -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \beta & \\
0 & \beta & -\alpha - 4\beta & 2\alpha + 6\beta & -\alpha - 4\beta & \\
0 & 0 & \beta & -\alpha - 4\beta & 2\alpha + 6\beta & \\
\vdots & & & & & \ddots
\end{vmatrix} \quad (4.34)$$

Si a la anterior discretización espacial añadimos una discretización de la evolución del modelo en intervalos temporales  $\tau$ , podemos aproximar la derivada temporal en 4.30 por una diferencia entre intervalos sucesivos, tal que

$$\frac{X^t - X^{t-1}}{\tau} + AX^t = F_{ext}(X^{t-1}), \quad (4.35)$$

por lo que la evolución del contorno deformable queda regida, a partir de su estado inicial  $X^0$ , por la ecuación:

$$X^t = (Id + \tau A)^{-1} (\tau F_{ext}(X^{t-1}) + X^{t-1}). \quad (4.36)$$

La ventaja de esta formulación es que la inversión de la matriz  $(Id + \tau A)$  sólo debe realizarse una vez, dado que no incluye ningún término dinámico. Como contrapartida, la elección del paso temporal  $\tau$  exige un compromiso que garantice la convergencia del modelo sin ralentizar excesivamente su evolución.

### 4.3.2. Difusión del vector gradiente

En este apartado presentamos un tipo de campos de fuerzas externas para modelos deformables que resuelve el problema del rango de captura a la vez que garantiza una buena convergencia hacia las concavidades del contorno.

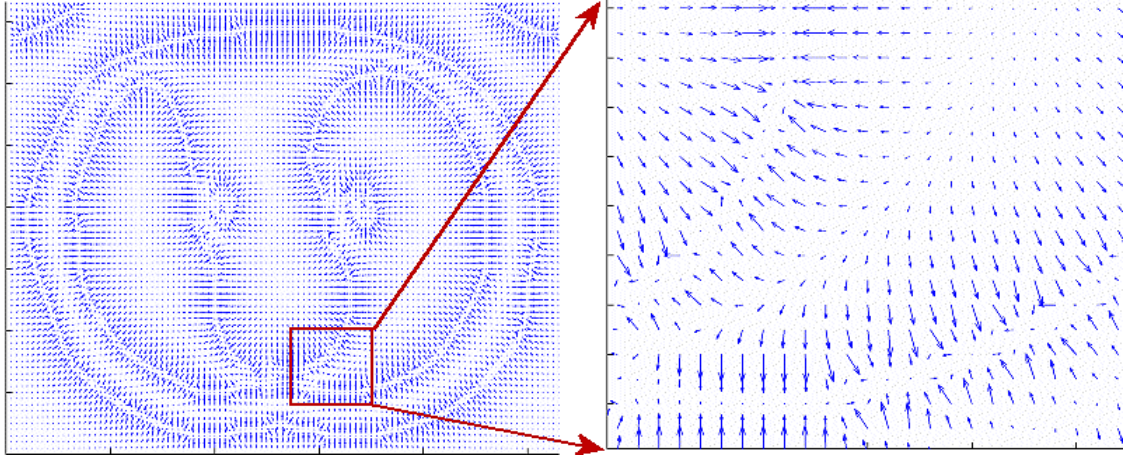
Estos campos, llamados de difusión del vector gradiente, o *gradient vector flow* (GVF), son campos densos de vectores derivados de las imágenes mediante la solución de una ecuación de difusión vectorial que distribuye los vectores de gradiente de un mapa de contornos binario o de niveles de gris calculado a partir de la imagen.

La figura 4.4 muestra un ejemplo de campo GVF bidimensional, calculado a partir de los contornos de un corte axial CT del tórax. La ampliación permite apreciar como los vectores convergen, de forma regular, hacia los contornos de la imagen, incluso en el caso de cavidades estrechas.

El GVF se introdujo por primera vez en [52] y se propuso una generalización del mismo en [53]. Aquí presentamos el GVF en su forma generalizada.

Los modelos deformables que utilizan GVF se distinguen de la mayor parte de las anteriores formulaciones de modelos deformables en que las fuerzas externas no pueden escribirse como el gradiente negativo de una función de potencial. Debido a esto, no pueden formularse utilizando la estructura estándar de minimización de energía; en lugar de ello se especifica directamente como una ecuación de dinámica de fuerzas.





**Figura 4.4:** Campo GVF calculado a partir de un corte axial del volumen CT. En la imagen ampliada se aprecia como el campo de fuerzas converge hacia los contornos de los pulmones.

Se define el campo GVF  $v(x)$  como la solución de equilibrio de la siguiente ecuación de difusión vectorial:

$$u_t = g(|\nabla f|) \nabla^2 u - h(|\nabla f|)(u - \nabla f) \quad (4.37)$$

$$u(x, 0) = \nabla f(x) \quad (4.38)$$

En la ecuación 4.37 el primer término de la derecha se denomina *término de suavizado*, ya que este término por si solo produce un campo vectorial que varía uniformemente. El segundo término es el *término de datos* ya que hace que el campo vectorial  $u$  se aproxime al  $\nabla f$  calculado a partir de los datos.

Las funciones de ponderación  $g(\cdot)$  y  $h(\cdot)$  se aplican al término de suavizado y de datos respectivamente. Dado que dichas funciones de ponderación dependen del gradiente del mapa de contornos que es variable en el espacio, generalmente los pesos en si mismos varían en el espacio. Dado que queremos que el campo vectorial  $u$  varíe lentamente (suavemente) en aquellas posiciones alejadas del contorno, pero que se ajuste a  $\nabla f$  cerca del mismo,  $g(\cdot)$  y  $h(\cdot)$  deben ser funciones monótonas no-crecientes y no-decrecientes de  $|\nabla f|$ , respectivamente.

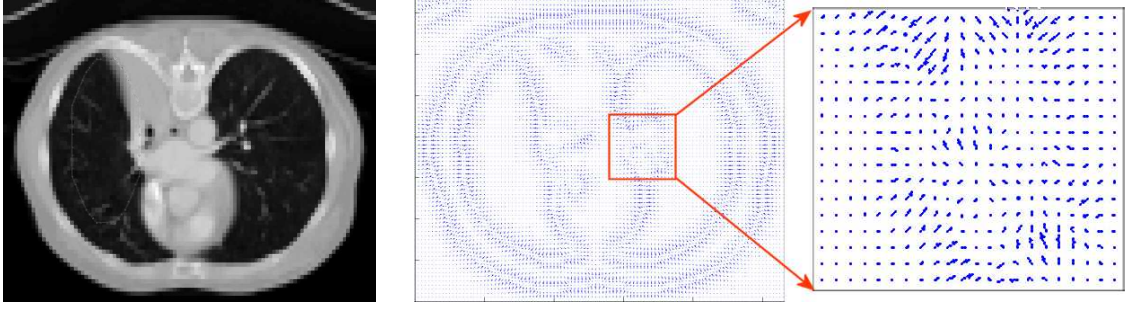
Los propios autores del GVF proponen las siguientes funciones de peso [53]:

$$g(|\nabla f|) = \mu \quad (4.39)$$

$$h(|\nabla f|) = |\nabla f|^2 \quad (4.40)$$

Dado que  $g(\cdot)$  es aquí constante, el suavizado se produce en toda la imagen; sin embargo  $h(\cdot)$  crece cerca de los bordes fuertes, y debería dominar en el contorno. Así pues, el GVF calculado con estas funciones de ponderación consigue una buena localización del contorno. El efecto del suavizado se aprecia, sin embargo, cuando hay





**Figura 4.5:** Ejemplo de error del GVF. En la ampliación se aprecia como un término de regularización excesivo ha provocado el promediado de dos contornos cercanos y opuestos.

dos bordes próximos, como cuando hay una indentación larga y estrecha en la frontera. En este caso el GVF tiende a suavizar entre los bordes opuestos, perdiéndose las fuerzas necesarias para dirigir el modelo deformable en esa zona. La figura 4.5 ilustra este efecto.

Para afrontar este problema, se proponen funciones de ponderación en las que  $g(\cdot)$  disminuye a medida que  $h(\cdot)$  crece. Así en la proximidad de gradientes fuertes el suavizado será pequeño, y el campo vectorial resultante será prácticamente igual al gradiente de la imagen de contornos.

Hay muchas maneras de especificar dicho par de funciones de ponderación. En [53] se utilizan las siguientes:

$$g(|\nabla f|) = e^{-\left(\frac{|\nabla f|}{K}\right)^2} \quad (4.41)$$

$$h(|\nabla f|) = 1 - g(|\nabla f|) \quad (4.42)$$

El GVF calculado utilizando estas funciones de ponderación se ajustará a la imagen de contornos en los bordes fuertes, pero variará suavemente lejos de los bordes. La especificación de  $K$  determina hasta cierto punto el compromiso entre suavidad del campo y conformidad al gradiente.

### Implementación numérica

La ecuación de difusión vectorial 4.37, que especifica el GVF dadas unas funciones de ponderación, puede implementarse mediante un esquema de diferencias finitas. Para el caso de las funciones de ponderación 4.39 y 4.40 esto consiste en resolver la ecuación:

$$u_t(x, y, t) = \mu \nabla^2 u(x, y, t) - [u(x, y, t) - f_x(x, y)] \cdot [f_x(x, y)^2 + f_y(x, y)^2] \quad (4.43)$$

$$v_t(x, y, t) = \mu \nabla^2 v(x, y, t) - [v(x, y, t) - f_y(x, y)] \cdot [f_x(x, y)^2 + f_y(x, y)^2] \quad (4.44)$$

La solución en estado estacionario de estas ecuaciones lineales parabólicas es la solución de las ecuaciones de Euler 4.37. Nótese que las ecuaciones son desacopladas, y por tanto pueden resolverse separadamente como ecuaciones diferenciales escalares con derivadas parciales en  $u$  y  $v$ .

Las ecuaciones 4.43 y 4.44 se conocen como las *ecuaciones de difusión generalizadas* y se pueden encontrar en campos tan dispares como la conducción de calor y la fluídica.

Para mayor comodidad podemos reescribirlas como sigue:

$$u_t(x, y, t) = \mu \nabla^2 u(x, y, t) - b(x, y)u(x, y, t) + c^1(x, y) \quad (4.45)$$

$$v_t(x, y, t) = \mu \nabla^2 v(x, y, t) - b(x, y)v(x, y, t) + c^2(x, y) \quad (4.46)$$

donde

$$b(x, y) = f_x(x, y)^2 + f_y(x, y)^2 \quad (4.47)$$

$$c^1(x, y) = b(x, y)f_x(x, y) \quad (4.48)$$

$$c^2(x, y) = b(x, y)f_y(x, y) \quad (4.49)$$

Cualquier operador digital de gradiente puede utilizarse para calcular  $f_x$  y  $f_y$ . Los coeficientes  $b(x, y)$ ,  $c^1(x, y)$  y  $c^2(x, y)$  pueden entonces calcularse y permanecen fijos durante el proceso iterativo.

Para preparar la solución iterativa, sean los índices  $i$ ,  $j$  y  $n$  correspondientes a  $x$ ,  $y$  y  $t$  respectivamente, y sean  $\Delta x$  y  $\Delta y$  las distancias entre píxeles y  $\Delta t$  el paso temporal entre iteraciones. En este caso se pueden aproximar las derivadas parciales como:

$$u_t = \frac{1}{\Delta t}(u_{i,j}^{n+1} - u_{i,j}^n) \quad (4.50)$$

$$v_t = \frac{1}{\Delta t}(v_{i,j}^{n+1} - v_{i,j}^n) \quad (4.51)$$

$$\nabla^2 u = \frac{1}{\Delta x \Delta y}(u_{i+1,j} + u_{i,j+1} + u_{i-1,j} + u_{i,j-1} - 4u_{i,j}) \quad (4.52)$$

$$\nabla^2 v = \frac{1}{\Delta x \Delta y}(v_{i+1,j} + v_{i,j+1} + v_{i-1,j} + v_{i,j-1} - 4v_{i,j}). \quad (4.53)$$

Sustituyendo estas aproximaciones en 4.45 y 4.46 se obtiene que la solución iterativa del GVF es como sigue:

$$\begin{aligned} u_{i,j}^{n+1} &= (1 - b_{i,j}\Delta t)u_{i,j}^n + \\ &+ r(u_{i+1,j}^n + u_{i,j+1}^n + u_{i-1,j}^n + u_{i,j-1}^n - 4u_{i,j}^n) + \\ &+ c_{i,j}^1\Delta t \end{aligned} \quad (4.54)$$

$$\begin{aligned}
v_{i,j}^{n+1} &= (1 - b_{i,j}\Delta t)v_{i,j}^n + \\
&+ r(v_{i+1,j}^n + v_{i,j+1}^n + v_{i-1,j}^n + v_{i,j-1}^n - 4v_{i,j}^n) + \\
&+ c_{i,j}^2\Delta t
\end{aligned} \tag{4.55}$$

donde

$$r = \frac{\mu\Delta t}{\Delta x\Delta y}. \tag{4.56}$$

La convergencia del proceso iterativo descrito queda garantizada por un resultado típico de la teoría de métodos numéricos. Suponiendo que  $b$ ,  $c^1$  y  $c^2$  son acotados, 4.54 y 4.55 son estables siempre que se cumpla la restricción de Courant-Friedrichs-Lewy sobre el tamaño del paso  $r \leq 1/4$ . Dado que normalmente  $\Delta x$ ,  $\Delta y$  y  $\mu$  son fijos, de la definición de  $r$  en 4.56 obtenemos que la siguiente restricción sobre el paso temporal debe cumplirse para garantizar la convergencia del GVF:

$$\Delta t \leq \frac{\Delta x\Delta y}{4\mu}. \tag{4.57}$$

### 4.3.3. Modelos deformables paramétricos 3D: Superficies activas

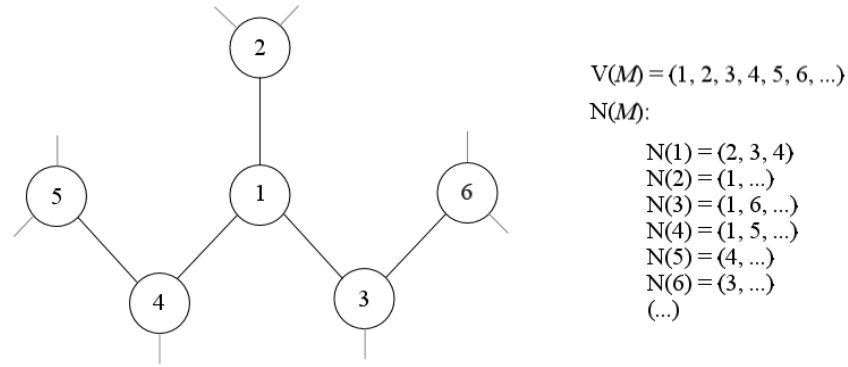
Si bien la formulación presentada en el apartado 4.3.1 es fácilmente generalizable al caso tridimensional, la implementación práctica de superficies activas presenta un conjunto de dificultades que abordamos en esta sección.

Podemos escribir un modelo deformable 3D como una superficie parametrizada tal que

$$\begin{cases} x_t(u; t) &= \alpha \nabla_u^2 x(u; t) - \beta \nabla_u^2 (\nabla_u^2 x(u; t)) + F_{ext}(x) \\ x(u; 0) &= x_0(u). \end{cases} \tag{4.58}$$

Para obtener una solución de esta ecuación diferencial en derivadas parciales, primero se debe representar la superficie deformable continua  $x$  como una malla discreta. Se puede resolver la ecuación 4.58 mediante métodos de elementos finitos o de Fourier. Sin embargo, dichos métodos son excesivamente lentos si el número de nodos de la malla es elevado. Por otra parte, los métodos de diferencias finitas basados en modelos de malla poligonal requieren tan sólo unas pocas operaciones aritméticas simples para cada nodo, siendo por tanto más adecuados que los anteriores para la implementación de modelos deformables con un elevado número de nodos, como es nuestro caso.

En este trabajo se ha implementado un modelo deformable 3D aplicando un método de diferencias finitas a una malla poligonal denominada *malla simplex*. La elección de este tipo de malla por encima de otras estructuras más simples, como las mallas rectangulares o triangulares, es que las primeras son poco adecuadas para el modelado



**Figura 4.6:** Esquema de una porción de una malla 2-simplex.

de estructuras cerradas, mientras que las segundas complican el cálculo de la energía interna.

A continuación se describe con detalle la construcción de la malla simplex, así como su codificación y la particularización para esta estructura de las ecuaciones que rigen el comportamiento del modelo deformable.

### Malla simplex

Una malla simplex es una malla poligonal en la que cada nodo (vértice) de la malla tiene un número constante de conexiones a sus vecinos. Una malla simplex puede definirse en cualquier dimensión. Una malla  $k$ -simplex se define como una malla  $(k+1)$ -conexa, donde cada vértice tiene exactamente  $(k+1)$  vecinos.

Formalmente, una malla  $k$ -simplex  $\mathcal{M}$  de  $\mathbf{R}^d$  se define como un par  $(V(\mathcal{M}), N(\mathcal{M}))$  donde  $V(\mathcal{M})$  es el conjunto de vértices de  $\mathcal{M}$  dados por

$$V(\mathcal{M}) = \{P_i\}, i \in \{1, \dots, n\}, P_i \in \mathbf{R}^d \quad (4.59)$$

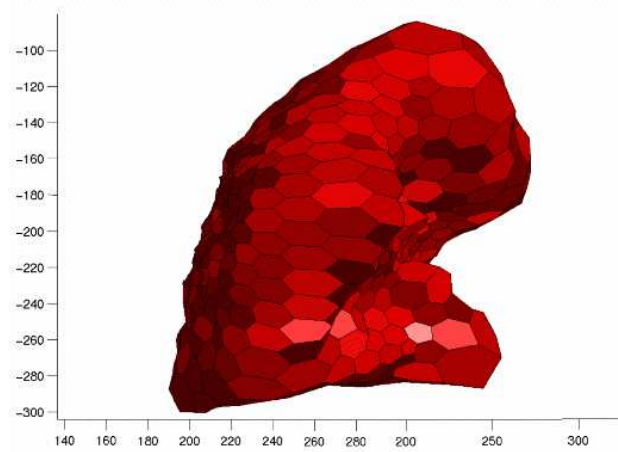
y  $N(\mathcal{M})$  es su función de conectividad asociada, definida como

$$N(\mathcal{M}) : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}^{k+1} \quad (4.60)$$

es decir, una función que a cada vértice de la malla le asigna otros  $k+1$  vértices, que denominaremos *vecinos* de dicho nodo. Esto se expresa como sigue

$$\forall i \in \{1, \dots, n\}, i \longmapsto (N_1(i), N_2(i), \dots, N_{k+1}(i)) \quad (4.61)$$

donde  $N_j(i)$  denota el  $j$ -ésimo vecino del nodo  $i$ . La figura 4.6 ilustra esto para el caso de un mallado 2-simplex.



**Figura 4.7:** Mallado 2-simplex de un pulmon extraido de un volumen CT.

Se restringen asimismo las posibles configuraciones que esta función de conectividad puede tomar, imponiendo que, para todo vértice, se cumpla lo siguiente:

$$\forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, k+1\}, \forall l \in \{1, \dots, k+1\}, l \neq j$$

$$(1) \quad N_j(i) \neq i \quad (4.62)$$

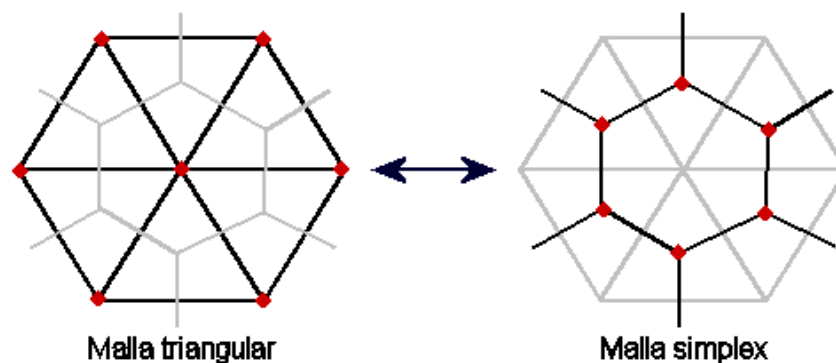
$$(2) \quad N_l(i) \neq N_j(i). \quad (4.63)$$

La condición 4.62 impide que la malla presente bucles, mientras que la 4.63 imposibilita la existencia de aristas dobles. Finalmente, se añade la restricción adicional de que una  $k$ -malla simplex debe estar conectada de tal modo que, para dos vértices cualquiera, exista un camino que los una. En este trabajo se considerarán únicamente mallas 2-simplex en  $\mathbf{R}^3$ .

