UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Modelado Sólido de Estructuras Óseas a partir de Imágenes de Tomografía

José Manuel Pardo López
Febrero, 1998

DEPARTAMENTO DE ELECTRÓNICA E COMPUTACIÓN

DIEGO CABELLO FERRER, catedrático del Área de Electrónica de la Universidad de Santiago de Compostela

INFORMA

Que procede la presentación de la memoria titulada **Modelado sólido de estructuras óseas a partir de imágenes de tomografía** realizada por D. José Manuel Pardo López bajo mi dirección en el Departamento de Electrónica y Computación de la Universidad de Santiago de Compostela y que dicha memoria constituye la Tesis que presenta para optar al grado de Doctor en Ciencias Físicas.

Santiago de Compostela, 16 de Febrero de 1998.

Fdo.: Senén Barro Ameneiro

Director del Departamento de Electrónica y Computación

Fdo.: Diego Cabello Ferrer

Director de la Tesis

A meus pais e a meu avó

Agradecementos

Quero expresar aquí o meu agradecemento a todas aquelas persoas que dun xeito ou outro teñen contribuído na realización deste traballo.

Ó director desta memoria, Profesor Diego Cabello Ferrer, por brindarme a oportunidade de realizar este traballo e pola axuda no desenvolvemento do mesmo.

Ós médicos D. Joaquín Heras e Profesor José Couceiro polos seus comentarios, axuda na interpretación das imaxes e o sumistro das mesmas.

A tódolos compañeiros do Departamento de Electrónica e Computación, en especial a Elisardo Antelo, José Manuel Mallo, José Antonio Vila, e ós membros do Grupo de Visión Artificial, polo agradable ambiente de traballo creado.

A Xosé Ramón Fernández Vidal e María José Varela Pose pola súa amistade e constante apoio.

Á miña familia e ós amigos por estar sempre preto.

Á Xunta de Galicia polo soporte económico a través do proxecto XUGA20603B96.

Índice General

Índice de Figuras

Índice de Tablas

Capítulo 1

Introducción

La posibilidad de ver dentro del cuerpo humano es una capacidad deseable y necesaria tanto para el estudio de procesos básicos de la vida como para el diagnóstico de los desórdenes que perturban la función normal de los procesos biológicos. Esta capacidad se puede alcanzar mediante intervenciones quirúrgicas que pongan al descubierto las estructuras de interés. Sin embargo, el riesgo y el dolor inherentes hacen que este procedimiento deba ser restringido, en lo posible, tan sólo al acto terapéutico y no como procedimiento de diagnóstico.

La transmisión de energía radiante a través del cuerpo produce imágenes del mismo, y a los niveles de dosis requeridos para obtener imágenes útiles, no afecta directamente a la función de los tejidos. Un haz de radiación que pase a través del cuerpo es dispersado y absorbido con diferentes grados por las diferentes estructuras que encuentra en su camino, dependiendo de su composición. Esta absorción diferencial y el patrón de dispersión de los tejidos se graba para obtener la imagen correspondiente. Dependiendo de la fuente de radiación y de la energía de la misma, cada tejido va a mostrar un comportamiento diferente. Las imágenes producidas por las fuentes de radiación después de atravesar el cuerpo proporcionan una vista de estructuras internas ocultas.

Ha pasado más de un siglo desde que en 1895 el físico alemán Wilhelm Conrad Röentgen obtuviese la primera imagen del esqueleto de la mano de su esposa mediante rayos-X. Desde ese momento el campo de las imágenes médicas no ha dejado de avanzar, tanto en el aspecto teórico como en el tecnológico. El primer sistema radiográfico para la proyec-

ción de estructuras anatómicas fue seguido, varias décadas más tarde, por la tomografía convencional, que permite una focalización dinámica de cortes (*slices*) particulares a lo largo del cuerpo. El desarrollo de esta técnica se basó en un tratado clásico de Radon [1915], quien describió matemáticamente una fórmula de inversión rigurosa para la reconstrucción de un objeto a partir de sus proyecciones. Sin embargo, el primer sistema capaz de obtener la imagen de un corte real no se obtuvo hasta principios de los años setenta, cuando G. N. Houndsfield diseñó el primer escáner de Tomografía Asistida por Computador (CAT). El uso de las imágenes en medicina ha avanzado desde un aspecto meramente cualitativo a otro más cuantitativo a medida que la naturaleza de la información ha pasado de analógica a digital y, en consecuencia, se ha mejorado la calidad de las imágenes. El actual reto está en la obtención de imágenes de volumen lo más precisas posible.