De la definición de malla  $k$ -simplex sabemos que debe ser  $(k+1)$ -conexa, o, para el caso que nos ocupa, que la malla 2-simplex debe ser 3-conexa. Esta es una propiedad sumamente importante, pues permite simplificar y acelerar considerablemente la implementación de la superficie deformable. En la figura 4.7 podemos ver un ejemplo de malla 2-simplex. Mediante una malla 2-simplex se puede representar cualquier superficie tridimensional, y nos restringiremos a este tipo de mallado de aquí en adelante.

Otro tipo de malla poligonal frecuentemente utilizada en gráficos por computador es la malla triangular, cuyas caras están constituidas exclusivamente por triángulos. Las mallas triangulares son generalmente más fáciles de construir, habiendo sido extensivamente estudiadas durante años.

Puede demostrarse que una malla simplex puede generarse a partir de una malla triangular mediante una operación dual. Por definición, una operación dual sobre una



**Figura 4.8:** Ejemplo de operación dual para pasar de un mallado triangular a un mallado simplex y viceversa.

La malla poligonal construye una nueva malla donde cada uno de los vértices coincide con una de las caras de la malla original, y cada una de sus caras corresponde a uno de los vértices originales. La figura 4.8 muestra un ejemplo de dicha operación dual.

La dualidad entre mallas triangulares y mallas simplex permite generar de forma sencilla mallas simplex mediante el cálculo de la malla dual de una determinada malla triangular. En las siguientes secciones se detalla como generar dicha malla triangular así como la implementación de la operación dual.

### Creación de la malla triangular

A continuación se describen dos métodos bien conocidos para la generación de mallas triangulares. El primer método comienza con una malla triangular de poca resolución para luego subdividir recursivamente cada triángulo en cuatro triángulos menores hasta alcanzar la densidad de mallado deseada. El segundo método parte de una imagen tridimensional y calcula isosuperficies para generar mallas triangulares.

- Método por subdivisión:

Al aplicar una subdivisión a una malla triangular, todos los triángulos que constituyen sus caras son sustituidos por cuatro triángulos menores. Dicha operación puede repetirse recursivamente para obtener cualquier resolución que se desee. Una aplicación de la subdivisión es la generación de una malla triangular que aproxime una esfera tan detalladamente como sea posible, como se muestra en la figura 4.9. Esta malla puede utilizarse para inicializar una superficie deformable. Para crear dicha malla se puede empezar construyendo directamente un icosaedro, para luego recursivamente subdividir sus caras y proyectar los nuevos vértices sobre la esfera. De este modo, con cada nueva iteración se obtiene una malla que aproxima mejor la esfera.

Un icosaedro es un sólido platónico  $P_5$  que tiene 12 vértices, 30 aristas y 20 caras triangulares equivalentes,  $20\{3\}$ . Un método para construir un icosaedro de lado  $a = \sqrt{50 - 10\sqrt{5}}/5$  consiste en situar los vértices extremos en  $(0, 0, \pm 1)$  y los vértices centrales a lo largo de dos circunferencias decaladas de radio  $\frac{2}{5}\sqrt{5}$  y alturas  $\pm\frac{1}{5}\sqrt{5}$ . Mediante una rotación adecuada los vértices de un icosaedro de lado  $a = 2$  pueden también colocarse en  $(0, \pm\phi, \pm 1), (\pm 1, 0, \pm\phi)$  y  $(\pm\phi, \pm 1, 0)$ , donde  $\phi$  es el *golden ratio* ( $\phi = 1,6180339$ ).

■ Método por isosuperficies:

Una *isosuperficie* es una superficie que pasa por todos los puntos del espacio donde un volumen de datos continuo toma un cierto valor constante, denominado umbral de isosuperficie. El resultado de la mayoría de algoritmos de isosuperficie es una superficie teselada mediante una malla triangular. El algoritmo *marching cubes* propuesto por Lorensen y Cline [54] ha sido extensamente utilizado en los últimos años. Este algoritmo y sus variantes se han convertido en el estándar para la generación de isosuperficies, siendo utilizados en numerosos paquetes de software comerciales de visualización. A continuación se presenta una descripción de funcionamiento de dicho algoritmo.

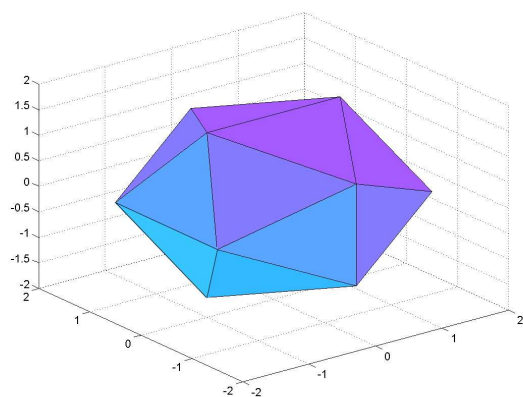
El *marching cubes* trabaja dividiendo el volumen de datos en *voxels*. Dichos voxels son paralelepípedos de dimensiones constantes, y tienen por tanto 8 vértices. En función de si el valor intensidad en un vértice es mayor o menor que el umbral de isosuperficie se etiqueta dicho vértice con un 1 o un 0 respectivamente. Es sencillo comprobar que hay en total 256 posibles configuraciones en las que una superficie puede cruzar un determinado voxel. Si se tienen en cuenta la simetría rotacional y la complementariedad, dicho número se reduce a tan solo 14 configuraciones. En función de dichas configuraciones se construyen uno o más triángulos para dicho cubo. Este proceso se aplica secuencialmente a todos los voxels hasta haber recorrido íntegramente el volumen de datos.

A pesar de ser el *marching cubes* un algoritmo efectivo, se ha demostrado que en ciertos casos puede producir agujeros en la superficie. Una variante del algoritmo llamada *Alligator* resuelve dicho problema utilizando información de conectividad para restringir la partición en triángulos de cada voxel.

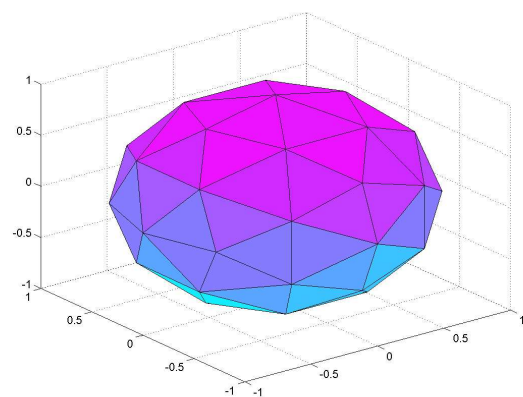
## Implementación de la operación dual

Si bien la definición que se ha hecho más arriba de la operación dual es sencilla, su implementación no lo es tanto. Si la malla poligonal no está representada de forma conveniente, la operación puede resultar complicada y costosa en tiempo. Asimismo, dicha representación debe ser lo bastante genérica como para describir tanto mallas triangulares (con número de aristas por vértice variable) como mallas simplex (con número de aristas por cara variable). Para hacer frente a este problema se ha empleado un tipo de representación de mallas, común en gráficos por ordenador, denominada

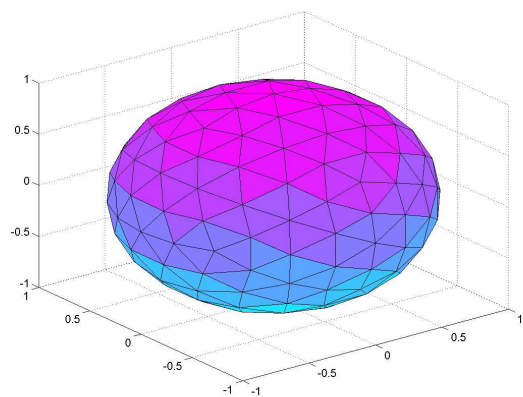




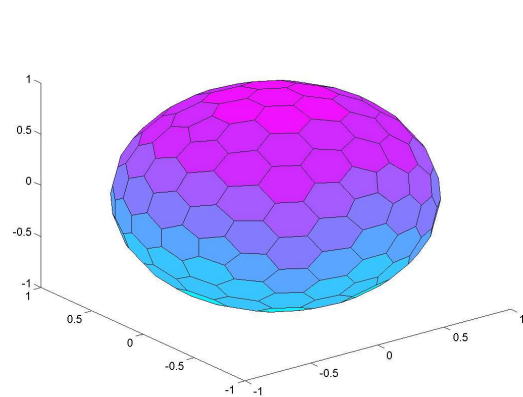
(a)



(b)



(c)



(d)

**Figura 4.9:** Construcción de la malla triangular por subdivisión de triángulos: Icosaedro (a), primera subdivisión (b), segunda subdivisión (c) y malla símplex construida a partir de esta última (d).

*winged-edge*. A continuación se incluye una breve descripción de la representación *winged-edge* de mallas que se ha utilizado, así como la implementación de la operación dual para esta estructura.

**La representación *winged-edge*:** Sea  $\mathcal{M} = \{V, E, F\}$  una malla donde  $V$  es el conjunto de vértices,  $E$  es el conjunto de aristas y  $F$  es el conjunto de caras. Denotamos asimismo el numero de vertices, aristas y caras como  $K_V$ ,  $K_E$  y  $K_F$  respectivamente. Una representación *winged-edge* de la malla  $\mathcal{M}$  queda descrita por las siguientes listas

$$\begin{aligned}
 \text{Lista de Vértices: } \{V_i\} &= \{(x_i, y_i, z_i)\}, & i &= 1, 2, \dots, K_V \\
 \text{Lista de Aristas: } \{E_i\} &= \{(v_{i_1}, v_{i_2})\}, & i &= 1, 2, \dots, K_E \\
 \text{Lista de Caras: } \{F_i\} &= \{(n_i, e_{i_1}, e_{i_2}, \dots, e_{i_{n_i}})\}, & i &= 1, 2, \dots, K_F \\
 \text{Vecinos Aristas: } \{N(E_i)\} &= \{(f_{i_1}, f_{i_2})\}, & i &= 1, 2, \dots, K_E \\
 \text{Vecinos Vértices: } \{N(V_i)\} &= \{(m_i, e_{i_1}, e_{i_2}, \dots, e_{i_{m_i}})\}, & i &= 1, 2, \dots, K_V
 \end{aligned} \tag{4.64}$$

donde  $V_i$ ,  $E_i$  y  $F_i$  son el vértice, arista y cara  $i$ -ésimos respectivamente.  $(x_i, y_i, z_i)$  son las coordenadas de  $V_i$ .  $v_i$ ,  $e_i$  y  $f_i$  son los índices de  $V_i$ ,  $E_i$  y  $F_i$  en sus respectivas listas.  $\{N(E_i)\}$  lista los planos adyacentes a una arista.  $\{N(V_i)\}$  lista el conjunto de aristas que parten de un vértice.  $n_i$  y  $m_i$  son respectivamente el número de elementos en  $\{F_i\}$  y  $\{N(V_i)\}$ . Finalmente, se impone a los elementos de los vectores que constituyen las listas  $\{F_i\}$  y  $\{N(V_i)\}$  un orden inverso al sentido de las agujas del reloj respecto al vector normal de la cara o vértice que apunta hacia el exterior de la superficie.

Este tipo de representación permite una implementación eficiente de diversas operaciones sobre la malla. A continuación se define un conjunto de operaciones necesarias para obtener la malla dual.

Dado un índice cualquiera  $i$ , podemos definir los operadores de búsqueda en la lista como

$$\begin{aligned}
 V(i) &= V_i \\
 E(i) &= E_i \\
 F(i) &= F_i.
 \end{aligned} \tag{4.65}$$

El operador que calcula el índice que sigue a  $i$  en una lista de  $n$  elementos será

$$\text{next}(i, n) = \text{mod}(i, n) + 1 \tag{4.66}$$

El operador que obtiene el vértice  $j$ -ésimo de la cara  $F_i$  se define como

$$V_j(F_i) = V(E_{i_j} \cap E_{i_{\text{next}(j, n_i)}}) = V(E(e_{i_j}) \cap E(e_{i_{\text{next}(j, n_i)}})). \tag{4.67}$$

Definimos asimismo el operador que calcula el centro de una cara como

$$\text{center}(F_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} V_j(F_i). \tag{4.68}$$

Empleando los operadores definidos, se puede calcular la malla dual  $\hat{\mathcal{M}} = \{\hat{V}, \hat{E}, \hat{F}\}$  de la malla  $\mathcal{M}$  ejecutando los pasos siguientes:

1. Vértices:  $\hat{V}_i = \text{center}(F_i)$
2. Aristas:  $\hat{E}_i = N(E_i)$
3. Caras:  $\hat{F}_i = N(V_i)$
4. Vecinos de las aristas:  $N(\hat{E}_i) = E_i$
5. Vecinos de los vértices:  $N(\hat{V}_i) = F_i$

### Implementación de la superficie deformable

Si se discretiza la ecuación dinámica 4.58 en la malla simplex mediante el método de las diferencias finitas, se obtiene la ecuación siguiente para la evolución de un vértice arbitrario  $x$

$$x^t = x^{t-1} + \tau(\alpha\tilde{\Delta}x^{t-1} - \beta\tilde{\Delta}^2x^{t-1} + F_{ext}(x^{t-1})) \quad (4.69)$$

donde  $\tilde{\Delta}x$  y  $\tilde{\Delta}^2x$  representan respectivamente el operador Laplaciano discreto y el operador biarmónico discreto implementados sobre una malla simplex. Sus definiciones son en este caso

$$\tilde{\Delta}x = \frac{1}{3}(x^1 + x^2 + x^3) - x \quad (4.70)$$

$$\tilde{\Delta}^2x = \tilde{\Delta}(\tilde{\Delta}x) = \frac{1}{9}\left(\sum_{i=4}^9 x^i - 6\sum_{i=1}^3 x^i + 12x\right) \quad (4.71)$$

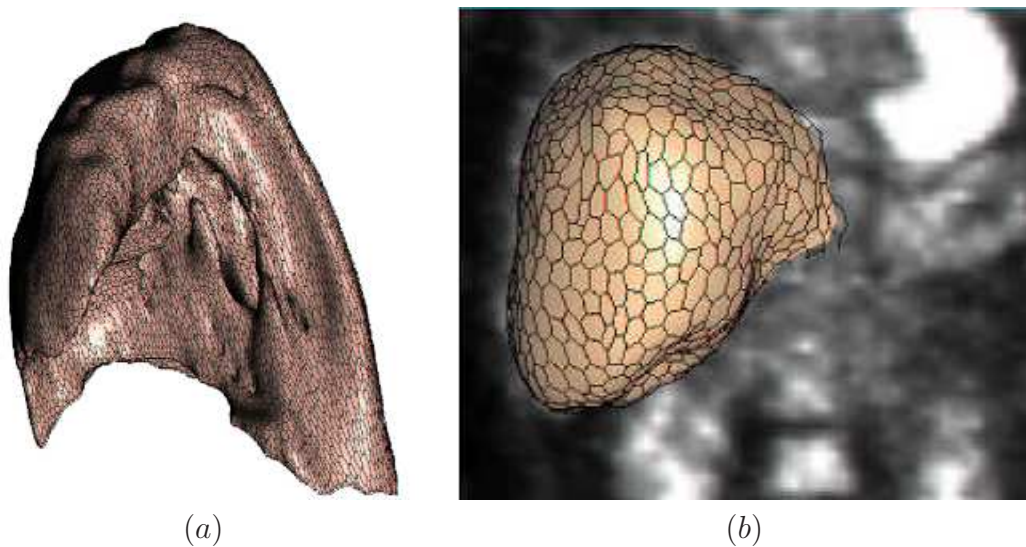
donde  $x^i, i = 1, 2, \dots, 9$  son los vecinos del vértice  $x$ .

Igual que para el caso de los contornos deformables, el término de energía externa  $F_{ext}$  puede obtenerse para cualquier posición de  $x$  mediante interpolación de los valores de la misma en las posiciones discretas contiguas a  $x$ . Es en este punto donde la regularidad y simplicidad de la malla simplex se hace notar, permitiendo una implementación sumamente eficiente de la superficie deformable.

## 4.4. Conclusión

En el presente capítulo se ha descrito el método desarrollado para extraer las distintas estructuras anatómicas de interés de los volúmenes CT y PET, detallándose cada una de las técnicas empleadas.

El diseño de este método ha estado condicionado por un conjunto de necesidades particulares de la presente aplicación, que lo apartan de los sistemas típicos de segmentación de imágenes anatómicas.



**Figura 4.10:** Pulmón (a) y hígado (b) extraídos de un volumen CT y PET, respectivamente, mediante un snake 3D símplex.

Así, se ha priorizado por encima de todo que el sistema sea robusto, incluyéndose numerosas rutinas de verificación de los resultados que permiten detectar y corregir de forma totalmente automática los errores de segmentación más comunes.

Por otra parte, se relega a un segundo plano la precisión de los órganos extraídos, dado que las imprecisiones así introducidas no se propagarán al resultado final. Esta condición ha permitido recurrir a técnicas de segmentación relativamente poco complejas, lo suficientemente genéricas para no verse afectadas por la gran variabilidad anatómica de las regiones estudiadas.

Finalmente, la estructuración en forma de búsqueda jerárquica hace posible restringir las sucesivas etapas de clasificación de modo que se saque partido de las relaciones espaciales entre los distintos órganos, aprovechando la información obtenida de la segmentación de las estructuras más visibles para facilitar la localización de aquellas de mayor complejidad.

Los resultados obtenidos demuestran que el método de segmentación de estructuras anatómicas desarrollado es capaz de extraer satisfactoriamente de ambas imágenes la información necesaria para que una etapa de registro basado en dicha segmentación converja hacia una buena aproximación de la posición de óptima correspondencia.

En el siguiente capítulo se discutirá la problemática asociada a la evaluación del registro elástico y se presentarán los resultados que prueban la eficacia del método propuesto.

# Capítulo 5

## Validación

En este capítulo se aborda el problema de la validación de los resultados obtenidos con los sistemas de registro de imágenes descritos en las secciones precedentes. Se trata de una tarea en absoluto trivial, dado que no existe una metodología comúnmente aceptada para medir el error en un registro elástico.

En primer lugar se presentarán las medidas numéricas empleadas para cuantificar la bondad del alineamiento entre dos volúmenes, estudiándose que información pueden aportar y justificando su conveniencia. Dichas medidas se aplicarán a los resultados obtenidos mediante el sistema de registro propuesto, obteniéndose de este modo una representación objetiva del rendimiento del mismo.

Dado que las anteriores medidas no bastan para evaluar el comportamiento del método frente a un caso clínico real, se describirá un protocolo diseñado para permitir una validación semicuantitativa de nuestros resultados por parte de oncólogos especialistas.

Finalmente, se mostraran los resultados de dicha validación, junto con algunos ejemplos, que prueban el buen funcionamiento del sistema de registro desarrollado en este trabajo de tesis.

### 5.1. Problemática

La validación cuantitativa de los resultados de una aplicación de *registro retrospectivo* (esto es, que no precisa de elementos de referencia implantados en el paciente), es, en general, un problema complejo, que ha sido objeto de numerosos congresos multidisciplinares, pero para el que no se ha alcanzado una solución globalmente aceptada.

El primer método en que se podría pensar, esto es, el uso de referencias anatómicas internas definidas por un especialista, ha demostrado ser demasiado impreciso en el caso que nos ocupa, debido a la excesiva diferencia entre las modalidades implicadas. Por ejemplo, a pesar de que las estructuras óseas son fácilmente identificables en el volumen CT, las estructuras visibles correspondientes que aparecen en el volumen PET son en

realidad el tejido vivo del interior de dichos huesos, por lo que cualquier medida basada en el emparejado de dichas referencias conllevará un cierto grado de imprecisión.

Se han realizado ciertos estudios [55] destinados a medir el error cometido por distintas aplicaciones de registro, empleando como referencia los parámetros de un supuesto registro óptimo obtenido mediante marcadores implantados (*gold standard*). Sin embargo, este tipo de criterio basado en referencias artificiales es poco adecuado para evaluar la calidad de un registro elástico, debido a su gran variabilidad, y no se tiene noticia de que exista un *gold standard* para el registro de la zona torácica o abdominal.

Otra opción es la comprobación visual de los resultados por parte de personal médico especializado. En un trabajo realizado por Wong et al. [56] se estudia la capacidad de detectar por este método errores de traslación y rotación en el registro de imágenes MRI-PET. Sus resultados determinan que el observador es capaz de apreciar errores traslacionales mayores de  $2\text{-}3\text{mm}$  y errores rotacionales mayores de  $3\text{-}4^\circ$ . A pesar de no ser estos datos directamente aplicables a la evaluación de un registro elástico, muy limitada debido a la incapacidad de efectuar una representación de los resultados que capte las complejas deformaciones tridimensionales presentes, nos hacen pensar que si es posible realizar estimaciones a nivel local de los errores de traslación entre estructuras anatómicas concretas.