De las muchas posibilidades que existen para la reconstrucción de imágenes a partir de sus proyecciones, la Tomografía Computerizada ha sido la que ha causado una mayor revolución en el momento de su aparición. Las imágenes tomográficas se obtienen a partir de la iluminación de un objeto desde varias direcciones, para cada una de las cuales se registran los campos transmitidos. Cada registro es una medida integrada de algún parámetro del objeto a lo largo de un particular camino de propagación. En la Tomografía Computerizada (CT) de rayos-X, el cuerpo se ilumina con rayos-X que se propagan en línea recta. La medida de los fotones transmitidos permite calcular la integral de línea del coeficiente de atenuación de los rayos-X a lo largo de la linea recta que conecta fuente con detector. Los datos obtenidos de esta forma constituyen una proyección del objeto. Su imagen final se obtiene a partir de las proyecciones tomadas desde diferentes direcciones, para lo cual existen técnicas perfectamente establecidas. La utilidad de la CT se debe al hecho de que diferentes tipos de tejidos tienen diferentes coeficientes de atenuación lineal. De este modo, los diferentes niveles de intensidad en la imagen tomográfica reflejarán la presencia de tejidos diferentes. Con el fin de obtener información médica útil se necesita una razón de muestreo alta de puntos en los cuales estimar dichos coeficientes; lo típico es encontrar del orden de un punto por milímetro, lo cual implica, en general, hablar de varios millones de puntos. En objetos estacionarios, como huesos, se puede simplificar el problema de computación recogiendo datos de una sección en cada momento y reconstruir esa sección a partir de los datos recibidos a través de ella. Este es el esquema típico en los dispositivos clínicos de CT de rayos X usados actualmente.

La Tomografía Computerizada ha revolucionado el diagnóstico radiológico desde su introducción hace dos décadas en la práctica médica diaria [Herman, 1995]. Este paso tecnológico desde proyecciones 2D de cuerpos 3D mediante técnicas de transmisión (rayos-X) a técnicas de obtención de imágenes verdaderamente 3D (CT) se puede considerar como el salto cualitativo más importante en el campo de la imagen médica. Esta capacidad de disponer de datos de volumen permite realizar medidas físicas en el mismo número de dimensiones espaciales que las del cuerpo humano original, a la vez que ofrece la posibilidad de descubrir lo que hay dentro de un objeto sin necesidad de entrar en él. La posibilidad de visualizar estructuras anatómicas 3D internas ha traido consigo la posibilidad del desarrollo de herramientas útiles para el diagnóstico y la planificación del tratamiento. Las técnicas 3D adquieren la información espacial en la forma de elementos de volumen (voxels) densamente muestreados. Generalmente, el muestreo consiste en la obtención de imágenes 2D de secciones transversales suficientemente próximas para que no se produzca una excesiva pérdida de información, y lo más distantes posibles con el fin de evitar un exceso de radiación, o de no sobrepasar limitaciones temporales. Si la distancia entre cortes es comparable a la dimensión del voxel dentro del corte, se puede considerar como un volumen captado de forma continua.

El estudio de las imágenes 3D ya se ha introducido en numerosos campos de la práctica médica diaria. El interés médico es obvio, sin más que ver la cantidad de trabajos sobre la materia que se pueden en la bibliografía. Esto concierne tanto al aspecto de diagnóstico (detección, caracterización y cuantificación de lesiones) como al de planificación quirúrgica; en ambos casos se está convirtiendo en una herramienta potente e indispensable [Roux y Coatrieux, 1990]. La disponibilidad de sistemas de reconstrucción 3D permite al médico la realización de una morfometría más precisa, mejor interpretación de datos de volumen y una mejor planificación en la terapia.

Por otra parte, la adecuada planificación preoperatoria es fundamental para alcanzar resultados exitosos en una intervención quirúrgica. Planificar o simular un procedimiento quirúrgico significa definir trayectorias de los instrumentos y actos operativos relacionados, anticipar posibles resultados morfológicos y funcionales de la intervención y evaluar el conjunto de circunstancias que condicionan la misma. Los sistemas de planificación requieren la reconstrucción 3D de la estructura anatómica de interés a partir de los datos de volumen y de la forma más precisa posible.

La construcción de modelos tridimensionales a partir de imágenes radiográficas y de

tomografía computerizada tiene aplicaciones clínicas en el análisis, simulación y diseño de una gran variedad de procedimientos de reconstrucción ortopédica [Murphy y col., 1986]. Aquí los esfuerzos suelen concentrarse en tres direcciones. Por una parte está el intento de mejora del conocimiento fundamental sobre la morfología de las estructuras de interés y sus propiedades mecánicas. Esto está relacionado con la cuantificación y medidas 3D, tanto sobre estructuras sanas como patológicas, especialmente en el caso de deformaciones traumáticas. Otro aspecto importante concierne al diseño de prótesis e implantes a medida del paciente específico mediante técnicas CAD/CAM. Por último, restan las técnicas de visualización 3D y la simulación y planificación preoperatoria de la intervención quirúrgica, que ya empiezan a fructificar en sistemas experimentales. Con la disponibilidad de reconstrucciones tridimensionales a partir de imágenes de CT se pueden analizar cambios en la forma, superficie de contacto, localización y orientación de las articulaciones. Estas reconstrucciones se pueden usar para simular osteotomías, testar sus mecanismos teóricos y cuantificar las correcciones a alcanzar por cada tipo de osteotomía.

Otro campo de la cirugía ortopédica particularmente adecuado para el análisis asistido por computador es la sustitución total de la articulación de rodilla (PTR). Las dos principales áreas de aplicación son la selección de prótesis estándares para individuos particulares y la manufactura de prótesis a medida.

Las prótesis de rodilla son ampliamente utilizadas en cirugía reconstructiva de la articulación de rodilla, en casos de enfermedad degenerativa articular (EDA) en estadío avanzado, artritis inflamatoria de la rodilla, artritis reumatoide y colagenopatías autoinmunes, así como en la EDA de secundaria a traumatismos. Se ha comprobado que si la técnica es correcta, los resultados son satisfactorios en más del 95% de los casos, incluso 10 años después de la operación.

La fijación de la prótesis puede realizarse a través de cemento acrílico o por medio de un sistema de fijación por presión, sin cemento. Con este segundo procedimiento es necesaria una extremada precisión en la preparación del hueso, de tal forma que el contacto con la prótesis sea lo más extenso y estable posible. De ahí que la planificación quirúrgica parezca altamente ventajosa a la hora de determinar la mejor localización y orientación de secciones y perforaciones, y de simular el resultado de la implantación [Dario 1995,1996]. En la implantación simulada de prótesis se pueden cuantificar de forma precisa la cantidad de hueso a eliminar y la superficie de contacto entre la prótesis y el hueso. Por otra parte,

se puede predecir el comportamiento de la interfaze entre prótesis implantada y hueso a través del análisis tensional. Toda esta información *a priori* permitiría alcanzar una mayor adecuación geométrica y mecánica, con lo que se reduciría la incidencia de fallo (aflojamiento).

El análisis tensional está justificado por el hecho de que gran cantidad de estructuras biológicas se ven significativamente influenciadas por su entorno mecánico [Hart y col., 1992] . El hueso, en particular, posee una estructura de crecimiento y auto-reparación, y puede adaptar su arquitectura a cambios en su uso mecánico (remodelación adapatativa). Los fallos en los implantes de prótesis en articulaciones tales como cadera y rodilla son ejemplos prácticos del interés clínico en el estudio de la remodelación adaptativa del hueso. Estos fallos son a veces debidos a la pérdida ósea en la vecindad del implante, y un factor decisivo en estos fallos es un entorno mecánico no favorable que induce la re-absorción y la consiguiente pérdida del implante. De esta forma, la previsión de la respuesta del hueso vivo a los cambios en la carga y a las prótesis depende de un correcto conocimiento de su entorno mecánico.

El método de los elementos finitos se ha convertido en el método más ampliamente usado en el estudio del entorno mecánico del hueso, debido a la posibilidad que ofrece de tratar sus complejidades inherentes (formas irregulares del hueso, la inhomogeneidad y anisotropía de los tejidos óseos, o la compleja naturaleza de variación con el tiempo de las condiciones de carga y contorno en el hueso vivo) [Keyak y col., 1990]. Para abordar todo este análisis se ha propuesto en varias ocasiones el uso de herramientas que incluyen el procesado y análisis de imágenes para la creación del modelo sólido. Estos sistemas tendrían numerosas ventajas, entre las que podemos considerar la realización de una planificación sobre un modelo de articulación específico del paciente, que incluiría simulación cinemática y dinámica. Por otra parte, el uso del computador permite al usuario considerar un mayor número de variables y parámetros para simular y resolver reconstrucciones particularmente complejas. Con el fin de crear un modelo tridimensional de un hueso es preciso ejecutar una serie de procesos que comprenden: exploración, modelado sólido (segmentación + mallado), modelado mediante elementos finitos y análisis cinemático, tal y como se muestra en la figura 1.

El primer paso en el uso del método de los elementos finitos para el análisis del entorno mecánico del hueso es construir un modelo sólido válido. Un modelo sólido es una descripción gráfica y matemática de algún objeto geométrico, a la vez completa y sin