En definitiva, si bien el único modo de conseguir medidas fiables de la bondad de esta aplicación será pasar por un procedimiento de validación clínica con marcadores implantables en número suficiente para captar adecuadamente las deformaciones que se experimenten, la definición de protocolo de evaluación por parte de personal médico especializado puede ofrecer valiosa información sobre el comportamiento del registro.

## 5.2. Base de datos

Para la realización de este trabajo de tesis, se ha recopilado una base de datos consistente en dieciocho estudios CT-PET de la región torácica y abdominal. Dichas imágenes han sido suministradas principalmente por *LifeScan Louisville*, el *Hospital Percy* y el *H.I.A. de Val de Grâce*.

El problema de utilizar imágenes provenientes de distintos centros es la gran diferencia de formatos y calidades, debida a los distintos protocolos e instrumental involucrados en cada adquisición. A pesar de ello, se ha optado por no reformatear estos datos a un estándar común, prefiriéndose diseñar el sistema de tal modo que sea independiente de los parámetros de adquisición.

De este modo, la base de datos está constituida por volúmenes CT, con aproximadamente  $(256 \times 256)$  voxels en el plano  $xy$ , y entre 60 y 125 cortes, con resoluciones de aproximadamente  $(1,0 \times 1,0 \times 5,5)\text{mm}^3$ , y por sus correspondientes imágenes PET, que comprenden un volumen de transmisión y otro de atenuación de idénticas dimensiones, que oscilan alrededor de los  $(144 \times 144)$  pixels en el plano  $xy$  y de 160 a 230 cortes, con resolución entorno a los  $(4,0 \times 4,0 \times 4,0)\text{mm}^3$ .

### 5.3. Medidas numéricas

Otra alternativa para obtener medidas cuantitativas de la calidad del registro es emplear los mismos criterios de similitud que se describían en el capítulo 2.4.3. Si bien estas medidas no permiten comparar el registro obtenido sobre pares de imágenes distintos, constituyen un método sencillo para evaluar los resultados obtenidos empleando diferentes técnicas de registro, dado un determinado par CT-PET.

Se han implementado dos medidas de este tipo, a fin de confirmar las mejoras predichas teóricamente para nuestro método de registro. Estas medidas son la *información mutua normalizada*, tal como se describe en el apartado 2.4.3, y el *solapamiento entre estructuras*, definido como

$$SE = \frac{E_A \cap E_B}{E_A \cup E_B}, \quad (5.1)$$

donde  $E_A$  y  $E_B$  representan los conjuntos de voxels pertenecientes a una misma estructura anatómica segmentada en una y otra imagen.

Tal como muestra la figura 5.1, se produce una mejora significativa de los resultados cuando el sistema se inicializa con la transformación elástica calculada a partir de las estructuras segmentadas, en lugar de emplear un método típico de registro rígido o una inicialización manual. Cabe mencionar que la significativa disparidad del paciente número 7 frente al resto de resultados se debe a la calidad significativamente menor de estas imágenes, así como a la importante disparidad en la posición del paciente entre ambas adquisiciones.

Estos resultados, si bien reflejan claramente tanto la viabilidad del método de registro propuesto (mejora del criterio de similitud), como la conveniencia de emplear el esquema de inicialización elástica basada en estructuras anatómicas (mejora del solapamiento), no garantizan que el comportamiento del sistema sea el requerido por la aplicación clínica para la que se ha diseñado.

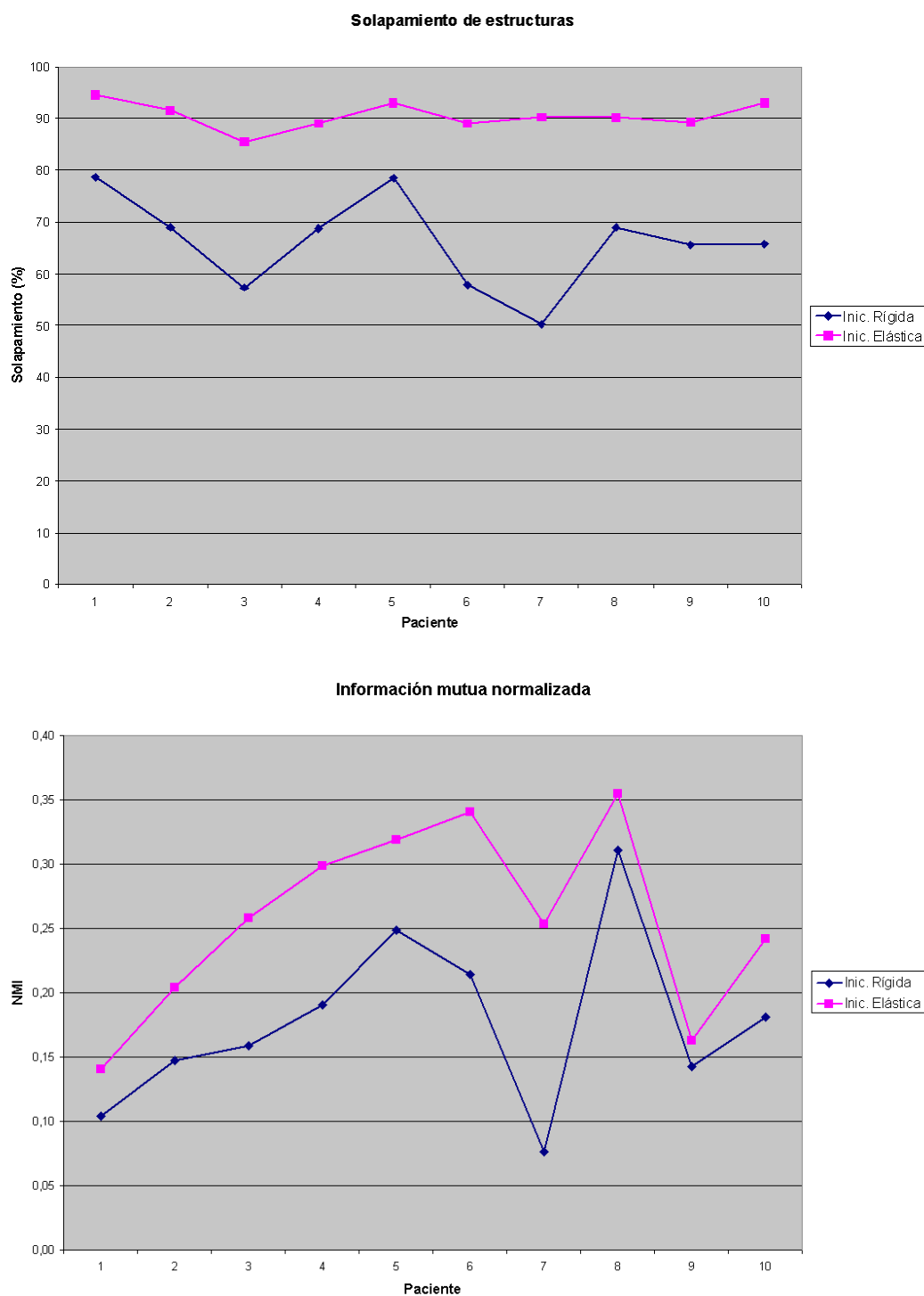
A fin de comprobar este punto, es preciso recurrir a una validación de los resultados por parte de personal médico especializado. No obstante, para realizar dicha validación debe definirse un protocolo que permita garantizar la objetividad y repetitividad de las medidas efectuadas.

### 5.4. Protocolo de evaluación

Se ha definido un protocolo de evaluación del registro elástico, bajo la dirección del *Dr. Hervé Foehrenbach* del *Hôpital d'Instruction des Armées du Val de Grace (Paris)*. Este protocolo permite que personal médico especialista genere una medida semicuantitativa de la precisión del registro sin necesidad de revisar la totalidad del volumen de datos fusionados.

Con este fin, se generan un conjunto de cortes anatómicamente significativos del volumen CT y su correspondiente volumen PET registrado, que se presentan al





**Figura 5.1:** Medidas numéricas de la calidad del registro obtenido para diez pacientes distintos. Se aprecia con claridad como en todos los casos el método de inicialización elástica propuesto mejora de forma significativa el resultado del registro.

**Cuadro 5.1:** Comprobación de la repetitividad del método de evaluación.

Region	Mean	Variance
Chest Wall	0.670	0.02
Mediastinal Wall (heart)	0.935	0.09
Diaphragmatic Wall (liver)	0.720	0.11

especialista. Esta representación se ha preferido a la de una única imagen fusionada, puesto que según nuestra experiencia así se reduce la subjetividad de las medidas.

Cada par de imágenes ha sido marcado con un conjunto de líneas graduadas que definen ciertos puntos de referencia a su cruce con las estructuras anatómicas más significativas. Así, para el caso de una aplicación de registro de imágenes torácicas, se evaluará en los distintos cortes la posición de la caja torácica, la pared mediastinal (que va ligada a la posición del corazón) y la pared diafragmática (que va ligado a la posición del hígado). Estas referencias permiten al especialista medir las diferencias de posición entre las estructuras mencionadas en una y otra imagen, y puntuar su registro en una escala que va de 0 (buen registro, error menor de 5mm) a 2 (registro inaceptable, error mayor de 15mm).

La figura 5.2 ilustra el proceso de evaluación del registro empleando las referencias sobreimpresas en la imagen.

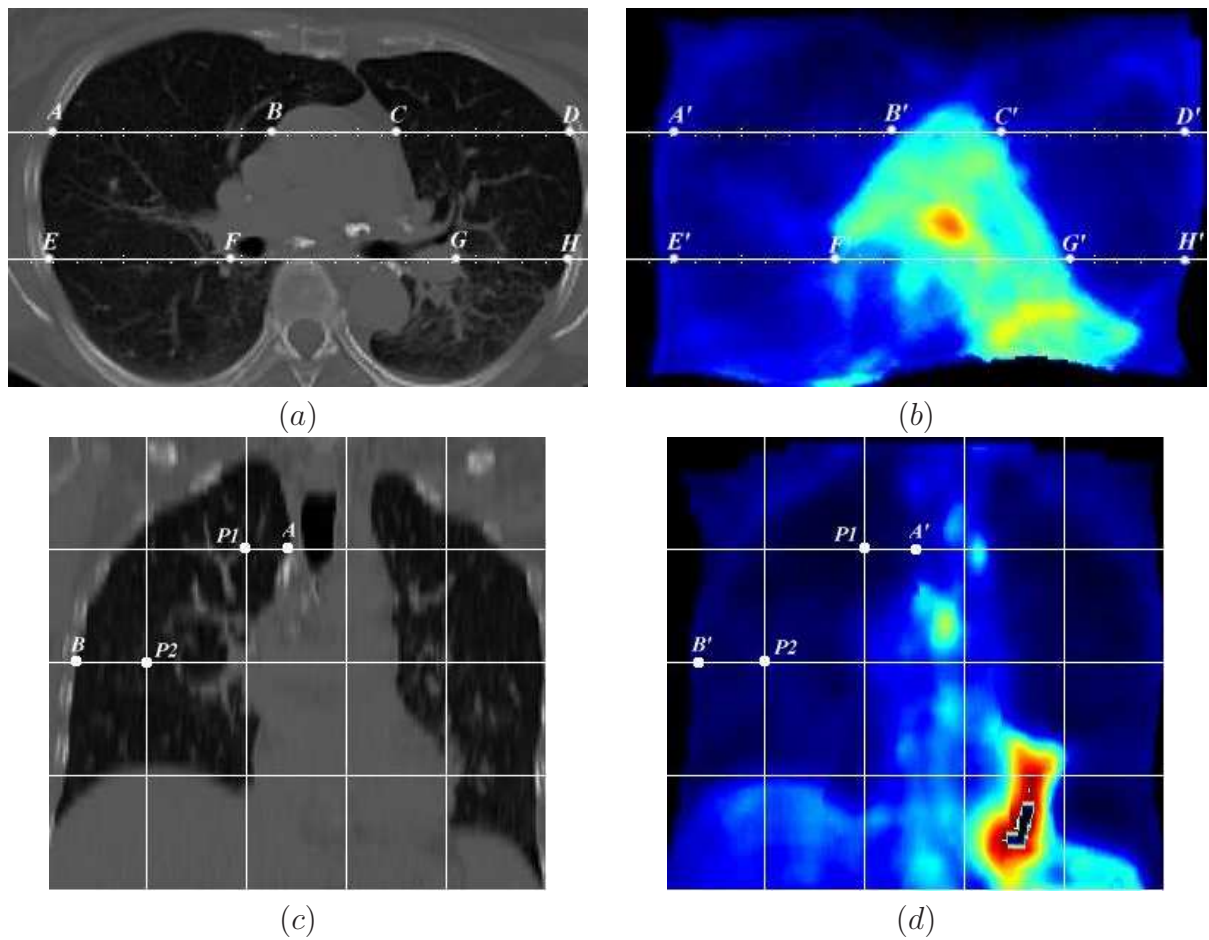
Para poner a prueba este protocolo, se ha pedido al siguiente grupo de médicos de distintos centros, todos ellos con una amplia experiencia en oncología, que realicen la evaluación del registro un par de volúmenes CT-PET:

- Dr. Hervé Foehrenbach, *Hôpital d'Instruction des Armées du Val de Grace (Paris)*
- Dr. Pierre Rigo, *Centre Hospitalier Princesse Grace (Monaco)*
- Dr. Xavier Marchandise, *Centre Hospitalier Régional Universitaire de Lille (Lille)*

Dicha evaluación se ha llevado a cabo de forma independiente, sin que ninguno de los participantes conociera de antemano las imágenes ni los resultados de los demás. Con tal fin, una interfaz gráfica permitía realizar la evaluación a través de internet, de forma que los resultados eran transmitidos y archivados automáticamente. La figura 5.3 muestra una captura de dicho interfaz.

Los resultados obtenidos, tal como se muestra en el cuadro 5.1, indican que el procedimiento descrito proporciona una medida lo bastante objetiva como para ser empleada en la validación de nuestro método de registro.

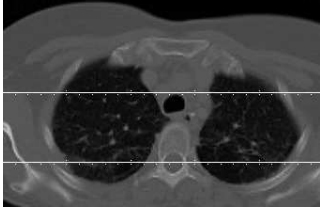
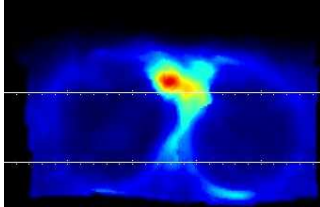
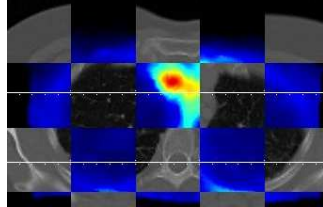
Una vez demostrada su validez, se ha empleado el procedimiento descrito para estimar la calidad del registro obtenido mediante el método desarrollado en esta tesis. El cuadro 5.2 muestra los resultados desglosados en las distintas regiones anatómicas de interés.

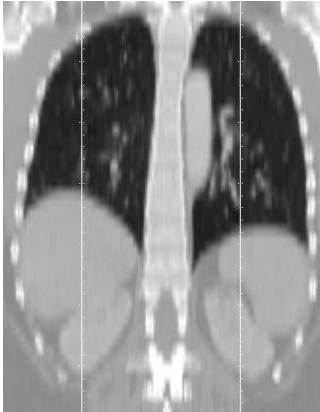
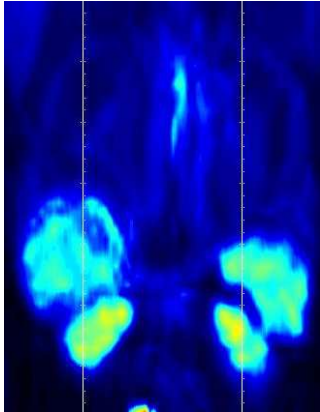
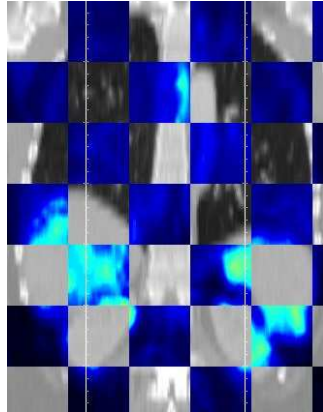


**Figura 5.2:** Ejemplo de evaluación del registro: Cortes axiales del volumen CT de referencia (a) y PET registrado (b) a los que se han superpuesto las referencias graduadas; cortes coronales de los mismos (c)(d) donde se muestra cómo medir el alineamiento de estructuras correspondientes.

**Cuadro 5.2:** Evaluación del registro de las estructuras de interés.

Region	Mean	Variance	Max	Min
Chest Wall	0.615	0.01	0.63	0.60
Mediastinal Wall (heart)	0.758	0.15	1.44	0.54
Diaphragmatic Wall (liver)	0.467	0.15	0.87	0.16

		
Anterior Left Chest Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Anterior Right Chest Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	
Anterior Left Mediastinal Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Anterior Right Mediastinal Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	
Posterior Left Mediastinal Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Posterior Right Mediastinal Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	
Posterior Left Chest Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Posterior Right Chest Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	

		
Left Superior Chest Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Right Superior Chest Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	
Left Diaphragmatic Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Right Diaphragmatic Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	
Inferior Liver Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Right Superior Mediastinal Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	
Superior Left Kidney Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Superior Right Kidney Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	
Inferior Left Kidney Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	Inferior Right Kidney Wall : <input type="radio"/> <0.5cm <input type="radio"/> <1.5cm <input type="radio"/> >1.5cm	

Submit

**Figura 5.3:** Ejemplo del interfaz de evaluación evaluación del registro (aquí reducido a un corte axial y otro coronal). Cada grupo de imágenes consiste en un corte del volumen CT de referencia, su corte correspondiente en el volumen PET registrado y una fusión por bloques de los mismos. Un conjunto de preguntas tipo test permiten puntuar el registro de distintas estructuras anatómicas significativas.

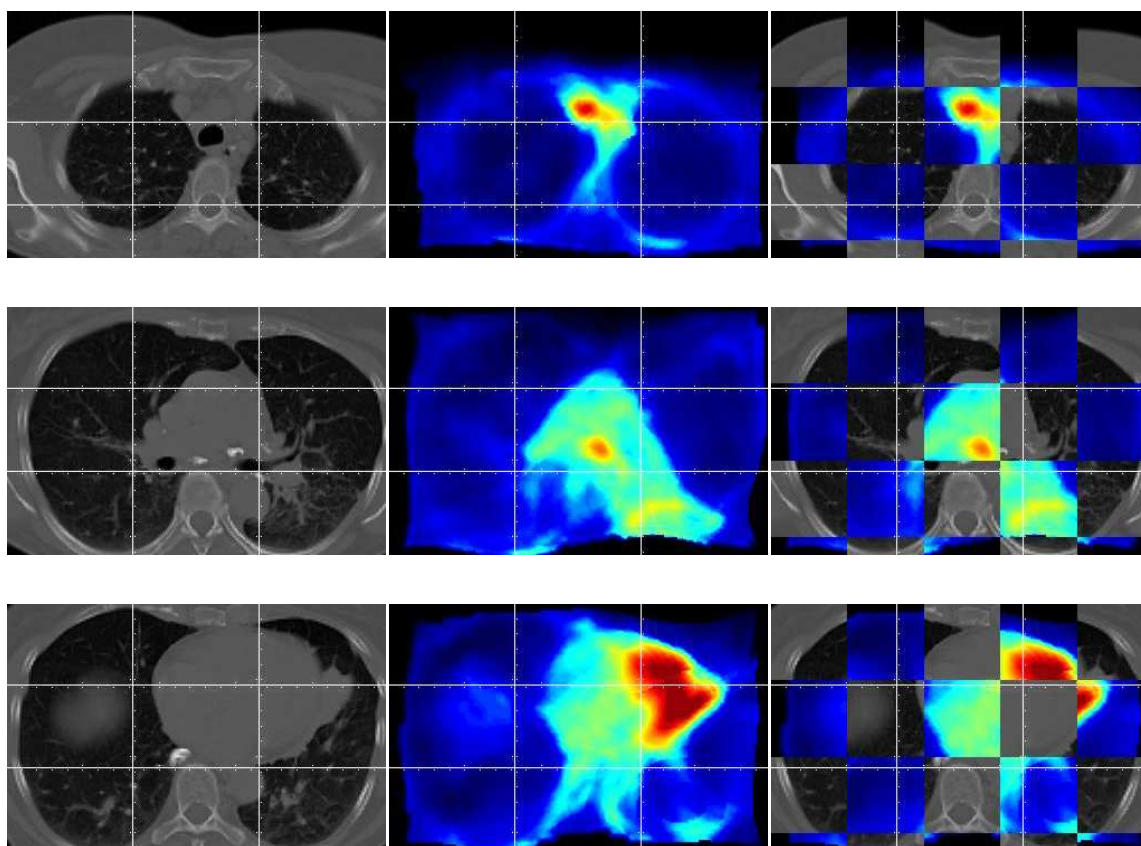
Se comprueba observando estos resultados que en todos los casos la media de los errores detectados se mantiene alrededor de un centímetro, y por tanto dentro del margen establecido como aceptable por los médicos.