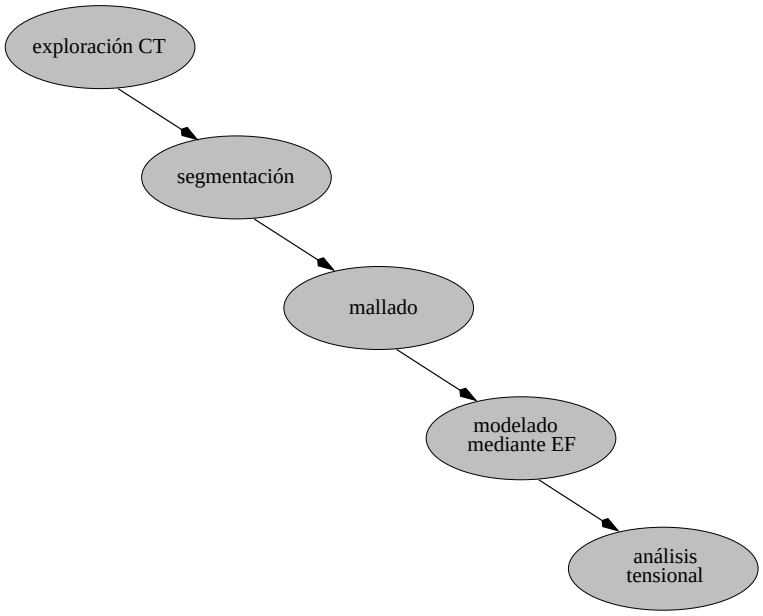


Figura 1.1: Procesos en el modelado 3D.

ambigüedades. Tal objeto, una vez definido como un modelo sólido, puede ser analizado y sometido a manipulaciones matemáticas. Propiedades tales como centro de gravedad, volumen, masa o momentos de inercia pueden ser calculadas directamente a partir de la descripción de los objetos en el computador. Los modelos sólidos también pueden ser manipulados gráficamente usando funciones Booleanas. Por ejemplo, se pueden unir en compuestos o ensamblados, diseccionados o separados en sub-objetos, reorientados para mejorar la visualización, etc. Se pueden simular también perturbaciones físicas: por ejemplo, se pueden aplicar fuerzas a un sistema de objetos y monotorizar la distribución de tensiones. En esencia, se puede introducir en el modelo información geométrica completa concerniente a un objeto, y esta información se puede usar para simular y evaluar las respuestas de los objetos ante las condiciones actuantes.

Para la representación 3D de la geometría de objetos físicos se han implementado métodos muy diversos. Los mejor conocidos son las representaciones de ocupación espacial, los modelos *wire frame*, los modelos de superficie, las representaciones de contorno o las representaciones de geometría sólida constructiva.

La información necesaria para poder obtener la geometría del hueso la proporciona una

exploración mediante Tomografía Computacional (CT). Por otra parte, para el análisis mediante elementos finitos se requiere disponer tanto del modelo geométrico anterior, como de información relativa a la estructura material del hueso. Estos datos se obtienen para un hueso vivo también a partir de los datos de intensidad proporcionados por los datos de CT. La precisión del análisis mediante elementos finitos depende en parte del modelo inicial. Éste debe ajustarse lo más posible al modelo real, de tal forma que no se eliminen de él partes del hueso, a la vez que no deben incorporarse regiones no pertenecientes al hueso. La falta de precisión del modelo sólido restaría valor al análisis tensional, ya que el caso analizado no se correspondería con el real.

Desafortunadamente, el principal inconveniente del método de análisis mediante elementos finitos está en la enorme cantidad de trabajo manual que se requiere para obtener una segmentación que permita generar el modelo tridimensional adecuado para cada paciente. El grado de dificultad en la automatización de una tarea de segmentación particular es frecuentemente proporcional a la importancia médica o científica del resultado. En la práctica, el uso efectivo de las imágenes médicas todavía requiere una cierta cantidad de edición manual. Con frecuencia, la edición manual es el único o el último método de selección en la segmentación de imágenes [Robb, 1995]. Por eso ha sido frecuente el uso del modelos 2D, más simples, en la investigación ortopédica [Keyak y col., 1990]. En otros casos, los estudios teóricos sobre la aplicación del método de los elementos finitos al análisis tensional sobre huesos se han efectuado sobre huesos *in vitro* o modelados manualmente [Rakotomanana y col., 1992; Keyak y col., 1993; Steele y col., 1994; Müller y Rüegsegger, 1995]. Sin embargo el modelado 3D permite una mejor adecuación al modelo anatómico real, al tiempo que proporciona una definición más precisa de los movimientos y las fuerzas actuantes.

El objetivo que perseguimos en este trabajo es implementar un sistema que de forma automática y a partir de una serie de imágenes de CT que correspondan a cortes transversales de una estructura anatómica, genere su modelo sólido. Sus aplicaciones más inmediatas serán la disposición de un modelo geométrico 3D de la estructura específica de cada paciente para el análisis tensional mediante elementos finitos, o la posibilidad de visualizar directamente en 3D información proporcionada originariamente como un conjunto de imágenes en 2D. Este modelo geométrico proporcionará, pues, ayuda para el diagnóstico y la planificación quirúrgica.