En particular se han obtenido resultados excelentes para la caja torácica, como cabía esperar, y, felizmente, el diafragma, una de las estructuras donde más crítico era el registro, tanto por las grandes deformaciones que experimenta como por el peligro de que un mal registro cree incertidumbre entre un cáncer pulmonar o de hígado.

Por otra parte, el registro de la zona mediastinal es ligeramente irregular en sus resultados, y hace pensar que tal vez sería deseable incorporar nueva información anatómica que cubra esa región, tal vez mediante una segmentación del corazón.

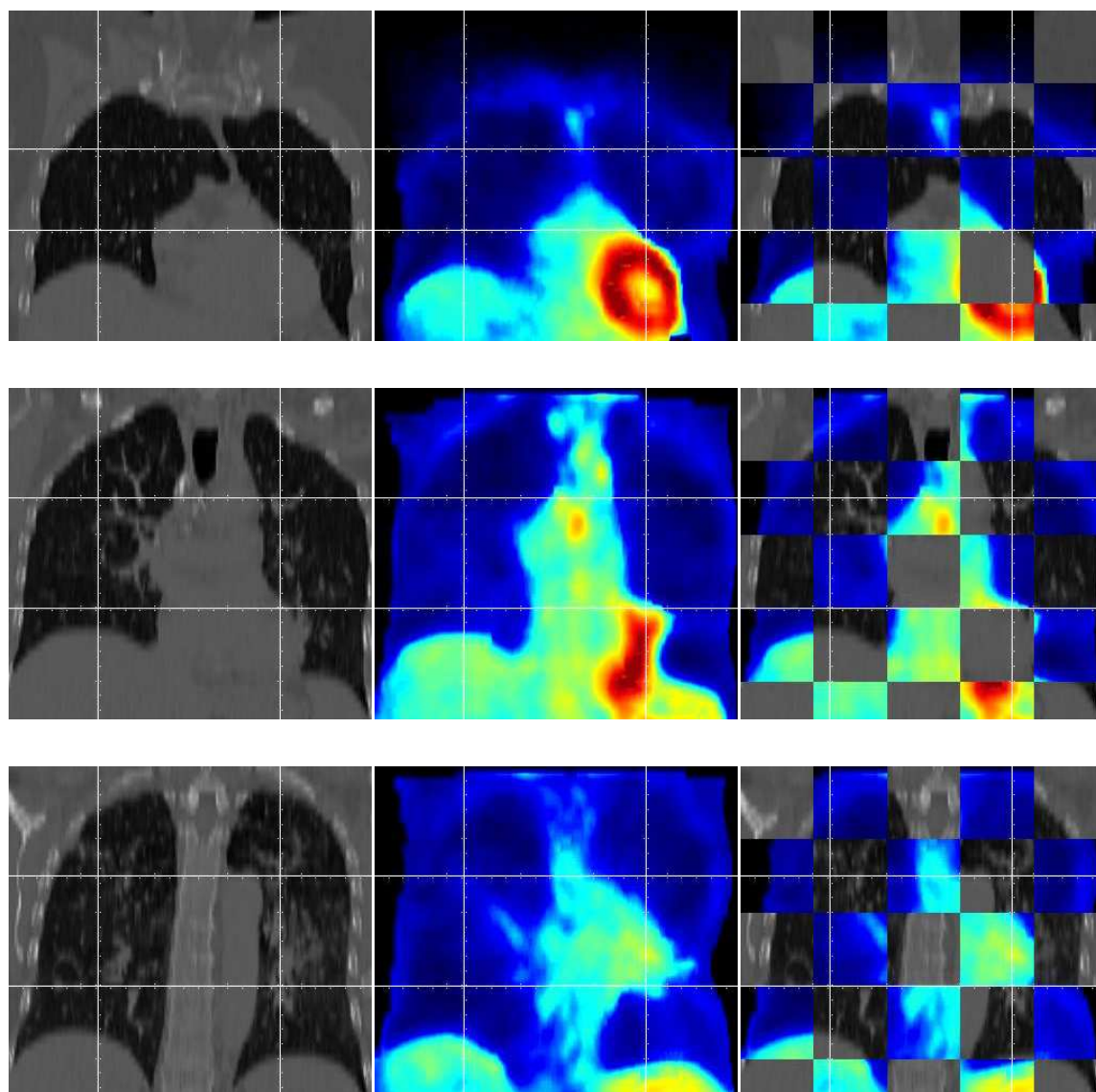
Para acabar, se presentan a continuación algunos resultados del registro para distintos pacientes (figuras 5.4 a 5.11), tal como se muestran a los especialistas durante el proceso de evaluación.

En cada caso se presentan tres cortes axiales y tres coronales de los volúmenes CT y PET registrado (las imágenes PET originales no se muestran por ser tan distintas de las anteriores que no es posible ofrecer cortes equivalentes). Se incluye asimismo para cada corte una imagen de superposición por bloques, para facilitar la comparación de los contornos de las distintas estructuras anatómicas.



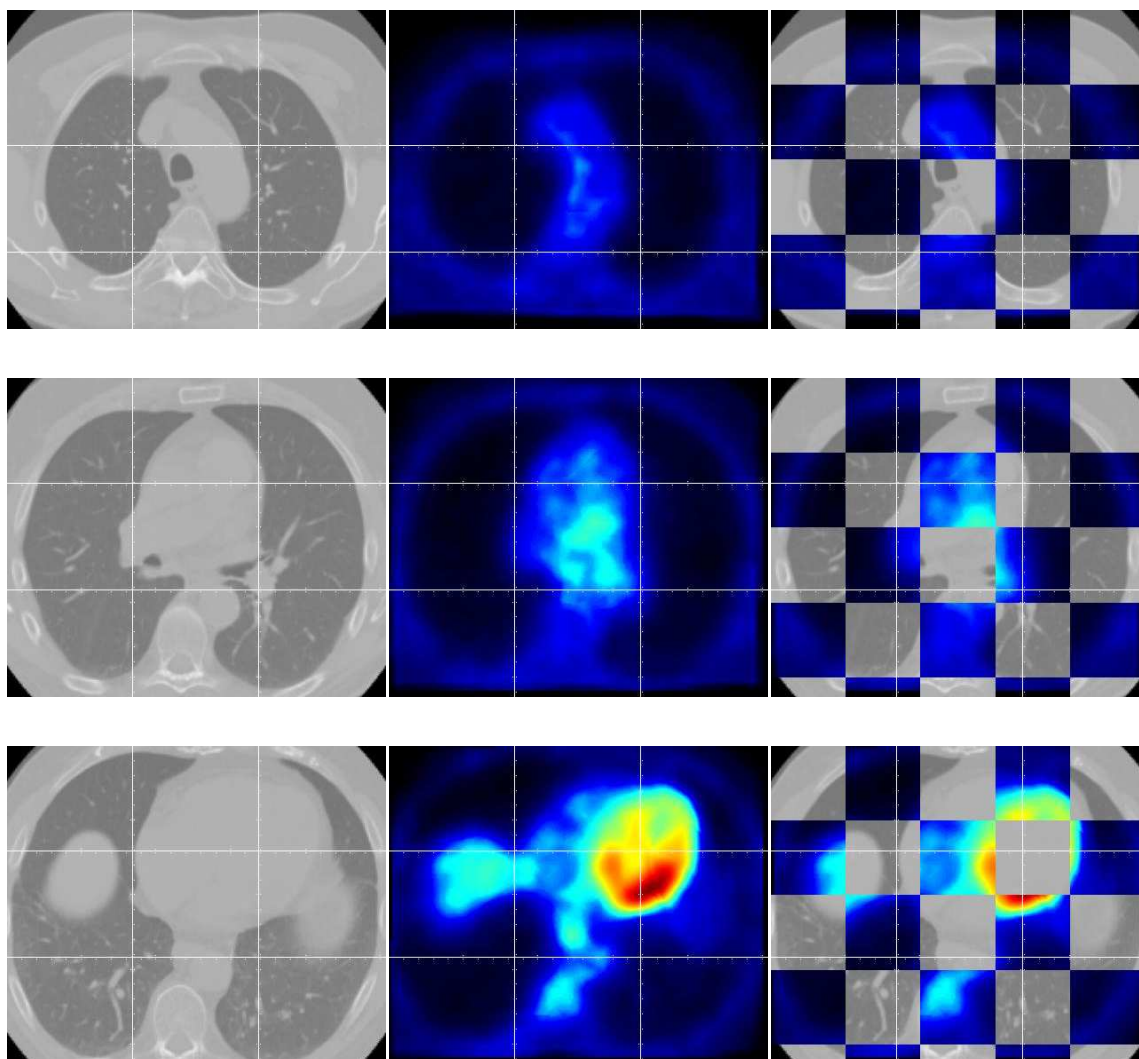
**Figura 5.4:** Paciente 1: Resultados del registro (cortes axiales).



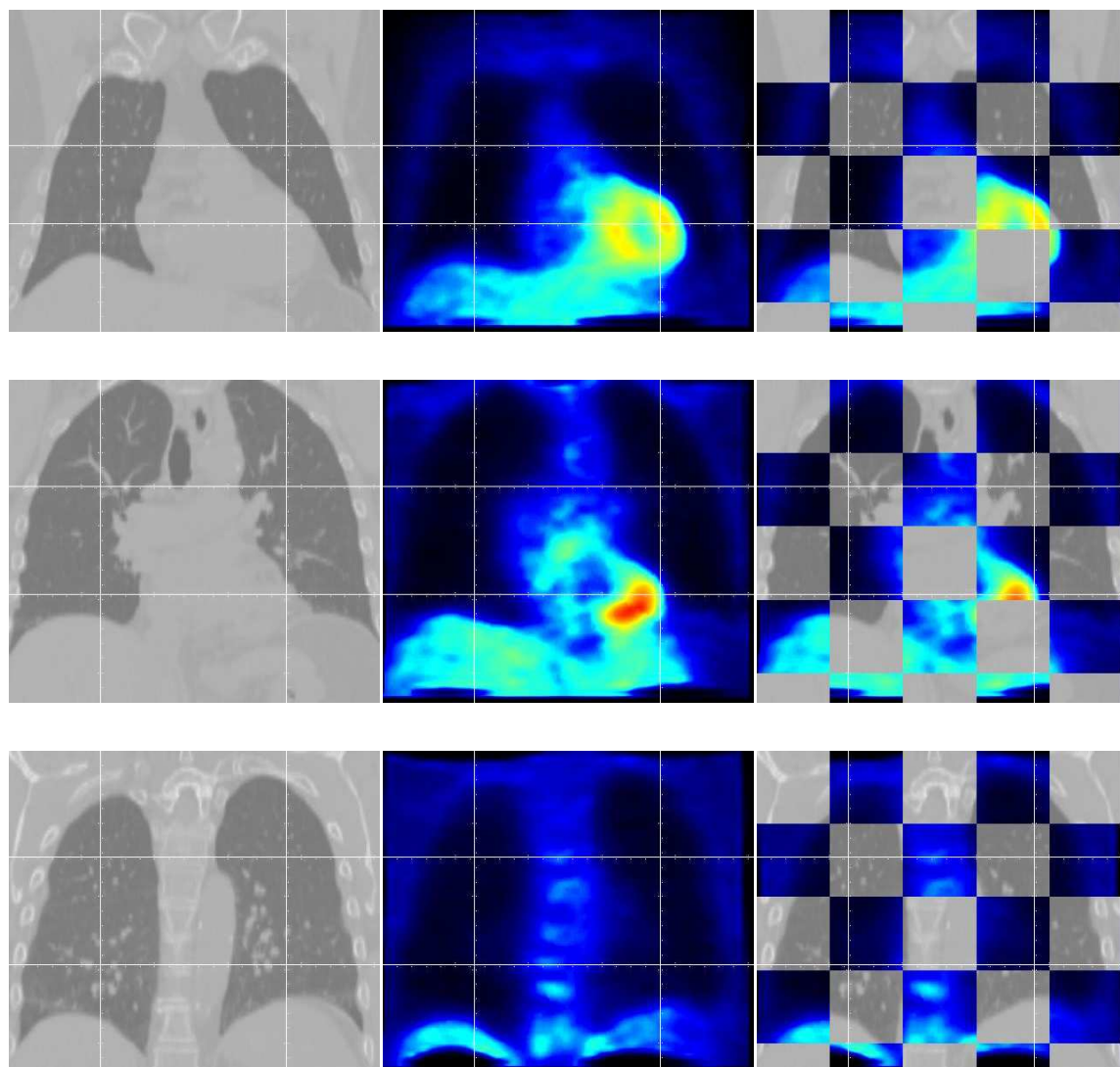


**Figura 5.5:** Paciente 1: Resultados del registro (cortes coronales).

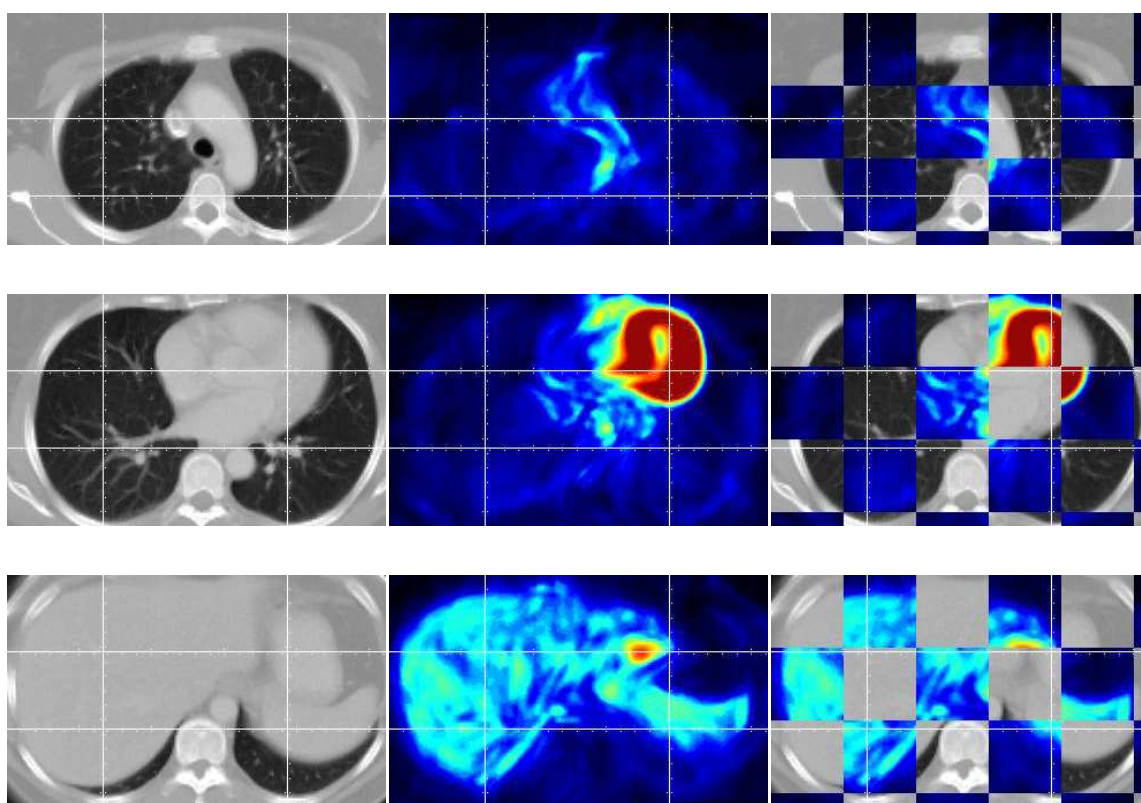




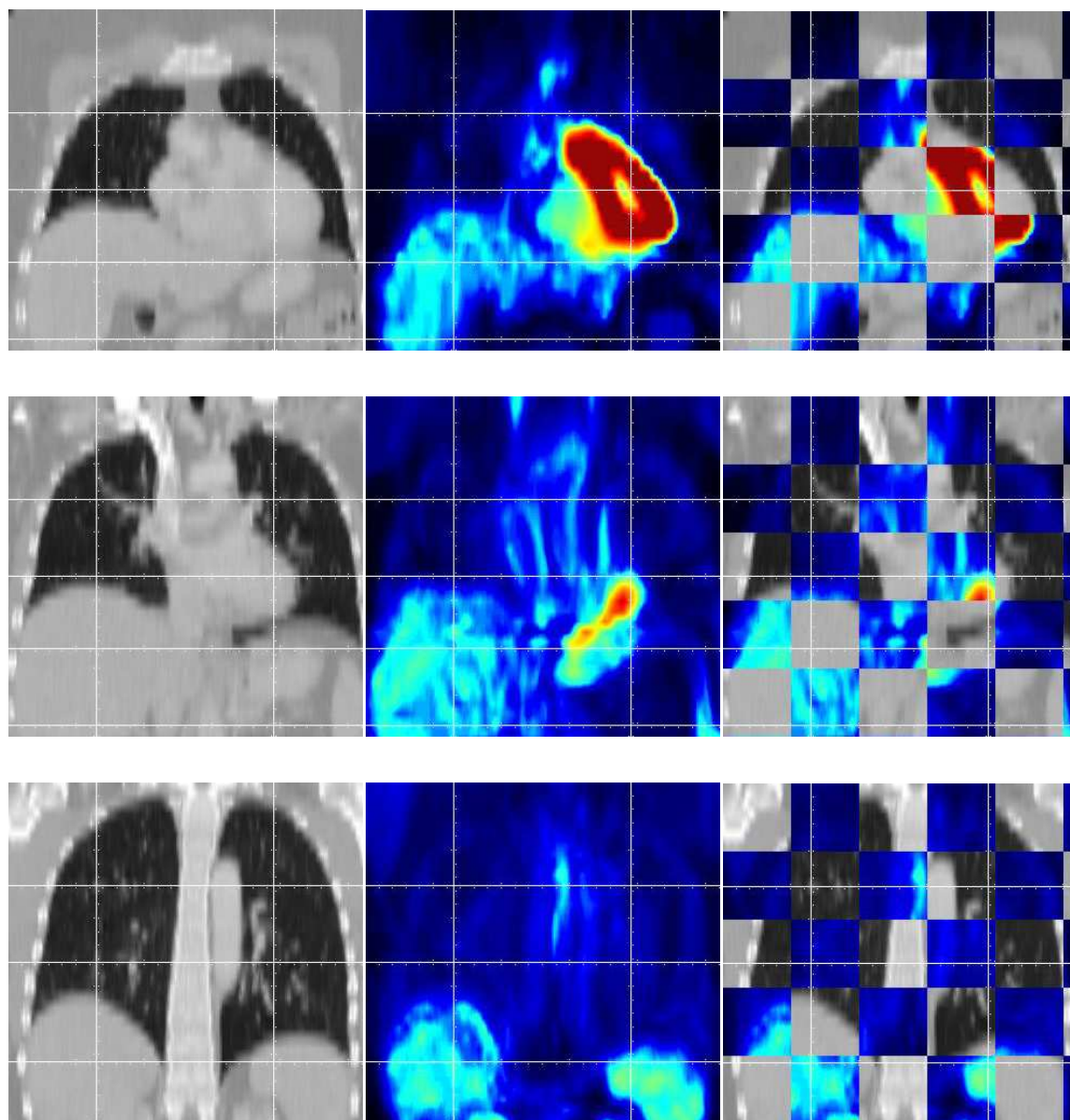
**Figura 5.6:** Paciente 2: Resultados del registro (cortes axiales).