Planteamiento general

En este trabajo perseguimos pues el desarrollo de un sistema automático para la reconstrucción 3D de la geometría del hueso específico de un paciente que libere al médico de la edición manual, al tiempo que proporcione una alta precisión en la definición de dicha geometría. Esta reconstrucción de la geometría constituirá el paso previo al análisis cinemático. La definición precisa de la geometría es esencial para modelar de forma adecuada la cinemática a partir de las superficies 3D. Los datos disponibles a partir de las exploraciones médicas de CT son, con frecuencia, de precisión insuficiente para proporcionar una buena definición de la superficie del hueso. Generalmente presentan baja resolución o contraste insuficiente entre hueso y tejidos blandos. Esto hace de la segmentación una tarea difícil y un proceso fundamental en el que se basará la calidad del modelo sólido obtenido.

Al tratarse de conjuntos de datos 3D, la segmentación podría efectuarse con métodos que operen directamente en 3D, o mediante el procesado corte a corte en el que la segmentación final vendría dada por el conjunto apilado de contornos 2D. Nosotros hemos optado por la segunda posibilidad por dos razones principales. Una de ellas es que la exploración CT proporciona en el eje longitudinal una menor definición que en el plano del corte tranversal. Aunque la resolución se puede igualar mediante la generación de cortes intermedios mediante interpolación, no se consigue una verdadera continuidad en los niveles de gris. Estos cortes interpolados facilitan la segmentación corte a corte, pero no resuelven el problema de falta de continuidad 3D necesaria para operadores 3D. Por otra parte, la natural variación de formas entre estructuras de individuos diferentes y la presencia de lesiones no permiten disponer de modelos 3D adecuados que guíen el proceso de segmentación. Estos modelos aproximados si están disponibles en una segmentación 2D corte a corte, en la que el contorno resultante de la segmentación en un corte sirve de modelo para la segmentación del siguiente corte de la secuencia.

La segmentación en cada corte consistirá en identificar en la imagen los tejidos de interés. Esta es una etapa crucial, necesaria para concentrar las siguientes operaciones en un conjunto pertinente y más compacto de datos. La segmentación de imágenes CT se hace particularmente difícil por la estructura inhomogénea de los huesos, los cuales presentan texturas en su interior. Esta textura provoca problemas en la obtención de los contornos externos del hueso y, por consiguiente, en la separación automática del hueso de

los tejidos que lo rodean. Las regiones se podrían obtener extrayendo áreas homogéneas con respecto a un criterio de homogeneidad dado, o mediante la detección de contornos, es decir, puntos de transición entre dos áreas homogéneas. En nuestro sistema hemos escogido combinar ambas metodologías con el fin de dotarlo de una mayor robustez.

Así, en un procesado de bajo nivel, las regiones se pueden definir en base a la agrupación en conjuntos conexos de puntos (pixels) que satisfagan una misma condición de homogeneidad sobre alguna de sus propiedades. Sin embargo, esto no es suficiente para el caso de estructuras óseas, debido a que éstas no son homogéneas, sino que están constituidas por tejidos de características muy diversas. Por consiguiente, en este tipo de análisis, tras una etapa de bajo nivel, se precisa una nueva etapa de alto nivel que incorpore conocimiento específico sobre el tipo de imágenes concreto para realizar un proceso de agrupación de las regiones iniciales en estructuras mayores con significado. Para lograrlo hemos implementado un sistema de unión y división de regiones basado en reglas.

Por otra parte, los bordes se definen como puntos de transición entre regiones homogéneas. Idealmente, los puntos de borde formarían contornos cerrados. Sin embargo, en la detección de estos puntos mediante operadores clásicos la propia complejidad de las imágenes o los problemas asociados al ruido inherente a ellas provocan, por una parte, la aparición de puntos de borde falsos, y por otra, la no detección de puntos de borde verdaderos. Los resultados de esta detección se pueden mejorar localmente mediante el desarrollo de métodos que permitan conectar pequeñas discontinuidades entre trozos del mismo contorno real, a la vez que eliminen puntos de borde espúreos. Sin embargo, en estos procedimientos se realiza un análisis local sobre los contornos detectados del objeto, pero sin tener en cuenta el posible conocimiento global *a priori* sobre la forma de los contornos buscados. Para poder incorporar este conocimiento consideraremos la aplicación del método de los modelos deformables. En este sentido, hemos realizado aportaciones a los métodos existentes en los siguientes aspectos: introducción de nuevos términos de energía, tanto interna como externa, que mejoren el rendimiento del método para nuestro caso particular, y simplificación del método de minimización, lo que permitirá una mayor velocidad de convergencia.

Una vez que se extrae el contorno de una estructura en una imagen 2D, podemos seguirlo de un corte a otro, y finalmente reconstruir la superficie 3D a partir de los contornos apilados. Sin embargo, el hecho de que los huesos sean estructuras heterogéneas, así como la proximidad existente entre diferentes estructuras óseas, o la presencia de

lesiones y deformaciones y la excesiva separación que, en algunos casos, se produce entre cortes sucesivos, pueden degradar el rendimiento del anterior sistema basado en bordes. Es por esto por lo que en general, aunque sea a costa de un mayor tiempo de computación, resulta ventajoso la integración de la información de regiones con la de contornos. Por ello propondremos un sistema basado en el modelo de los Campos Aleatorios de Markov que integre la información de las dos metodologías diferentes desarrolladas, una basada en regiones y otra en contornos, con el fin de obtener una segmentación más precisa.