**Figura 5.7:** Paciente 2: Resultados del registro (cortes coronales).

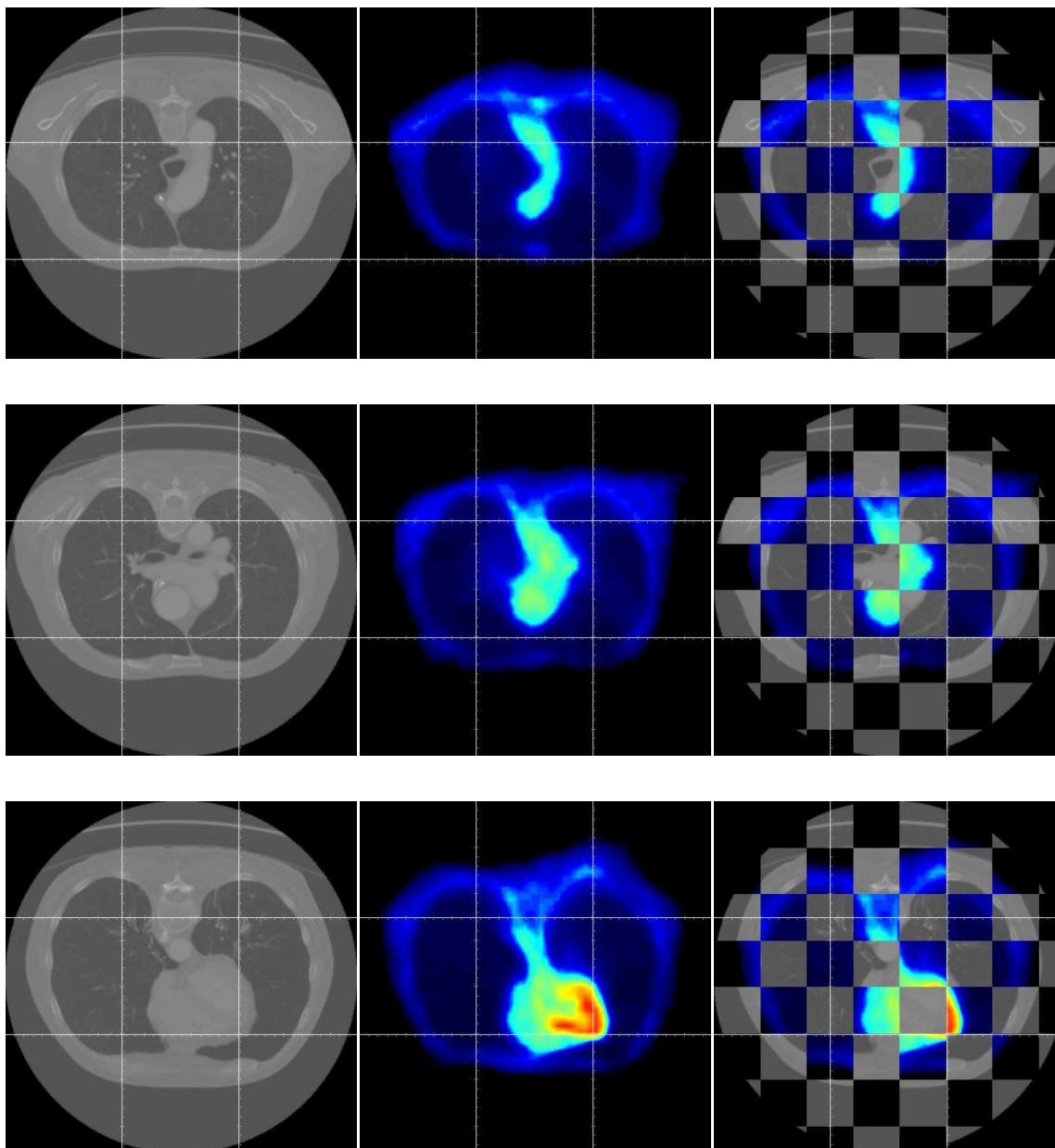


**Figura 5.8:** Paciente 3: Resultados del registro (cortes axiales).

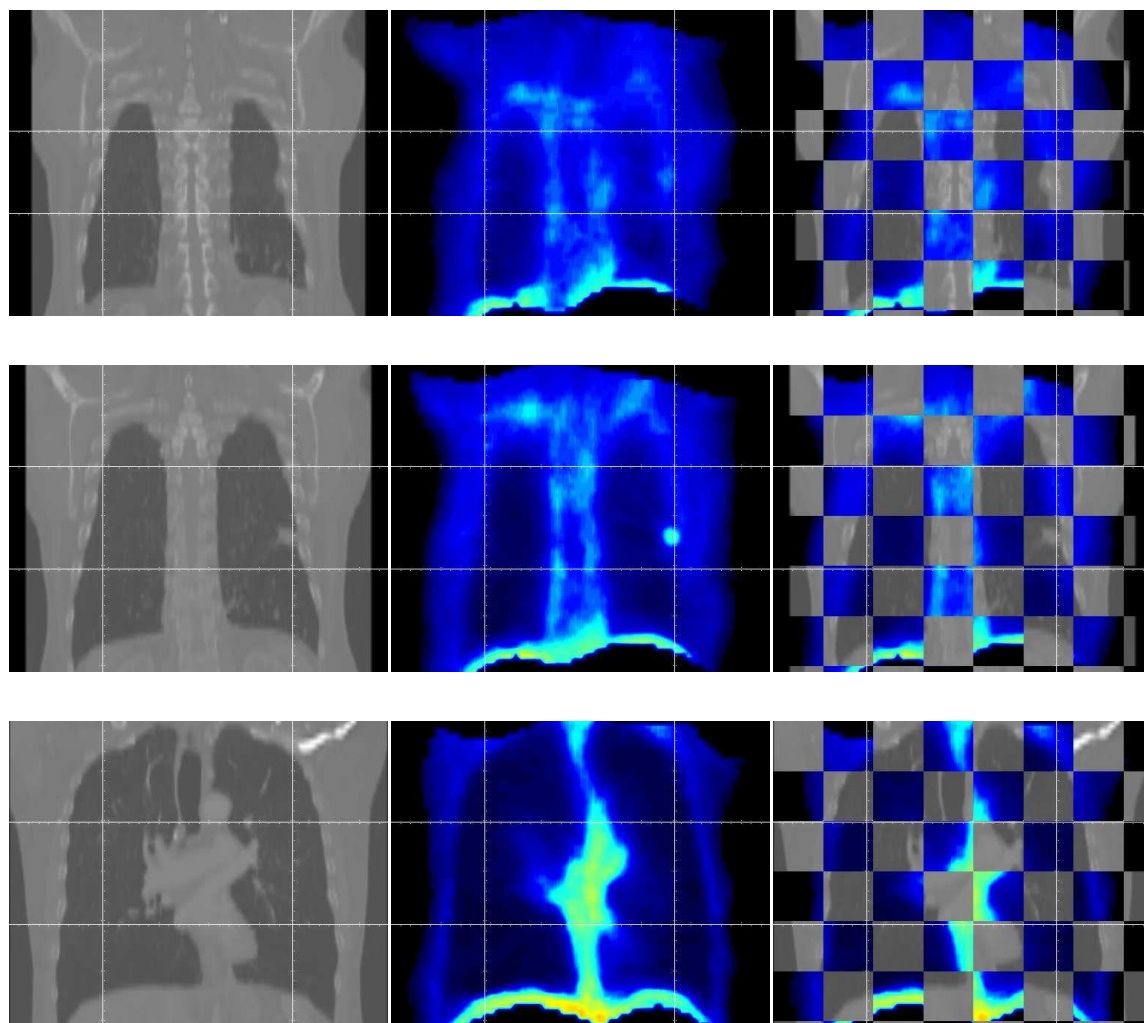


**Figura 5.9:** Paciente 3: Resultados del registro (cortes coronales).





**Figura 5.10:** Paciente 4: Resultados del registro (cortes axiales).



**Figura 5.11:** Paciente 4: Resultados del registro (cortes coronales). Nótese el nódulo presente en el segundo corte, y la fijación del marcador radiactivo en el mismo, indicando una actividad metabólica anormal.

# Capítulo 6

## Conclusiones

Habiéndose planteado el problema del registro elástico de imágenes y su interés en el campo de la diagnosis e intervención médicas, se ha descrito un método capaz de establecer automáticamente dicho registro. Se han mostrado asimismo resultados que prueban la validez de dicho método en un entorno de aplicación de reconocida complejidad, como es el registro de imágenes CT-PET de tórax y abdomen para diagnosis en oncología.

Si bien el método dista mucho de ser perfecto, y en el apartado 6.2 citamos algunas de las modificaciones que consideramos más necesarias para mejorar su rendimiento, su diseño se ha basado en una idea original que combina dos técnicas de registro hasta ahora consideradas antagónicas, y que ha permitido obtener unos resultados que superan en ciertos aspectos a los del resto de trabajos que han abordado esta aplicación.

Llegados a este punto, es el momento de revisar los objetivos cumplidos en este trabajo de tesis, y las posibles líneas de investigación que se podrían emprender a continuación.

### 6.1. Objetivos cumplidos

- Detectar una aplicación para la cual no se ha resuelto satisfactoriamente el problema del registro

No obstante el planteamiento inicial de la investigación, encaminado a desarrollar un método de registro que permitiera eliminar los elementos marcadores invasivos en una operación de cirugía robotizada, el estudio del estado del arte pronto reveló la difícil viabilidad de dicho objetivo, debido a la imposibilidad de realizar el registro en tiempo real con los medios de cálculo de que se disponía.

Se optó en este caso por redirigir el esfuerzo hacia un tema todavía poco investigado, el registro elástico de volúmenes multimodales de datos. Tras establecer contacto con el grupo formado por la ENST de París y la empresa Segami, se concretó el ámbito médico en el que se desarrollaría la investigación,



el registro elástico de imágenes PET-CT de tórax y abdomen para su uso en diagnóstico oncológica.

- Diseñar e implementar un método original de registro que se adapte a las necesidades de dicha aplicación

Una vez definido el tema central de la tesis, se revisó la bibliografía en busca de otros trabajos que trataran la misma problemática o pudieran serle aplicables. Se buscaba un método de registro elástico multimodal de tórax y/o abdomen, totalmente automático y que no precisara de elementos marcadores.

Tal como se ha descrito en la sección 3.2, son contados los métodos que trabajan sobre zonas de deformación elástica tan severas, y no se halló ninguno que reuniera las prestaciones de precisión y robustez deseadas.

Se decidió entonces diseñar un método original, que adaptara ciertas técnicas empleadas con éxito en el registro de imágenes mamográficas [27] para su uso en zonas anatómicas con severas deformaciones elásticas. La idea principal consistía en integrar las prestaciones de un registro basado en la segmentación de estructuras anatómicas con un registro basado en el contenido íntegro de las imágenes. Para conseguirlo, se llevaron a cabo las siguientes tareas:

- Desarrollo de un sistema de segmentación automática de estructuras anatómicas en la región torácica y abdominal:
    - Diseño y evaluación de una estructura jerárquica de segmentación por umbralizado.
    - Diseño y evaluación de modelos deformables 2D y 3D destinados a refinar los resultados de la segmentación por umbralizado.
  - Desarrollo y estudio comparativo de dos métodos de inicialización elástica del registro:
    - Un método basado en una extensión del algoritmo *Iterative Closest Point*, finalmente desestimado debido a su difícil integración en el sistema de registro.
    - Un método basado en *Free Form Deformations* destinadas a minimizar el error cuadrático medio entre las estructuras segmentadas. Se optó finalmente por este método debido a la facilidad, tanto conceptual como práctica, de su integración en el sistema de registro como una etapa más de la multiresolución.
  - Adaptación de un método de registro elástico multimodal, empleado con éxito en estudios mamográficos, para su uso en diagnóstico oncológica de tórax y abdomen.
- Realizar la evaluación del método diseñado, comparando sus prestaciones con la de aquellos métodos similares existentes

Una vez definido el nuevo método de registro y completado el sistema para su uso en diagnóstico oncológica, fue preciso diseñar, de común acuerdo con el personal médico implicado en el proyecto *Oncomatching* del ministerio francés de investigación, un protocolo que permitiera la evaluación de los resultados del registro elástico (véase la sección 5).

De acuerdo con dicho protocolo se desarrolló un sistema que permitiera la generación automática de los cuestionarios de evaluación a partir de los resultados del registro, y se empleó este para validar el nuevo método de registro, gracias a la colaboración de especialistas de diversos centros hospitalarios.

## 6.2. Futuras líneas de investigación

A causa de la amplitud del tema de esta tesis, previsiblemente, no ha sido posible profundizar en algunos aspectos del mismo tanto como sería deseable. Detallamos a continuación aquellas líneas de trabajo que consideramos más interesantes, sea por la mejora que comportan en el método de registro propuesto, sea por su originalidad dentro de este campo:

- Mejorar segmentación con atlas anatómico.

En efecto, la extracción de estructuras para la primera etapa de registro puede beneficiarse en gran medida de una mayor utilización de información anatómica.

Actualmente dichos conocimientos se emplean únicamente como relaciones espaciales entre los distintos órganos, que permiten definir en cierta medida las regiones de interés. El uso de un atlas anatómico ya segmentado, como los que se han confeccionado dentro del marco del *Visible Human Project* de la *National Library of Medicine (USA)*, debidamente parametrizado para adaptarse a cada caso a medida que el proceso de segmentación jerárquica obtiene información sobre el paciente, posibilitaría una extracción más precisa y robusta tanto de los órganos que actualmente se toman como referencia (pulmones, riñones) como de otros que de momento no es posible utilizar (corazón, hígado).

- Incorporar modelo respiratorio para compensar el efecto de promediado en la imagen PET.

Mediante un modelo del movimiento respiratorio, como el *Mathematical Cardiac Torso (MCAT) Phantom* de la *University of North Carolina*, podrían corregirse de forma anatómicamente coherente las diferencias entre los volúmenes de datos debidas a dicho movimiento, y que se cuentan entre las deformaciones más severas que el registro debe compensar.

- Acelerar la convergencia del registro de las imágenes segmentadas mediante un campo de GVF.

Debido al deseo de integrar el registro de las estructuras anatómicas segmentadas dentro del mismo esquema de deformación y optimización que el registro basado en intensidades, esta primera etapa adolece de una cierta lentitud.

Si bien no incluido en esta tesis por estar aún en fase de pruebas, se han obtenido resultados prometedores guiando la optimización de la FFD mediante un *gradient vector flow* generado a partir de los contornos de los órganos segmentados.

- Incorporar información de elasticidad en la FFD.

Otro modo de conseguir deformaciones anatómicamente más realistas consiste en aprovechar la segmentación que se ha hecho de las imágenes para introducir en los nodos de control de la FFD un parámetro que regule su comportamiento en función de las propiedades del tejido subyacente.

Naturalmente, esto supone que la malla de control es lo suficientemente fina para que la mayoría de nodos afecten a un único tipo de tejido.

### 6.3. Contribuciones

Todo buen trabajo de investigación científica, y en consecuencia todo trabajo de tesis, debe contener una aportación original que contribuya al avance de su campo de estudio. El presente trabajo, por bien que se trate de una tesis en ingeniería, y por consiguiente orientada a proporcionar soluciones aplicadas a un problema existente, no ha de ser una excepción.

A continuación se presenta una lista que resume las que consideramos constituyen las principales contribuciones originales de esta tesis.

#### **Integración en una estructura de deformación única de un método de registro basado en la segmentación y un método de registro basado en las intensidades.**

Este esquema, del que no se conocen antecedentes, permite sacar partido a las prestaciones de dos conceptos de registro hasta ahora antagónicos (ver clasificaciones de registro hasta la fecha [57] [1] [55]) a la vez que se compensan efectivamente sus principales inconvenientes, a saber, los errores de segmentación y de convergencia respectivamente.

En efecto, al emplear la misma estructura de deformación-optimización para ambos métodos de registro, la transición entre ellos se reduce a un cambio del criterio de similitud. Gracias a ello se pueden emplear los resultados del registro elástico de estructuras segmentadas como inicialización de la etapa basada en intensidades, de forma casi instantánea y sin pérdidas de precisión.

Por bien que inusitado, los beneficios de emplear un método de registro elástico basado en referencias anatómicas para inicializar otro basado en intensidades resultan evidentes, tanto en precisión y robustez como en tiempo de convergencia de este último. A pesar de que dicha estructura requiera la complejidad adicional de una etapa de segmentación, así como el coste computacional añadido que ésta representa, las dificultades propias del registro de las zonas torácica y abdominal justifican plenamente su elección.

El hecho de independizar totalmente la segunda etapa de registro de los resultados de las estructuras anatómicas segmentadas, más allá de la inicialización arriba mencionada, hace al sistema prácticamente inmune a las imprecisiones y errores de segmentación inherentes al trabajo con imágenes médicas. No es este el caso de la mayoría de métodos similares, que optan por incluir las referencias anatómicas como un término ponderado más en el criterio de similitud a optimizar.

#### **Diseño de una metodología de segmentación jerárquica de estructuras anatómicas de la región torácica y abdominal.**

El hecho de incluir una etapa de registro basada en estructuras anatómicas de referencia impone la tarea, nada trivial, de desarrollar un mecanismo de segmentación capaz de extraer dichas estructuras a la vez de los volúmenes CT y PET. No obstante, la particular configuración de nuestro sistema nos ha permitido explotar el hecho de

que los pequeños errores de segmentación no tienen un impacto directo en el resultado final de registro, como se ha explicado en el punto anterior.

En consecuencia, el objetivo de esta segmentación consistía en extraer el mayor número de estructuras o referencias anatómicas de forma paralela en ambos volúmenes, siendo su prioridad no tanto la precisión conseguida, sino la capacidad de analizar sus propios resultados y detectar cuando éstos no presentan las características esperadas. En este último caso, el sistema debía ser capaz de ajustarse para corregir el error o, caso de no conseguirlo, advertir al usuario y proseguir con el registro prescindiendo de la estructura en cuestión.

El sistema que se ha propuesto cumple con los requisitos citados, permitiendo una extracción razonablemente precisa y, ante todo, robusta, de distintas estructuras anatómicas presentes en los volúmenes CT y PET a registrar. El principal atractivo de este método radica en el uso de una estructura jerárquica centrada en el aprovechamiento de la información obtenida de las estructuras ya segmentadas para restringir la búsqueda de las demás.

#### **Desarrollo de un entorno de evaluación semicuantitativa del registro elástico.**

Se ha puesto a punto, bajo la dirección del *Dr. Hervé Foehrenbach*, del *Service Hospitalier Frédéric-Joliot (Paris)*, un protocolo que permite la evaluación de los resultados del registro elástico por parte de médicos especialistas.

La concepción de éste surgió como una necesidad a la hora de validar los resultados obtenidos, dada la ausencia de un *gold standard* que tomar como referencia. En efecto, la evaluación visual de una deformación elástica en tres dimensiones es sumamente compleja, requiriendo de personal familiarizado a la vez con la anatomía y la transformación geométrica implicadas. Por otra parte, el gran número de grados de libertad de la transformación hace poco viable su validación mediante elementos marcadores implantables, y la naturaleza multimodal de las imágenes dificulta el uso de referencias anatómicas con tal fin.

# Apéndice A

## Splines

Dada su repetida aparición a lo largo de este trabajo de tesis, y el importante papel que las splines juegan dentro del modelo de transformación seleccionado para nuestro sistema de registro, se ha juzgado oportuno dedicar el presente apéndice a profundizar en las distintas propiedades de estas funciones.

Para mayor información sobre este tema, remitimos al lector al artículo original de M.Unser [58], que describe de forma más detallada y exhaustiva las características de las splines en su aplicación al tratamiento de imágenes y señales.