Una vez determinados los contornos exteriores de las estructuras de interés (huesos), se procederá con un sistema de etiquetado que permite identificar y extraer las zonas de lesión presentes en el hueso. Para ello se implementará de nuevo un modelo de Campos Aleatorios de Markov. Finalmente, se realizará una descripción de la superficie que encierra el volumen del hueso que sea fácilmente transferible como modelo geométrico de entrada a un sistema de análisis mediante elementos finitos.

La memoria la hemos estructurado en 5 capítulos. Así, tras este primero de introducción y planteamiento general, en el segundo capítulo realizamos una revisión de las diferentes técnicas para el análisis de imágenes propuestas en la bibliografía. Se hace en él una descripción de los diferentes esquemas, avanzando en el sentido de complejidad creciente. Finalmente, se muestran diversos ejemplos de sistemas de análisis de aplicación específica, principalmente en el campo de las imágenes médicas. En los capítulos 3 y 4 se describe nuestro sistema, diferenciando entre la parte correspondiente a la determinación de la geometría externa del hueso y la que corresponde con la identificación de lesiones presentes en el mismo. En el capítulo 5 se describe el proceso de definición del modelo sólido a través de la puesta en correspondencia de los puntos de contorno en cortes diferentes, lo cual permite la definición de la superficie que delimita la estructura de interés. Esta definición del modelo sólido servirá como definición geométrica del problema para el análisis tensional mediante elementos finitos. Finalmente, se mostrarán las conclusiones y principales aportaciones del trabajo.

Capítulo 2

Análisis de imágenes. Reconstrucción 3D.

2.1 Introducción

El análisis de imágenes es un proceso que persigue descubrir, identificar y comprender aquellos patrones en la imagen que son relevantes en un dominio de aplicación. Uno de los principales objetivos al implementar un sistema de análisis de imágenes consiste en dotarle de una capacidad de percepción similar, en algún sentido, a la de los seres humanos. El sistema debería poseer la capacidad de extraer la información de interés, separándola de los detalles irrelevantes, la capacidad de aprender a partir de ejemplos y de generalizar este conocimiento para aplicar en nuevas circunstancias, y la capacidad de realizar inferencias a partir de información incompleta. Sin embargo, los sistemas informáticos actualmente disponibles, así como los conocimientos biológicos, aún resultan insuficientes para emular el sistema visual humano a la hora de realizar tareas genéricas de análisis automático de imágenes. Las técnicas más avanzadas de análisis de imágenes por computador se basan en su mayor parte en fórmulas heurísticas, adaptadas para la resolución de problemas específicos.

En el campo de las imágenes médicas, en la última década se ha hecho un considerable esfuerzo para la obtención de técnicas de adquisición capacidades avanzadas de procesado, análisis y visualización de imágenes que aseguren una extracción fiable de información con

significado clínico [Baker, 1990; Udupa y col. 1991; Narayan y col. 1994]. Estos esfuerzos se han orientado fundamentalmente hacia la obtención de métodos de visualización de la información contenida en ellas para una mejor percepción por parte del experto. Existe pues una necesidad, aún no totalmente cubierta, de analizar y modelar de forma automática la información proveniente de las imágenes, proporcionando así una ayuda eficiente al clínico.

Por otra parte, la visualización de datos de volumen de imágenes médicas proporciona información muy valiosa para la evaluación clínica. Ya hemos comentado la necesidad de disponer de imágenes 3D en procesos tales como planificación quirúrgica, diseño e implante de prótesis, seguimiento de la evolución de la terapia, etc. En este contexto, el modelado con imágenes médicas debe proporcionar información estructural de los diferentes elementos. Nuestro interés, además de visualizar la reconstrucción 3D de las estructuras, está en la simulación de la deformación física de dichas estructuras sometidas a cargas. La manipulación estructural de datos o el análisis dinámico requieren de un formalismo de modelado y herramientas de diseño. La reconstrucción y el modelado sólido 3D implica tanto una segmentación 2D ó 3D de las estructuras de interés en un conjunto de imágenes 3D, como un modelado 3D posterior a partir de dicha segmentación. En este capítulo describiremos en qué consiste cada uno de estos procesos y efectuaremos una revisión de los métodos mas habituales para abordarlos. Así, en primer lugar, abordaremos la segmentación de bajo nivel, describiendo una serie de esquemas de segmentación que no hacen uso de conocimiento sobre el dominio de aplicación. A continuación introduciremos los sistemas de análisis de imágenes basados en conocimiento. Estos sistemas hacen uso de características de bajo nivel (carácter general), junto con otras de nivel medio alto (específicas del dominio) para la interpretación de las imágenes. Posteriormente abordaremos el problema de la reconstrucción 3D de un objeto a partir de una secuencia de contornos planos paralelos. Finalmente completaremos el capítulo con la descripción de algunos sistemas para el análisis de imágenes de dominios restringidos, haciendo énfasis en aquellos desarrollados en el campo de las imágenes médicas.