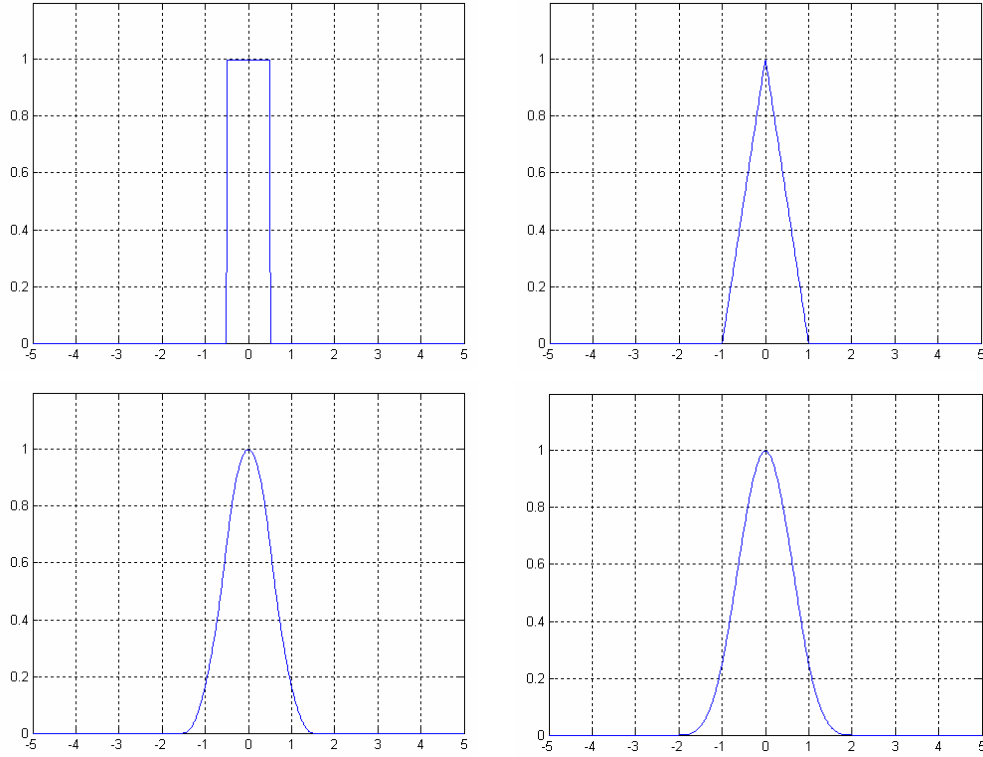
### A.1. Introducción

Las splines son polinomios a tramos conectados entre sí de forma suave. Los puntos de unión entre los polinomios se denominan *nudos*. Para un spline de grado  $n$ , cada segmento es un polinomio de grado  $n$ , cosa que parecería implicar que se precisa de  $(n+1)$  coeficientes para describir cada segmento. Sin embargo, existe una restricción de suavidad que impone la continuidad de la spline y sus derivadas hasta el orden  $(n-1)$  en los nudos, por lo que a efectos prácticos no hay más que un grado de libertad por tramo. En adelante consideraremos tan sólo splines con nudos regularmente espaciados a distancia unidad.

Un importante resultado indica que este tipo de splines se pueden caracterizar de forma única en términos de una expansión de B-splines

$$s(x) = \sum_{k \in \mathbf{Z}} c(k) \beta^n(x - k), \quad (\text{A.1})$$

que implica las traslaciones enteras de la B-spline central de grado  $n$  indicada como  $\beta^n(x)$ ; los parámetros del modelo son los coeficientes B-spline  $c(k)$ . Las B-splines, definidas más abajo, son funciones acampanadas, simétricas, construidas mediante la convolución  $(n+1)$ -ésima de un pulso rectangular  $\beta^0$ :



**Figura A.1:** B-Splines de grados 0 a 3. Nótese como equivalen a convoluciones sucesivas de la B-Spline de grado cero.

$$\beta^0(x) = \begin{cases} 1, & -\frac{1}{2} < x < \frac{1}{2} \\ \frac{1}{2}, & |x| = \frac{1}{2} \\ 0, & \text{en otro caso} \end{cases} \quad (\text{A.2})$$

$$\beta^n(x) = \underbrace{\beta^0 * \beta^0 * \dots * \beta^0(x)}_{(n+1)\text{veces}}. \quad (\text{A.3})$$

Las B-splines de grados 0 a 3 se muestran en la figura A.1. Gracias a esta representación, cada spline queda representado sin ambigüedad por su secuencia de coeficientes B-spline  $c(k)$ , que tiene la muy conveniente estructura de una señal discreta, aunque el modelo subyacente sea continuo.

Dado que el modelo de B-Spline A.1 es lineal, estudiar las propiedades de sus elementos básicos puede dar mucha información sobre los splines en general.

## A.2. B-splines

Las B-splines (donde B puede leerse como *base* o *básicas*) son los elementos constructivos básicos de las splines. Su utilidad deriva de tener soporte compacto,



siendo de hecho las splines polinomiales más cortas posibles.

Aquí consideraremos la B-spline simétrica respecto a su centro de grado  $n$ ,  $\beta_n(x)$  definida por A.2 y A.3. El modo más simple de obtener una fórmula explícita es comenzar escribiendo su transformada de Fourier

$$\hat{\beta}^n(\omega) = \left( \frac{\sin(\omega/2)}{(\omega/2)} \right)^{n+1} = \frac{(e^{j\omega/2} - e^{-j\omega/2})^{n+1}}{(j\omega)^{n+1}}, \quad (\text{A.4})$$

donde hemos expresado la convolución  $(n+1)$ -ésima que veíamos en A.3 como un producto en el dominio frecuencial. Consideremos ahora la función potencia para el semieje positivo

$$(x)_+^n = \begin{cases} x^n, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad (\text{A.5})$$

cuya transformada de Fourier denotaremos  $X_+^n(\omega)$ . Para evitar la evaluación explícita de  $X_+^n(\omega)$ , que implica utilizar deltas de Dirac, derivamos  $(x)_+^n$  repetidamente hasta que encontramos la discontinuidad de orden  $(n+1)$  en el origen:  $D^{n+1}(x)_+^n = n!\delta(x)$ . En el espacio de Fourier, esto resulta  $(j\omega)^{n+1}X_+^n = n!$ . Usando esta identidad manipulamos A.4 de la manera siguiente

$$\begin{aligned} \hat{\beta}^n(\omega) &= \frac{(e^{j\omega/2} - e^{-j\omega/2})^{n+1}}{(j\omega)^{n+1}} \underbrace{(j\omega)^{n+1}X_+^n(\omega)}_{=1} \\ &= \frac{1}{n!}(e^{j\omega/2} - e^{-j\omega/2})^{n+1}X_+^n(\omega), \end{aligned}$$

A continuación expandimos el termino entre paréntesis utilizando el teorema binomial, con lo que queda

$$\hat{\beta}^n(\omega) = \frac{1}{n!} \sum_{k=0}^{n+1} \binom{n+1}{k} (-1)^k e^{-j\omega[k-(n+1)/2]} X_+^n(\omega). \quad (\text{A.6})$$

Finalmente, interpretamos las exponenciales complejas como términos de fase (traslaciones temporales), y obtenemos su formula correspondiente en el dominio temporal

$$\beta^n(x) = \frac{1}{n!} \sum_{k=0}^{n+1} \binom{n+1}{k} (-1)^k \left( x - k + \frac{n+1}{2} \right)_+^n. \quad (\text{A.7})$$

Este resultado demuestra claramente que  $\beta^n(x)$  es un polinomio a tramos de grado  $n$ . Implica también que la derivada  $(n+1)$ -ésima de  $\beta^n(x)$  es una serie de impulsos de Dirac, indicando que  $\beta^n(x)$  es derivable hasta orden  $n$ . Para  $n$  impar, los nudos (o puntos de discontinuidad en la derivada  $n$ -ésima) están en los enteros, mientras que para  $n$  par están en los medios enteros.

La manipulación de las B-splines resulta sumamente simple. Por ejemplo, podemos obtener derivadas a través de la siguiente formula

$$\frac{d\beta^n(x)}{dx} = \beta^{n-1}\left(x + \frac{1}{2}\right) - \beta^{n-1}\left(x - \frac{1}{2}\right), \quad (\text{A.8})$$

que reduce el grado en uno. Del mismo modo, se calcula la integral como

$$\int_{-\infty}^x \beta^n(x) dx = \sum_{k=0}^{+\infty} \beta^{n+1}\left(x - \frac{1}{2} - k\right). \quad (\text{A.9})$$

Una vez se conoce el efecto de operadores lineales tales que A.8 y A.9 en las funciones de base, resulta trivial aplicarlos a cualquier spline vía la representación B-spline.

Dentro de la familia de splines polinomiales, las splines cubicas tienden a ser las más populares en la práctica, tal vez debido a su propiedad de mínima curvatura, que discutimos más adelante. Utilizando A.2, obtenemos la siguiente representación cerrada de la B-spline cubica:

$$\beta^3(x) = \begin{cases} \frac{2}{3} - |x|^2 + \frac{|x|^3}{2}, & 0 \leq |x| < 1 \\ \frac{(2-|x|)^3}{6}, & 1 \leq |x| < 2 \\ 0, & 2 \leq |x| \end{cases}, \quad (\text{A.10})$$

utilizada muy a menudo en el tratamiento de señales y de imágenes para realizar interpolaciones de buena calidad.

### A.3. Interpolación B-spline por filtrado digital

Por todo lo dicho hasta este punto, se aprecia que la principal tarea consiste en determinar el modelo B-spline de la señal de entrada  $s(k)$ . Nos centraremos ahora en el problema de la interpolación spline, en la cual que los coeficientes se determinan para que la función pase exactamente por la posición de los datos.

Para splines de grado 0 (constantes a tramos) y splines de grado 1 (lineales a tramos) este es una tarea trivial porque los coeficientes B-spline son idénticos a las muestras de la señal:  $c(k) = s(k)$ . Para splines de grados superiores, sin embargo, el problema es más complejo.

Tradicionalmente el problema de interpolación B-spline se trata mediante una operación matricial. A principios de los 90, se apreció que este problema podía resolverse también mediante técnicas más sencillas de filtrado digital.

Para deducir este algoritmo, debemos introducir el núcleo B-spline discreto  $b_m^n$ , que se obtiene muestreando la B-spline del grado  $n$  expandida por un factor  $m$ :

$$b_m^n(k) = \beta^n(x/m)|_{x=k} \quad \xleftrightarrow{z} \quad B_m^n(z) = \sum_{k \in \mathbf{Z}} b_m^n(k) z^{-k}. \quad (\text{A.11})$$

Así pues, dadas las muestras  $s(k)$  de la señal, queremos determinar los coeficientes  $c(k)$  del modelo B-spline para conseguir una coincidencia exacta en los enteros, es decir,  $\forall k \in \mathbf{Z}$

$$\sum_{l \in \mathbf{Z}} c(l) \beta^n(x-l)|_{x=k} = s(k). \quad (\text{A.12})$$

Utilizando las B-splines discretas tal como las hemos definido arriba, esta restricción puede reescribirse como una convolución

$$s(k) = (b_1^n * c)(k). \quad (\text{A.13})$$

Si definimos el operador convolución inversa tal que

$$(b_1^n)^{-1}(k) \xleftrightarrow{z} 1/B_1^n(z), \quad (\text{A.14})$$

la solución se obtiene por filtrado inverso

$$c(k) = (b_1^n)^{-1} * s(k). \quad (\text{A.15})$$

Dado que  $b_1^n$  es FIR (respuesta impulsional finita) simétrico, el denominado filtro B-spline directo  $(b_1^n)^{-1}$  es un sistema todo polos que puede implementarse muy eficientemente mediante una cascada de filtros recursivos de primer orden causal y anticausal. Este algoritmo es numéricamente estable y resulta más rápido y fácil de implementar que cualquier otra técnica numérica.

Por otra parte, puede demostrarse desde el punto de vista de la teoría de muestreo que el error en la aproximación de señales mediante coeficientes spline tiene un decaimiento  $O(T^L)$ , siendo  $T$  el periodo de muestreo y  $L = n + 1$  el orden del spline más uno, tendiendo para  $n \rightarrow \infty$  al comportamiento de un filtro paso-bajo ideal.

## A.4. Ejemplo: Aplicación a la transformación geométrica de imágenes

Se pueden extender fácilmente los splines a dimensiones superiores mediante el producto tensorial de funciones de base. En particular el modelo spline para una cierta posición  $(x, y)$  de la imagen viene dado por

$$f(x, y) = \sum_{k=k_1}^{(k_1+K-1)} \sum_{l=l_1}^{(l_1+K-1)} c(k, l) \beta^n(x-k) \beta^n(y-l), \quad (\text{A.16})$$

con

$$k_1 = k_1(x) = \left\lceil x - \frac{n+1}{2} \right\rceil \quad (\text{A.17})$$

$$l_1 = l_1(y) = \left\lceil y - \frac{n+1}{2} \right\rceil \quad (\text{A.18})$$

y

$$K = \text{soporte } \{\beta^n\} = n + 1. \quad (\text{A.19})$$

Nótese que se ha restringido el rango de la adición a aquellas contribuciones que son distintas de cero para esta localización  $(x, y)$  en particular.

Supongamos ahora que se quiere realizar una transformación geométrica de la imagen de acuerdo a un mapeo  $(x, y) = G(u, v)$ , donde  $(u, v)$  son las nuevas coordenadas de destino en la imagen. Por ejemplo, el mapeo podría consistir en una rotación, en cuyo caso el operador  $G$  sería descrito por una matriz ortogonal  $2 \times 2$ . El algoritmo de transformación B-spline sería como sigue:

En primer lugar, se precálculan los coeficientes B-spline  $c(k, l)$  mediante filtrado separable de los valores de los pixels  $f(k, l)$ . Es decir, se aplica el algoritmo de filtrado monodimensional sucesivamente a lo largo de las filas y columnas de la imagen.

Hecho esto, se recorren los puntos de la imagen transformada. Para cada posición  $(u, v)$  de la imagen de destino, se determina su posición correspondiente  $(x, y)$  en la imagen origen y se calcula su valor de acuerdo con A.16.

Este proceso típicamente requiere el cálculo de  $2(n+1)$  valores de las funciones de base más  $2(n+1)^2$  multiplicaciones por punto. Nótese que el coste de la etapa de prefiltrado es negligible en comparación. Es posible usar ordenes tan elevados como se desee, pero normalmente no hay muchos beneficios a partir de las splines cúbicas.

## A.5. Otras propiedades óptimas

### A.5.1. Propiedades variacionales

Las splines tienen ciertas propiedades extremas que las hacen interesantes. Un resultado importante es la *primera relación integral*, que afirma que para cualquier función  $f(x)$  cuya derivada  $m$ -ésima es integrable, tenemos

$$\int_{-\infty}^{+\infty} (f^{(m)})^2 dx = \int_{-\infty}^{+\infty} (s^{(m)})^2 dx + \int_{-\infty}^{+\infty} (f^{(m)} - s^{(m)})^2 dx, \quad (\text{A.20})$$

Donde  $s(x)$  es el interpolador spline de grado  $n = 2m - 1$  tal que  $s(k) = f(k)$ .

En concreto, si se aplica esta descomposición al problema de la interpolación de una secuencia de datos  $f(k)$ , se constata que de todos los posibles interpoladores  $f(x)$ , el interpolador spline  $s(x)$  es el único que minimiza la norma de la derivada  $m$ -ésima, resultado muy significativo. Ello se debe a que el segundo término de A.20 es distinto de cero en los puntos no enteros para  $f(x) \neq s(x)$ .

El spline es pues el interpolador que oscila menos. Para  $m = 2$  la integral A.20 es una aproximación a la curvatura integral de la curva  $y = f(x)$ . Así pues, el interpolador spline cúbico posee una *propiedad de mínima curvatura*, un comportamiento similar al de un cuerpo elástico.

### A.5.2. Suavizado mediante splines

La interpolación no es el único modo de ajustar un modelo continuo a una señal. En ocasiones, por ejemplo cuando los datos son ruidosos, un ajuste exacto puede no ser deseable. Tales situaciones pueden tratarse relajando las restricciones de la interpolación y utilizando cuanta información a priori del problema se disponga. La extensión natural del anterior problema de interpolación consiste en hallar la función  $s(x)$  que minimiza

$$\sum_{k \in \mathbf{Z}} (f(k) - s(k))^2 + \lambda \int_{-\infty}^{+\infty} (s^{(m)}(x))^2 dx. \quad (\text{A.21})$$

Este es un problema de mínimos cuadrados, regularizado, bien planteado (*well-posed*), donde el primer término cuantifica el error entre el modelo  $s(x)$  y los datos  $f(k)$ ; el segundo término impone una restricción de suavidad a la solución.

La elección del valor del factor de regularización  $\lambda$  refleja nuestra información a priori; puede basarse tanto en el conocimiento de la varianza del ruido como en una medida de la suavidad de la señal original.

Se puede demostrar que entre todas las funciones posibles la solución óptima es una spline de grado  $n = 2m - 1$ . Esto se puede entender a la luz de la anteriormente descrita propiedad de mínima curvatura; así, el resultado obtenido mediante cualquier función distinta a una spline puede mejorarse sustituyéndolo por su interpolación spline, que reducirá el segundo término del criterio a la vez que conserva los mismos valores  $s(k)$  en los puntos de muestreo.

La solución al anterior problema se denomina una *spline de suavizado*, ya que es equivalente a un cierto tipo de suavizado de los datos. De forma similar a la interpolación exacta que corresponde al caso  $\lambda \rightarrow 0$ , los coeficientes B-spline del spline de suavizado pueden calcularse eficientemente mediante filtrado recursivo.

### A.5.3. Regularidad máxima y mínimo soporte

Es un hecho bien conocido de la teoría de wavelets que las B-splines son las funciones de escalado más cortas de orden de aproximación  $L$  (esto es, cuyo error de aproximación decae como la  $L$ -ésima de potencia del periodo de muestreo). Son también las más regulares: su regularidad de Sobolev ( $r$  derivadas en  $L_2$ ) es  $r_{\max} = n + \frac{1}{2}$  y su exponente de Hölder es  $\alpha = n$ . Esto último significa que la B-spline de grado  $n$  tiene derivada continua casi  $n$  veces; para ser estrictos, la derivada  $n$ -ésima de una spline de grado  $n$  tiene algunas discontinuidades aisladas (los nudos), pero sin embargo es acotada.

Así pues, en el sentido más general, las B-splines son las funciones más cortas y suaves de orden  $L$ . Dado que el rendimiento de un algoritmo de aproximación está estrechamente relacionado con el orden de aproximación y hasta cierto punto la regularidad de las funciones de base, esto tiene consecuencias importantes, especialmente en la interpolación de imágenes. En este tipo de aplicaciones, donde el coste computacional viene dado principalmente por el tamaño de las funciones de base, tiene

mucho sentido utilizar aquellas más cortas posible entre las que cumplan los requisitos de orden, esto es, las B-Splines.

#### A.5.4. Splines fraccionarias

Un hecho interesante es que las B-splines pueden generalizarse a ordenes fraccionarios. Las splines fraccionarias son funciones potencia a tramos cuyos elementos constructivos tienen la forma  $(x-x_k)_+^\alpha$ , con  $\alpha > -\frac{1}{2}$  real. Las B-splines correspondientes proporcionan una transición uniforme entre las polinómicas. Conservan todas las propiedades de las B-splines convencionales -basta sustituir  $n$  por  $\alpha$  en todas las fórmulas-, excepto el soporte compacto. Una justificación del interés de las B-splines fraccionarias es que ofrecen la misma facilidad conceptual para tratar las derivadas fraccionarias que las splines convencionales ofrecían para las derivadas.

### A.6. Aplicaciones

Entre las principales aplicaciones de las splines, cabe mencionar:

- visualización de imágenes y *zooming*
- realización de transformaciones geométricas (traslaciones no enteras, rotaciones, reescalados y transformaciones afines tanto de imágenes como de volúmenes)
- implementación de filtros reescalables y transformación wavelet rápida
- compresión de imágenes
- procesado multiescala
- detección de contornos
- implementación de modelos deformables
- conversión analógico-digital

### A.7. Conclusión

Resumimos a continuación las principales ventajas de las splines:

- Siempre se puede obtener una representación continua de una señal discreta, en una o más dimensiones, aproximándola mediante una spline. Dicha aproximación puede ser exacta (interpolación) o aproximada (mínimos cuadrados o spline de suavizado). La aproximación mediante splines es en general preferible a otras formas de aproximación dado que tiene una menor tendencia a oscilar (propiedad de mínima curvatura).

- Las splines polinómicas pueden expresarse como combinación lineal de funciones de base B-spline. Para el caso de nudos regularmente espaciados, los coeficientes B-spline pueden determinarse mediante un simple filtrado digital, sin necesidad de manipular matrices.
- La razón principal para trabajar con B-splines es que tienen un soporte compacto. Son las funciones más cortas con un orden de aproximación  $L = n + 1$ . Esta propiedad es esencial para la eficiencia computacional. Su forma analítica simple facilita asimismo las manipulaciones.
- Las splines son funciones con un comportamiento suave (polinomios a tramos). Las splines de grado  $n$  son  $(n - 1)$  veces continuamente diferenciables. En consecuencia, las splines tienen propiedades de aproximación excelentes. Se dispone de tasas de convergencia y estimaciones de error precisas.
- Las splines tienen propiedades multiresolución que las vuelven muy apropiadas para la construcción de bases wavelet y para realizar procesamiento multiresolución.
- Las B-splines y sus contrapartidas wavelet tienen propiedades de localización excelentes, siendo adecuadas para análisis de señales tiempo-frecuencia.
- La familia de las splines polinómicas ofrece flexibilidad en el diseño. Incrementando el grado  $n$  se conmuta progresivamente de las más simples representaciones constante a tramos ( $n = 0$ ) y lineal a tramos ( $n = 1$ ) al otro extremo, que corresponde al modelo de señal limitado en frecuencia ( $n \rightarrow \infty$ ).
- El procedimiento de muestreo convencional puede modificarse fácilmente para obtener una representación spline de una señal analógica. Esto se reduce esencialmente a sustituir el filtro paso-bajo ideal de Shannon por otro prefiltro óptimo especificado por la representación. En principio no hay razón para preferir el modelo limitado en frecuencia -y su correspondiente interpolador sinc- a los otros.

Finalmente, destacar que existen técnicas similares para el caso de datos no uniformemente espaciados. El precio a pagar es, sin embargo, la pérdida de la conveniente estructura invariante al desplazamiento (filtros) que antes hemos enfatizado.





# Apéndice B

## Interpolación

En este apéndice se examinan brevemente las diversas técnicas de interpolación de datos, dado el importante papel que éstas desempeñan en todo sistema de registro de imágenes.

En efecto, a lo largo de este trabajo de tesis ha sido preciso considerar los métodos de interpolación existentes a la hora de realizar tareas como la transformación de la imagen flotante, el cálculo del histograma conjunto o, implícitamente, la creación de mapas de deformación mediante FFD.

Por esta razón se ha estimado oportuno incorporar una breve referencia sobre las técnicas de interpolación más comunes en el tratamiento de imágenes. Para un estudio más detallado nos remitimos al artículo de Lehmann [59], de donde se ha extraído la mayor parte de la información que se expone a continuación.

### B.1. Introducción

Son muchas las aplicaciones de la interpolación en el tratamiento de imágenes. Se trata de una etapa esencial en el remuestreo (*resampling*) ya que permite transformar una matriz discreta en una imagen continua, cuyo muestreo permite obtener la imagen remuestreada.

Dicho remuestreo es necesario en numerosas operaciones de tratamiento de imágenes, tanto en alineamiento geométrico y puesta en correspondencia como para mejorar la calidad de las imágenes en dispositivos de visualización, en compresión de imágenes, generación de volúmenes tomográficos mediante retroproyección y manipulación de imágenes para diagnóstico, diseño asistido por ordenador y cirugía asistida por ordenador.

Existen numerosos métodos destinados a la interpolación de imágenes. Inicialmente se recurrió a métodos simples como la interpolación lineal o por vecino más próximo. A raíz de la teoría de la información introducida por Shannon a finales de los años 40, se adoptó la función *sinc* como función de interpolación preferente. Sin embargo, este interpolador ideal presenta una respuesta impulsional infinita (IIR), poco conveniente para interpolación local con respuesta impulsional finita (FIR). Desde el punto de vista matemático, se han propuesto aproximaciones de la función *sinc* mediante polinomios

de Taylor o Lagrange. Posteriormente, dada su eficiencia computacional, distintas familias de splines los han ido sustituyendo.

En la literatura se pueden hallar gran variedad de métodos de confusa nomenclatura. Las B-splines se denominan a veces splines cúbicas, mientras que la interpolación cúbica se conoce también como convolución cúbica, interpolación spline de alta resolución e interpolación spline bicúbica. Para evitar confusiones en adelante se hará una distinción entre lo que denominaremos aproximación B-spline y la interpolación B-spline. Remitimos al lector al apéndice A para una explicación más detallada de estas técnicas.

A continuación se presenta una revisión de algunos de los métodos de interpolación más populares, junto con datos comparativos de su rendimiento en aplicaciones de procesamiento de imágenes médicas. Con esto se pretende justificar las elecciones realizadas en el desarrollo de este trabajo, así como ofrecer al lector una referencia rápida a la hora de seleccionar uno u otro método para su propia aplicación.

## B.2. Métodos de interpolación

Para el remuestreo de la imagen, la interpolación debe reconstruir una señal continua bidimensional  $s(x, y)$  a partir de sus muestras discretas  $s(k, l)$  con  $s, x, y \in \mathbf{R}$  y  $k, l \in \mathbf{N}$ . Así pues, la amplitud en la posición  $(x, y)$  debe estimarse a partir de sus vecinos discretos. Esto puede ser descrito formalmente como la convolución de las muestras discretas de la imagen con la respuesta impulsional continua  ${}_{2D}h(x, y)$  de un filtro reconstructor bidimensional

$$s(x, y) = \sum_k \sum_l s(k, l) \cdot {}_{2D}h(x - k, y - l). \quad (\text{B.1})$$

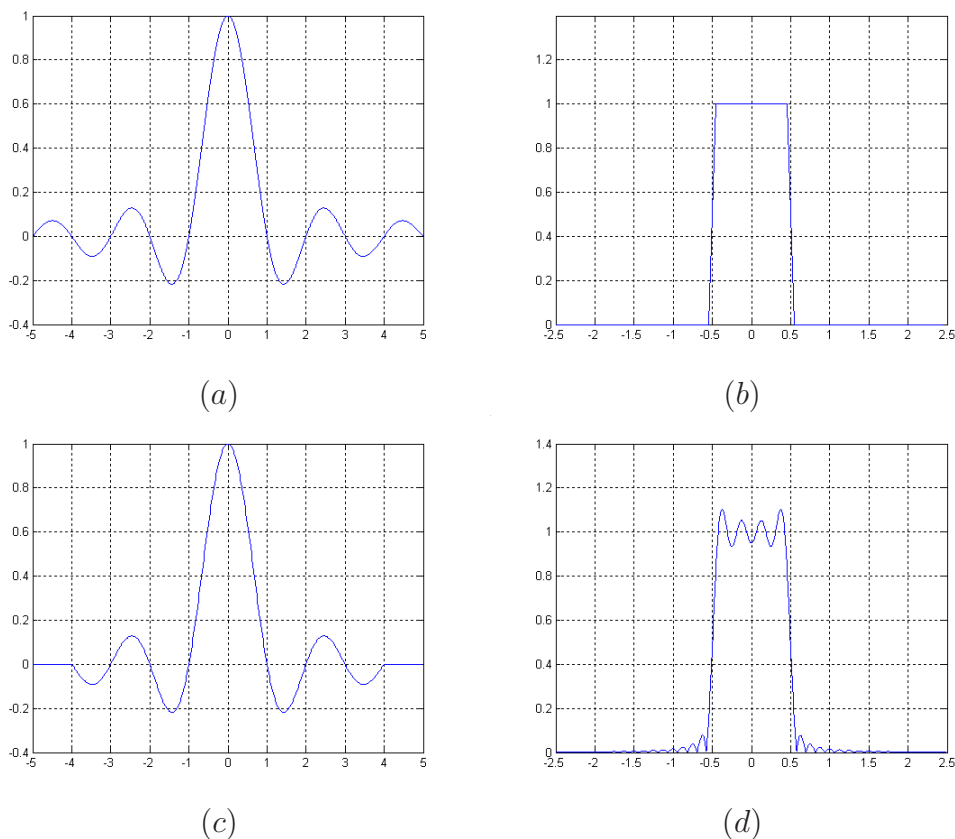
Normalmente, se usan *kernels* de interpolación simétricos y separables para reducir el coste computacional

$${}_{2D}h(x, y) = h(x) \cdot h(y). \quad (\text{B.2})$$

### Interpolación ideal

Según la teoría del muestreo, muestrear una imagen continua  $s(x, y)$  genera en el espacio Fourier infinitas repeticiones de su espectro continuo  $S(u, v)$ , que no se solapan si se satisface el criterio de Nyquist. En tal caso y sólo entonces la imagen original  $s(x, y)$  puede ser perfectamente reconstruida a partir de sus muestras  $s(k, l)$  mediante multiplicación en el dominio de Fourier por un pulso rectangular apropiado.

La interpolación ideal 1D equivale a la multiplicación con una función *rect* en el dominio de Fourier, y puede realizarse en el dominio espacial mediante convolución con la función *sinc*



**Figura B.1:** Interpolador ideal (a) y módulo de su transformada de Fourier (b). Se aprecia el efecto que un truncado de dicho interpolador (c) tiene sobre su espectro frecuencial (d).

$$Ideal h(x) = \frac{\sin(\pi x)}{\pi x} = sinc(x). \quad (B.3)$$

De esta función ideal de interpolación se pueden derivar algunas propiedades fundamentales de todo interpolador:

$^{Ideal}h(x)$  es positivo de cero a uno, negativo de uno a dos, positivo de dos a tres y así sucesivamente. Para  $h(0) \equiv 1$  estos cruces por cero garantizan que la imagen no se modifica si se remuestrea en la misma malla. Así pues, los kernels que satisfacen

$$\begin{cases} h(0) \equiv 1 \\ h(x) \equiv 0, \quad |x| = 1, 2, \dots \end{cases} \quad (B.4)$$

evitan el suavizado y preservan las altas frecuencias. Se denominan interpoladores. Veremos más adelante que las funciones de kernel más adecuadas tienden a tener esta forma general. Al contrario que los interpoladores, los kernels que no cumplen B.4 se denominan aproximadores.

Muestrear la imagen interpolada (continua) es equivalente a interpolar la imagen discreta con una función de interpolación muestreada. El muestreo de la función de interpolación crea alias de sus altas frecuencias sobre las bajas. Tan sólo en el caso de la interpolación ideal no existen las altas frecuencias y por tanto la función de interpolación muestreada tiene el mismo espectro de Fourier en el intervalo  $-0,5 < f < 0,5$  que la misma función sin muestrear.

Sin embargo, es necesario examinar no sólo la función de interpolación continua  $h(x)$ , sino también muestreos típicos de la función de interpolación  $h(k)$ . En particular, la suma de todas las muestras debería ser uno para todo desplazamiento  $0 \leq d < 1$

$$\sum_{k=-\infty}^{\infty} h(d+k) \equiv 1. \quad (\text{B.5})$$

Esto significa que para cualquier desplazamiento  $d$  la amplificación de la componente continua es la unidad, por lo que el remuestreo no altera la energía de la imagen. En otras palabras, el brillo de la imagen no cambia si ésta es muestreada o interpolada. Por ello, las funciones de kernel se denominan DC-constantes o DC-inconstantes dependiendo de si cumplen o no la condición B.4. Más adelante veremos que los mejores kernels son DC-constantes.

La ecuación B.4 se llama también condición de partición de la unidad, y puede ser fácilmente evaluada en el dominio de Fourier. En efecto, según la teoría de la información la suma de muestras discretas de el kernel  $h(x)$  equivale al área bajo la función continua que se obtiene multiplicando  $h(x)$  por un tren de funciones delta  $\delta(x)$

$$\sum_{k=-\infty}^{\infty} h(d+k) = \int_{-\infty}^{\infty} h(x) \cdot \Pi(x+d) dx \quad (\text{B.6})$$

donde

$$\Pi(x) = \sum_{k=-\infty}^{\infty} \delta(x+k) \quad (\text{B.7})$$

y, por definición, el peso de un impulso delta corresponde a la amplitud del kernel  $h$  en la posición de la  $\delta$ . Si añadimos al integrando en B.6 la exponencial  $e^{-i2\pi fx}$ , cuyo valor es 1 para  $f = 0$ , podemos reinterpretar la integral como una transformada de Fourier. Así, de las ecuaciones B.4 y B.6 obtenemos

$$H(f) * \Pi(f) \cdot e^{i2\pi fd}|_{f=0} \equiv 1 \quad \Leftarrow \quad \begin{cases} H(0) \equiv 1 \\ H(f) \equiv 0, \quad |f| = 1, 2, \dots \end{cases} \quad (\text{B.8})$$

donde  $H(f)$  denota la transformada de Fourier de  $h(k)$ .

Debido a que las condiciones en B.8 son necesarias pero no suficientes en el contexto de la interpolación, se utilizan para distinguir en el dominio de Fourier los kernels DC-constantes de los DC-inconstantes.

### Interpolación sinc

Si bien la función sinc permite una reconstrucción exacta de  $s(x, y)$ , tiene el problema de ser ilimitada en el tiempo. Este problema puede resolverse bien truncando, bien multiplicando por una ventana  $w(x)$  de modo que

$$\text{sinc}h(x) = \begin{cases} \text{sinc}h(x) \cdot w(x), & 0 \leq |x| < N/2 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.9})$$

donde  $N$  indica el número de muestras de la función de kernel finita.

Dado que el truncado en el tiempo equivale en dominio frecuencial a una convolución del interpolador ideal con una señal sinc, ello produce un efecto de *ringing*, también conocido como fenómeno de Gibbs. Asimismo ello conlleva una violación del principio de partición de la unidad, por lo que los kernels truncados serán DC-inconstantes. El uso de ventanas  $w(x)$  menos severas que el pulso rectangular (coseno, Hanning, Kaiser-Bessel, Blackman-Harris, ...) puede contribuir a atenuar estos efectos.

### Interpolación por vecino más próximo

Esta es la forma más simple de aproximar la función sinc mediante un kernel finito. El valor interpolado  $s(x)$  se determina tomando únicamente el valor conocido  $s(k)$  más próximo, lo que es lo mismo que convolucionar con una señal rectangular.

$$h_1(x) = \begin{cases} 1, & 0 \leq |x| < 0,5 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.10})$$

Se trata, naturalmente, de un interpolador DC-constante. Por otra parte, esta técnica conlleva importantes efectos de aliasing y difuminado.

### Interpolación lineal

En este caso se genera el valor interpolado  $s(x)$  mediante una ponderación de los vecinos inmediatos por el inverso de su distancia a la posición  $x$ . Así, el interpolador aproxima la señal sinc mediante una función triangular:

$$h_2(x) = \begin{cases} 1 - |x|, & 0 \leq |x| < 1 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.11})$$

Se trata también de un interpolador DC-constante, cuyos principales problemas son la atenuación de los componentes de alta frecuencia y el aliasing, si bien en menor medida que para el caso de la interpolación por vecino más próximo.

### Aproximación cuadrática

Un método muy común para la creación de interpoladores de tipo sinc es su aproximación mediante polinomios. Si bien las funciones cuadráticas han sido poco usadas en general debido a que suelen introducir distorsión de fase, recientemente se

ha demostrado que esto no es siempre cierto, y se ha propuesto alguna familia de funciones con un comportamiento más adecuado.

Estas funciones se basan en una vecindad simétrica de  $3 \times 3$ , desde  $-3/2$  hasta  $3/2$ , con puntos de unión en las fracciones. Desde un punto de vista de kernels separados, en la interpolación se emplean los tres puntos más próximos, siendo estos los dos vecinos inmediatos y un vecino segundo que varía en cada caso.

Un kernel cuadrático simétrico puede escribirse de la siguiente forma:

$${}^{Quad}h_3(x) = \begin{cases} A_1|x|^2 + B_1|x| + C_1, & 0 \leq |x| < 1/2 \\ A_2|x|^2 + B_2|x| + C_2, & 1/2 \leq |x| < 3/2 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.12})$$

donde  $A_i, B_i$  y  $C_i \in \mathbb{R}$  y  $B_1 = 0$  para garantizar la continuidad C1. A las anteriores ecuaciones se añaden un conjunto de restricciones tales como que se satisfaga la ecuación B.5, y que los polinomios encajen perfectamente en sus puntos inicial y final así como en las uniones.

De este modo, se reduce dicha aproximación cuadrática a la siguiente expresión con un único grado de libertad  $a \in \mathbb{R}$ .

$${}^{Quad}h_3(x) = \begin{cases} -2a|x|^2 + 1/2(a+1), & 0 \leq |x| < 1/2 \\ a|x|^2 - (2a+1/2)|x| + 3/4(a+1), & 1/2 \leq |x| < 3/2 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.13})$$

Una ventaja de esta familia de polinomios cuadráticos es que todos sus miembros tienen fase lineal en el dominio frecuencial, por lo que no introducen distorsión de fase.

Se puede forzar que las derivadas de estos polinomios encajen en los puntos de unión fijando este último grado de libertad de forma que  $a = 1/2$ . Dado que el kernel resultante no verifica la condición B.4, se le denomina aproximador cuadrático. Su principal inconveniente es el fuerte efecto de difuminado que produce en las imágenes.

### Interpolación cuadrática

Una alternativa al anterior caso es emplear el grado de libertad de la familia de funciones descrita para forzar la condición de paso por cero B.4 del interpolador. Esto se consigue cuando  $a = 1$ , obteniéndose un interpolador DC-continuo con un mejor perfil pasa bajos que en el caso anterior, pero perdiéndose la continuidad C1, con el consiguiente empeoramiento de los efectos de aliasing.

### Aproximación B-Spline

Tal como se ha descrito en el apéndice A, las splines básicas o B-Splines pueden utilizarse como interpolador. Mediante B-Splines pueden conseguirse los interpoladores correspondientes a la interpolación por vecino más próximo, lineal y a la aproximación cuadrática.



El caso más empleado es el de la B-Spline cúbica,

$$h_4(x) = \begin{cases} (1/2)|x|^3 - |x|^2 + 2/3, & 0 \leq |x| < 1 \\ -(1/6)|x|^3 + |x|^2 - 2|x| + 4/3, & 1 \leq |x| < 2 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.14})$$

Dado que este kernel no satisface los mismos cruces por cero que la función ideal sinc, se trata de un aproximador más que de un interpolador, por lo que provoca un efecto muy fuerte de difuminado. Por otra parte, tiene la ventaja de ser DC-constante y presentar un excelente rechazo de las frecuencias fuera de la banda de paso. Si bien se puede mejorar la calidad de la interpolación aumentando el orden de la B-Spline, esto conlleva un empeoramiento del difuminado.

### Interpolación B-Spline

Es posible crear un kernel interpolador mediante B-Splines, aplicando el aproximador descrito arriba a un determinado conjunto de muestras  $t(k)$  generadas a partir de los datos  $s(k)$  de modo que la curva resultante cumpla las condiciones deseadas:

$$s(k) = \sum_{m=k-2}^{k+2} t(m) \cdot h_4(k-m) \quad (\text{B.15})$$

Esto da lugar a un sistema de ecuaciones cuya solución brinda los elementos  $t(k)$ .

al contrario que en el resto de métodos de interpolación aquí descritos, este esquema hace que el kernel  $^{Spline}h(x)$  sea infinito. En el apéndice A se detalla cómo implementar este interpolador mediante filtrado recursivo.

La interpolación B-Spline cúbica ofrece unas excelentes características pasa bajos, buena cancelación de las frecuencias fuera de la banda de paso y verifica la condición de partición de la unidad, por lo que se trata de un interpolador DC-constante.

### Interpolación cúbica

Igual como se hacia para el caso de la interpolación cuadrática, pueden emplearse polinomios cúbicos para construir interpoladores que aproximen la función sinc. Dichos interpoladores pueden tener soportes de dos, cuatro, seis u ocho puntos.

En la interpolación de dos puntos se construye un kernel simétrico de la forma:

$$^{Cubic}h_2(x) = \begin{cases} A|x|^3 + B|x|^2 + C|x| + D, & 0 \leq |x| < 1 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.16})$$

con  $A, B, C$  y  $D \in \mathbb{R}$ . Si se fuerzan continuidad  $C1$  en los extremos y la condición B.4, se obtiene el kernel siguiente:

$$^{Cubic}h_2(x) = \begin{cases} 2|x|^3 - 3|x|^2 + 1, & 0 \leq |x| < 1 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.17})$$

Incrementando el soporte del interpolador cúbico, se mejora la calidad de la interpolación obtenida, obteniéndose kernels con un grado de libertad para  $N = 4$ , dos para  $N = 6$ , tres para  $N = 8$ , etc. Existen numerosos estudios que proponen ajustar los parámetros correspondientes a dichos grados de libertad para potenciar unas u otras prestaciones del interpolador, que no se describirán aquí por escapar al objetivo de este apéndice.

### Método de Mitchell y Netravali

Se trata de una familia de filtros cúbicos de  $N = 4$  puntos, en los que se han impuesto continuidad  $C1$  y la condición de partición de la unidad (B.5):

$${}^{Mitch}h_4(x) = \frac{1}{6} \cdot \begin{cases} -(9b + 6c - 12)|x|^3 + \\ + (12b + 6c - 18)|x|^2 - (2b - 6), & 0 \leq |x| < 1 \\ -(b + 6c)|x|^3 + (6b + 30c)|x|^2 - \\ -(12b + 48c)|x| + (8b + 24c), & 1 \leq |x| < 2 \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.18})$$

Ajustando los parámetros  $(b, c)$  es posible obtener kernels interpoladores o aproximadores con distintas características de difuminado, anisotropía y *ringing*, entre ellos algunos de los anteriormente descritos.

Empíricamente se ha determinado que el aproximador con parámetros  $(1/3, 1/3)$  es el que ofrece una mejor calidad de la imagen tratada.