2.2 Segmentación

Uno de los problemas más importantes en un sistema de visión es el de la segmentación [Ballard y Brown, 1982; Rosenfeld y Kak, 1982, González y Wintz, 1987]. La idea de la segmentación tiene su origen en los trabajos de psicólogos de la Gestalt, que estudiaron las preferencias exhibidas por los seres humanos en la organización de conjuntos de formas dispuestos en el campo visual. Los principios de la Gestalt dictan ciertas preferencias en la agrupación basadas en características tales como proximidad, similaridad y continuidad. En visión, la agrupación de partes de una imagen en unidades que resulten homogéneas con respecto a una o más características origina una imagen segmentada. Mediante la segmentación se obtiene un primer nivel descriptivo de la imagen en el que los modelos internos de los objetos, dependientes del dominio, empiezan a influir en la agrupación de las estructuras de la imagen en unidades con significado en el dominio.

Formalmente, la segmentación de una imagen la podemos definir como una partición de la misma en regiones, tal que:

- La unión de las partes es igual a la imagen completa: $\bigcup_{i=1}^{k} P_i = I$.
- Las partes son disjuntas: $P_i \bigcap P_j = \emptyset \quad$ si $i \neq j$
- Cada parte P_k satisface un predicado, $\mathcal{P}(P_k) = \mathcal{V}$; es decir, todos los puntos de la partición tienen una propiedad común.
- El predicado es falso para la unión de cualquier número de partes adyacentes: $\mathcal{P}(P_k \bigcup P_l) = \mathcal{F}$, si $k \neq l$ y $\mathcal{ADY}(k, l) = \mathcal{V}$.

Existen métodos muy diversos de abordar la segmentación de una imagen. Atendiendo a la clasificación propuesta por Wilson [1985], podemos distinguir entre tres tipos de metodologías básicas: clasificación estadística, métodos basados en detección de contornos y crecimiento de regiones. La forma más simple de obtener una segmentación en una imagen monocroma consiste en asociar un rango de niveles de gris con cada clase de objetos. Cuando este método simple falla es natural buscar bases estadísticas más rigurosas para la segmentación. Una forma obvia de mejorar el resultado es tratar la segmentación como un problema de decisión, en el cual las estadísticas de la población total de pixels se consideran como una mezcla de las de las regiones componentes en

las que se va a segmentar la imagen. Tales métodos proceden por estimar primero los parámetros relevantes, tales como media y varianza, para las distribuciones componentes de la mezcla, y después seleccionan los umbrales óptimos para separar los pixels que pertenecen a las diferentes poblaciones basándose en algún tipo de decisión, tales como mínimo coste o mínima probabilidad de error de clasificación. Las distribuciones componentes podrían estimarse usando un modelo paramétrico, por ejemplo el Gaussiano, y aplicando técnicas de estimación Bayesianas o de máxima semejanza, o usando un método derivado heurísticamente como la detección de agrupaciones (*clustering*). Sin embargo, los métodos estadísticos tienen serias deficiencias cuando se aplican a la segmentación. Entre ellas está la incapacidad de los métodos estadísticos de primer orden para tratar con las relaciones espaciales de los datos. Para solventar estas deficiencias se han propuesto métodos iterativos o de relajación, en los cuales se hace probabilística una clasificación inicial de los pixels, y ésta va siendo adaptada sucesivamente en referencia a las clases de los pixels vecinos. Un método alternativo es el de división y unión de regiones (*split-and-merge*). Volviendo al problema de incertidumbre, uno de los puntos débiles de la mayoría de los procedimientos de clasificación estadísticos es la no consideración de información de localización espacial, por lo que, en general, la coherencia espacial y la localización de las regiones resultantes de la segmentación es fortuita.

La carencia de la localización espacial y los problemas relacionados se pueden superar con el uso de los métodos basados en la detección de bordes. Estos son métodos basados en la convolución de la imagen con una función escogida para dar respuesta máxima en áreas de la imagen donde el nivel de gris cambia rápidamente. De esta forma, la segmentación se hace encontrando los pixels que están sobre el contorno de una región. Además de la localización espacial, se necesita alguna forma de localidad en el dominio de la frecuencia para evitar que el ruido y la textura fina degraden el comportamiento de los algoritmos. Estos requerimientos de localización espacial e inmunidad al ruido entran en conflicto por el principio de incertidumbre, puesto que el aumento de la inmunidad al ruido implica mayor promediado espacial, lo cual reduce la calidad de la localización espacial. Para solventar este problema se ha propuesto el uso de filtros para diferentes escalas, obteniéndose así una estimación más fiable de la posición del borde de la que se puede obtener mediante una simple operación de filtrado. No obstante, de entre las técnicas de segmentación basadas en la detección de bordes, la más sofisticada corresponde a la segmentación basada en modelos deformables. Esta técnica, introducida por Kass

y col. [1988], ha experimentado una gran popularidad en los últimos años debido a la facilidad con la que integra información obtenida a partir de la imagen con conocimiento a priori sobre el dominio.