### Interpolación de Lagrange

Es posible emplear polinomios de Lagrange para aproximar la función sinc, forzando el paso por  $N$  puntos de un polinomio de grado  $N - 1$ :

$${}^{Lagra}h_N(x) = \begin{cases} \prod_{j=0, j-\frac{N}{2}+1 \neq n}^{N-1} \frac{n-i-x}{n-i}, & n-1 \leq x < n \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.19})$$

con  $n \in \{-N/2 + 1, -N/2 + 2, \dots, N/2\}$  y  $i = j - N/2 + 1$ .

Los kernels así generados son siempre DC-constantes, y las características de su banda de paso mejoran al aumentar el orden del polinomio. Por otra parte, el hecho de no presentar continuidad  $C1$  en las uniones provoca lóbulos laterales importantes fuera de dicha banda de paso.

### Interpolación Gaussiana

Dado que la función Gaussiana es recurrente con respecto a la transformación de Fourier, es posible emplearla para crear kernels compactos tanto en el dominio temporal como en el frecuencial. Sea

$$G^0(x, \beta) = \frac{1}{\sqrt{2\pi\beta}} e^{-x^2/2\beta} \quad (\text{B.20})$$

la función Gaussiana unitaria con media cero y varianza  $\beta$ , y sea

$$G^M(x, \beta) = \frac{\partial^M}{\partial x^M} G^0(x, \beta) \quad (\text{B.21})$$

su derivada parcial M-ésima. El kernel Gaussiano de orden  $M$  se define como:

$$_{Gauss}h_N^M(x) = \begin{cases} \sum_{m=0}^M \alpha_m G^m(x, \beta_m), & 0 \leq |x| < \frac{N}{2} \\ 0, & \text{en otro caso} \end{cases} \quad (\text{B.22})$$

Los parámetros  $\alpha_m$  y  $\beta_m$  se obtienen forzando que el kernel sea igual al interpolador ideal en la posición cero del dominio espacial y frecuencial, así como que su perfil frecuencial sea lo más llano posible en el origen.

De este modo se consiguen kernels aproximadores DC-inconstantes, pero con excelentes propiedades pasa bajos en el dominio frecuencial, que mejoran con el orden  $M$  de la función.



# Apéndice C

## Publicaciones

Dentro del marco de esta tesis se han realizado un cierto numero de artículos, algunos de los cuales han sido presentados en conferencias o publicaciones internacionales especializadas en el tratamiento de imágenes médicas. A continuación se incluye una relación de aquellos que consideramos de mayor relevancia.

- "Improving thoracic elastic registration in oncology by using anatomical constraints"  
O.Camara, G.Delso, V.Frouin, I.Bloch  
Medical Imaging, Understanding and Analysis. MIUA'02
- "Evaluation of a thoracic elastic registration method using anatomical constraints in oncology"  
O.Camara, G.Delso, I.Bloch  
*2nd Joint Conference of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society. EMBS-BMES'02*
- "Computer-assisted echocardiographic diagnosis"  
Enric X. Martin, Gaspar Delso, Alicia Casals, F. Xavier Borrás  
*15th IFAC World Congress on Automatic Control. IFAC'02*
- "Elastic thoracic registration with anatomical multi-resolution."  
O.Camara, G.Delso, H.Foehrenbach, I.Bloch  
*International Journal of Pattern Recognition and Artificial Intelligence. IJPRAI*
- "Free-Form Deformations guided by Gradient Vector Flow: a surface registration method in thoracic and abdominal PET-CT applications."  
O.Camara, G.Delso, I.Bloch  
*2nd International Workshop on Biomedical Image Registration. WBIR'03*

- "3D nonlinear PET-CT image registration algorithm with constrained Free-Form Deformations."

O.Camara, O.Colliot, G.Delso, I.Bloch

*3rd IASTED International Conference on Visualization, Imaging, and Image Processing. VIIP'03*

- "Entorno de desarrollo para aplicaciones de registro de imágenes"

G.Delso, A.Casals

*Segundas Jornadas de Investigación en Ingeniería Biomédica. JIIB'00.*

- "Apport de contraintes anatomiques au recalage non linéaire d'images TDM et TEP dans les régions thoraciques et abdominales"

O.Camara, O. Colliot, G.Delso, I.Bloch

Propuesto para: *Reconnaissance de Formes et Image Analysis. RFIA'04.*

La mayor parte de los artículos aquí citados se hallan disponibles para su descarga bajo la rúbrica 'Articles' de la siguiente dirección web: <http://www.tsi.enst.fr/~delso/>.

Cabe también mencionar que ciertas de las ideas presentadas en esta tesis, concebidas dentro del marco de la colaboración con la empresa Segami SARL, han llevado al registro de una patente, pendiente de resolución, cuyos datos se facilitan a continuación:

- US Provisional Patent Appl. No. 60/369, 681

Titulada: "*Registration of thoracic and abdominal imaging modalities.*"

I. Bloch, G. Delso and O. Camara

[20] [49] [60] [61] [45] [62] [44] [57] [5] [63] [30] [64] [65] [28] [66] [34] [67] [10] [68]  
[69] [48] [33] [46] [70] [24] [39] [23] [25] [8] [9] [71] [22] [21] [19] [72] [51] [73] [74] [59] [54]  
[75] [47] [16] [76] [32] [1] [35] [37] [77] [4] [78] [79] [2] [80] [81] [82] [43] [27] [17] [31] [83]  
[84] [13] [11] [15] [36] [50] [18] [26] [38] [58] [29] [14] [12] [85] [6] [55] [56] [7] [52] [53]





# Bibliografía

- [1] J. B. A. Maintz; M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
- [2] G. P. Penney; J. Weese; J. A. Little; et al. A comparison of similarity measures for use in 2-d/3-d medical image registration. *IEEE Transactions on Medical Imaging*, 17(4):586–595, 1998.
- [3] H. Barrow; J. Tenenbaum; R. Bolles; H. Wolf. Parametric correspondence and chamfer matching: two new techniques for image matching. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 659–663, 1977.
- [4] R. Mohr; T. Pavlidis; A. Sanfeliu. *Structural Pattern Analysis*. S World Scientific Pub. Co., Teaneck, NJ., 1990.
- [5] H Bunke; A Sanfeliu. *Syntactic and structural pattern recognition: Theory and applications*. World Scientific Pub. Co., Teaneck, NJ., 1990.
- [6] J. Weese; T. M. Buzug; C. Lorenz; C. Fassnacht. An approach to 2d/3d registration of a vertebra in 2d x-ray fluoroscopies with 3d ct images. In *Proceedings of CVRMed/MRCAS*, pages 119–128, 1997.
- [7] R. P. Woods; J. C. Mazziotta; S. R. Cherry. Mri-pet registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17:536–546, 1993.
- [8] D. L. G. Hill; D. J. Hawkes; N. Harrison; C. F. Ruff. A strategy for automated multimodality registration incorporating anatomical knowledge and imager characteristics. In *Information Processing in Medical Imaging IPMI'93*, volume 46, pages 182–196, 1993.
- [9] D. L. G. Hill; D. J. Hawkes. Voxel similarity measures for automated image registration. In *Proc. Visualisation in Biomedical Computing*, pages 205–216, 1994.
- [10] A. Collignon; F. Maes; D. Delaere; D. Vandermeulen; P. Suetens; G. Marchal. Automated multimodality image registration using information theory. In *Proceedings of the International Conference on Information Processing Medical Imaging, IPMI*, pages 263–264, 1995.

- [11] C. Studholme; D. Hill; D. Hawkes. Multiresolution voxel similarity measures for mr-pet registration. In *Proceedings of the Information Processing in Medical Imaging Conference*, 1995.
- [12] M. P. Wachowiak; R. Smolikova; G. D. Tourassi; A. S. Elmaghraby. Generalized mutual information similarity metrics for multimodal biomedical image registration. In *Proceedings of second joint meeting EMBS-BMES*, 2002.
- [13] C. E. Shannon. A mathematical theory of communication. *Bell System Technology Journal*, 27:379–423, 623–656, 1948.
- [14] P. Viola. *Alignment by maximization of mutual information*. PhD thesis, MIT, Cambridge, Ma, 1995.
- [15] C. Studholme; D. Hill; D. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.
- [16] F. Maes; A. Collignon; D. Vandermeulen; G. Marchal; P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16:187–198, 1997.
- [17] D. Sarrut. *Recalage multimodal et plate-forme d'imagerie médicale à accès distant*. PhD thesis, Université Lumière Lyon 2, France, 2000.
- [18] P. Thévenaz; M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.
- [19] J. R. Hurley; R. B. Cuttall. The procrustes program: direct rotation to test a hypothesized factor structure. *Behavioral Science*, 7:258–262, 1962.
- [20] K. S. Arun; T. S. Huang; S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9, 1987.
- [21] B. K. P. Horn. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society of America*, 5:1127–1135, 1988.
- [22] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America*, 4:629–642, 1987.
- [23] P. Hellier. *Recalage non rigide en imagerie cérébrale: méthodes et validation*. PhD thesis, Université de Rennes I, France, 2000.
- [24] A. Guimond; A. Roche; N. Ayache; J. Meunier. Multimodal Brain Warping Using the Demons Algorithm and Adaptive Intensity Corrections. *IEEE Transactions on Medical Imaging*, 20:58–69, 2001.

- [25] C. Chéfd'Hotel; G. Hermosillo; O. Faugeras. Flows of diffeomorphisms for multimodal image registration. In *Proceedings of IEEE International Symposium on Biomedical Imaging*, Washington, USA, 2002.
- [26] J.-P. Thiran; T. Butz. Fast non-rigid registration and model-based segmentation of 3d images using mutual information. In *Proceedings SPIE Medical Imaging*, pages 1504–1515, San Diego, USA, 2000.
- [27] D. Rueckert; I. Somoda; C. Hayes; D. Hill; M. Leach; D. Hawkes. Nonrigid Registration Using Free-Form Deformations: Applications to Breast MR Images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
- [28] R. Chandrashekar; R. H. Mohiaddin; D. Rueckert. Analysis of myocardial motion in tagged MR images using nonrigid image registration. In *Proceedings of Medical Image Understanding and Analysis, MIUA*, pages 1–4, Portsmouth, United Kingdom, 2002.
- [29] J. Vansteenkiste; S. Stroobants; P. Dupont; et al. Fdg-pet scan in potentially operable non-small cell lung cancer : do anatomometabolic pet-ct fusion images improve the localisation of regional lymph node metastases? *European Journal of Nuclear Medicine*, 25(11):1495–1501, 1998.
- [30] C. B. Caldwell; K. Mah; Y. C. Ung; C. E. Danjoux; J. M. Balogh. Fdg-pet/ct integration: Impact on tumour localization and dose volume histograms in radiation therapy. In *Proceedings of the 22nd Annual EMBS International Conference*, pages 1051–1053, Chicago, IL, 2000.
- [31] M. Sato; A. E. Hassanien; M. Nakajima. Non-linear registration of medical images using Cauchy-Navier splines transformation. In *SPIE Conference on Image Processing*, pages 3661–3677, San Diego, California, 1999.
- [32] F. Mahmoud; A. Ton; E. L. Kramer; M. E. Noz. Comparison of three methods for registration of abdominal/pelvic volume data sets from functional-anatomic scans. In *Proceedings of Visualization Development Environments*, pages 205–216, Princeton, NJ, 2000.
- [33] J. S. Fleming; J. H. Conway; L. Bolt; M. Quint; A. A. Salam. Template matching in radionuclide imaging of the lung. In *Proceedings of Medical Image Understanding and Analysis*, 2002.
- [34] T. M. Chung; X. H. Liu; C. H. Chen; X. N. Sun; N. T. Chiu; J. X. Lee. Intermodality registration and fusion of liver images for medical diagnosis. In *Proceedings of International Conference on Intelligent Information Systems*, pages 42–46, 1997.

- [35] Y. Masutani; F. Kimura. Modally Controlled Free Form Deformation for Non-rigid Registration in Image-Guided Liver Surgery. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention, MICCAI*, pages 1275–1278, 2001.
- [36] Y. C. Tai; K. P. Lin; C. K. Hoh; S. C. H. Huang; E. J. Hoffman. Utilization of 3D Elastic Transformation in the Registration of Chest X-ray CT and Whole-Body PET. *IEEE Transactions on Nuclear Science*, 44(4):1606–1612, 1997.
- [37] D. Mattes. Automatic Multimodality Image Registration with Deformations. Technical report, Thesis for the degree of MS in Electrical Eng., Univ. of Washington Medical Center, Seattle, 2000.
- [38] D. W. Townsend; T. Beyer; P. E. Kinahan; T. Brun; R. Roddy; R. Nutt; L. G. Byars. The smart scanner: A combined pet/ct tomograph for clinical oncology. In *Proceedings of IEEE Medical Imaging Conference*, 1998.
- [39] T. F. Hany; H. C. Steinert; G. K. von Schulthess. Improvement of diagnostic accuracy of pet imaging in non-small cell lung cancer using a high performance in-line pet-ct system: Preliminary results. In *European Congress of Radiology*, 2002.
- [40] D.L. Collins. *Brain Warping*, chapter ANIMAL: Automatic Nonlinear Image Matching and Anatomical Labeling. Harcourt Brace College Publishers, 1998.
- [41] O. Camara. *Non-linear registration of thoracic and abdominal CT and PET images: methodological study and application in clinical routine*. PhD thesis, ENST, Paris, 2003.
- [42] O. Colliot. *Représentation, évaluation et utilisation de relations spatiales pour l'interprétation d'images. Application la reconnaissance de structures anatomiques en imagerie médicale*. PhD thesis, ENST, Paris, 2003.
- [43] A. Rosenfeld; J.L. Pfaltz. Distance functions on digital pictures. *Pattern Recognition*, 1:33–61, 1968.
- [44] G. Borgefors. Distance transformations in arbitrary dimensions. *Computer Vision, Graphics and Image Processing*, 27:321–345, 1984.
- [45] P. Besl; N. McKay. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(14):239–256, 1992.
- [46] T. Géraud. *Segmentation des structures internes du cerveau en imagerie par résonance magnétique tridimensionnelle*. PhD thesis, Telecom Paris, 1998.
- [47] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297, 1967.

- [48] E. Diday. Une Nouvelle méthode de classification automatique et de reconnaissance des formes : la méthode des nuées dynamiques. *Revue de Statistique Appliquée*, 19(2), 1970.
- [49] G. Ball; D. Hall. ISODATA, a Novel method of data analysis and pattern classification. Technical report, Stanford Research Institute, 1965.
- [50] M. Kass; A. Witkin; D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1987.
- [51] M. Kass; A. Witkin; D. Terzopoulos. Snakes: Active contour models. *International Journal on Computer Vision*, 1:321–331, 1987.
- [52] C. Xu; J.-P. Prince. Gradient vector flow: A new external force for snakes. In *IEEE Proceedings Conference on Computer Vision Pattern Recognition, CVPR97*, pages 66–71, Puerto Rico, 1997.
- [53] C. Xu; J.-P. Prince. Generalized gradient vector flow external forces for active contours. *Signal Processing, An International Journal*, 71(2):131–139, 1998.
- [54] W. E. Lorensen; H. E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 163–169, 1987.
- [55] J. B. West; J. M. Fitzpatrick; M. Y. Wang; et al. Comparison and evaluation of retrospective intermodality image registration techniques. *Journal of Computer Assisted Tomography*, 21:554–566, 1997.
- [56] J. C. H. Wong; C. Studholme; D. J. Hawkes; M.Ñ. Maisey. Evaluation of the limits of visual detection of image misregistration in a brain fluorine-18 fluorodeoxyglucose pet-mri study. *European Journal of Nuclear Medicine*, 24:642–650, 1998.
- [57] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, Dec 1992.
- [58] M. Unser. Splines. a perfect fit for signal and image processing. *IEEE signal processing magazine*, 16(6):22–38, 1999.
- [59] T. M. Lehmann; C. Gonner; K. Spitzer. Survey: Interpolation methods in medical image processing. *IEEE Transactions on Medical Imaging*, 18:1049–1075, 1999.
- [60] C. Barillot. Fusion de données appliquée à la modélisation de structures cérébrales. Technical report, Laboratoire SIM faculté de médecine, Université de Rennes I, France, 1999.
- [61] O.H. Beahrs, D.E. Henson, R.V. Hutter, and M.H. Myers. *Manual for staging of cancer, 5th ed.* Lippincott Williams Wilkins, 1997.

- [62] T. Blodgett; T. Beyer; G. Antoch; S. Mueller; L. Freudenberg; T. Akhurst. The effect of respiration motion on pet/ct image quality. In *SNM'02*, page 58, june 2002.
- [63] T. Butz; J.-P. Thiran. Feature-space Mutual Information for Multi-modal Signal Processing with Application to Medical Image Registration. In *Proceedings of XI European Signal Processing Conference, EUSIPCO*, volume 1, pages 3–10, Toulouse, France, September 2002.
- [64] O. Camara; G. Delso; I. Bloch. Improving thoracic elastic registration in oncology by using anatomical constraints. In *Proceedings of Medical Image Understanding and Analysis, MIUA*, pages 205–208, Portsmouth, United Kingdom, 2002.
- [65] O. Camara; G. Delso; I. Bloch. Evaluation of a thoracic elastic registration method using anatomical constraints in oncology. In *Proceedings of 2nd Joint Conference of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society, EMBS-BMES'02*, Houston, USA, October 2002.
- [66] M. Charron; T. Beyer; P. E. Kinahan; C. C. Meltzer; D. W. Townsend. Whole-body fdg pet and ct imaging of malignancies using a combined pet/ct scanner. *Journal of Nuclear Mmedicine*, 1999.
- [67] L.D. Cohen and I. Cohen. 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, 1993.
- [68] D. Delbeke. Oncological applications of FDG PET imaging. *Nuclear Medicine*, 40(10):1706–1714, 1999.
- [69] H. Delingette. General object reconstruction based on simplex meshes. *International Journal of Computer Vision*, 32(2):111–146, 1999.
- [70] G. W. Goerres; E. Kamel; T.-N. H. Heidelberg; M. R. Schwitter; C. Burger; G. K. von Schulthess. PET-CT image co-registration in the thorax: influence of respiration. *European Journal of Nuclear Medicine*, 29(3), 2002.
- [71] D. L. G. Hill; P. G. Batchelor; M. Holden; D. J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46:1–45, 2001.
- [72] P. Jannin; J.M. Fitzpatrick; D.J. Hawkes; X. Pennec; R. Shahidi; M.W. Vannier. White paper: Validation of medical image processing in image-guided therapy. In *CARS'02*, pages 299–305, 2002.
- [73] S. Lavallée. Registration for computer-integrated surgery: methodology, state of the art. In R. H. Taylor, S. Lavallée, G. C. Burdea, and R. Mosges, editors, *Computer-integrated surgery, technology and clinical applications*, chapter 5, pages 77–97. MIT Press, 1996.



- [74] S. Lee; G. Wolberg; K. Y. Chwa; S. Y. Shin. Image metamorphosis with scattered feature constraints. *IEEE Transactions on Visualization Computer Graphics*, 2:337–354, 1996.
- [75] J.M.P. Lotjonen. Segmentation of mr images using deformable models: Application to cardiac images. *International Journal of Bioelectromagnetism*, 3(2):37–45, 2001.
- [76] F. Maes; D. Vandermeulen; P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.
- [77] D. Mattes; D. R. Haynor; H. Vesselle; T. K. Lewellen; W. Eubank. Pet-ct image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128, 2003.
- [78] J. Montagnat. *Modèles déformables pour la segmentation et la modélisation d’images médicales 3D et 4D*. PhD thesis, INRIA Sophia-Antipolis, 1999.
- [79] J. Montagnat; H. Delingette; Ñ. Ayache. A review of deformable surfaces: topology, geometry and deformation. *Image and Vision Computing*, 19:1023–1040, 2001.
- [80] J. P. W. Pluim; J. B. A. Maintz; M. A. Viergever. Image Registration by Maximization of Combined Mutual Information and Gradient Information. *IEEE Transactions on Medical Imaging*, 19(8):809–814, 2000.
- [81] A. Roche; X. Pennec; G. Malandain; Ñ. Ayache. Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging*, 20(10):1038–1049, October 2001.
- [82] P. Rogelj; S. Kovacic; J.C. Gee. Validation of a non-rigid registration algorithm for multi-modal data. In *SPIE Conference on Image Processing*, volume 4684, San Diego, California, 2002.
- [83] T. Sederberg and S. Parry. Free form deformation of solid geometric models. In *SIGGRAPH’86*, volume 20, pages 151–160, August 1986.
- [84] R. Shahidi; L. Clarke; R.D. Bucholz et al. White paper: Challenges and opportunities in computer-assisted interventions. *Computer Aided Surgery*, 6(3):176–181, January 2001.
- [85] H.Ñ. Wagner. Fused Image Tomography: An Integrating Force. *Nuclear Medicine*, 40(8):13N–32N, 1999.