El otro tipo de técnicas de segmentación consiste en el crecimiento de regiones. La idea de esta técnica es simple: dada una serie de núcleos iniciales en la imagen, se examinan las particiones adyacentes, una a una, para ver si están lo suficientemente próximas a las propiedades actuales de la región núcleo y ser aceptadas como parte de esa región. La partición inicial puede ser la imagen original, con lo cual cada partición consta de un pixel, o bien estar formada por la clasificación obtenida mediante la aplicación de una segmentación previa. La motivación detrás de este planteamiento es clara: es un intento de considerar a la vez la localización en el espacio y la similaridad de las propiedades. La dificultad con la que se encuentran la mayoría de estos métodos es que, como en muchos problemas de optimización no lineal, tanto la solución final como el número de iteraciones requeridas para alcanzarla son altamente dependientes de las condiciones iniciales: número y localización de los centros iniciales de crecimiento. En general no está garantizado que, dada una imagen, se obtenga el mismo resultado para el mismo algoritmo y dos conjuntos de puntos iniciales diferentes.

Las técnicas de particionado de una imagen en regiones se pueden clasificar también en base a otros criterios. Otros autores, Ballard y Brown [1982] entre ellos, distinguen sólo entre métodos basados en regiones y métodos basados en bordes.

El proceso de segmentación particiona la imagen en regiones uniformes con respecto a alguna de sus propiedades. El siguiente paso en el análisis de imágenes consiste en la identificación de las estructuras de interés. Estas estructuras se corresponderán, en el caso general, con varias de las regiones obtenidas tras el proceso de segmentación. Es por ello necesario la inclusión de una etapa que asocie regiones con objetos. Este proceso necesita la incorporación de conocimiento específico del dominio de trabajo. Se acostumbra a dividir el espectro de procesos involucrados en el análisis automático de imágenes en tres grupos básicos: procesado de bajo nivel, procesado de nivel medio, y procesado de alto nivel. Esta clasificación se hace de acuerdo con la cantidad de conocimiento incorporado en la operación, donde *conocimiento* significa ligaduras implícitas o explícitas sobre la posibilidad de un determinado agrupamiento. Estas ligaduras pueden obtenerse a partir de argumentos físicos generales o de restricciones más fuertes sobre la imagen, obtenidas a partir de consideraciones dependientes del dominio. Mucho conocimiento significa un

alto grado de restricción sobre algunas de las características de la estructura a identificar y que sus relaciones con las demás estructuras están también muy bien definidas. Poco conocimiento a priori significa que el análisis debe proceder más en base a pistas locales y suposiciones de tipo general (independientes del dominio), con pocas restricciones sobre el resultado final. El procesado de bajo nivel sólo se aplica a situaciones muy simples. Cuando el nivel de complicación aumenta es necesario incluir conocimiento específico sobre el dominio. Así, los sistemas de análisis de imágenes especializados incorporan procesos de bajo nivel, basados en características locales (de los pixels), que permiten obtener una descripción preliminar de la escena. A continuación viene una etapa de nivel medio en el que se toman en cuenta forma y relaciones de las estructuras percibidas y que reducen el número de objetos por unión de varios de ellos. Este proceso de reducción en el número de estructuras se consigue mediante la introducción de restricciones impuestas sobre características de mayor complejidad que los pixels de la etapa anterior. Las etapas de alto nivel contienen características e interpretaciones globales, que pueden provocar la modificación sobre parámetros de las etapas previas, precisando más el conocimiento de los niveles medio y bajo, y finalmente identificarán las estructuras presentes en la imagen.

En los siguientes apartados de esta sección describiremos los métodos más habituales, entre los muchos que existen, para la segmentación de bajo nivel. En la siguiente sección abordaremos los sistemas para interpretación de imágenes basados en conocimiento de alto nivel. Al final del capítulo describiremos una serie de sistemas de este tipo que aparecen descritos en la bibliografía.

2.2.1 Segmentación basada en la clasificación de pixels

La segmentación basada en la clasificación de pixels pretende clasificar los pixels de la imagen en un número mas reducido de clases que el que presenta la imagen original. Para ello evalúa para cada pixel un conjunto de propiedades, y es dentro de este espacio de propiedades donde se establece la clasificación de los pixels en diferentes clases. A cada pixel se le asigna la etiqueta de la partición del espacio de medidas al que pertenece. Como consecuencia de esta clasificación se obtiene un mapa de regiones en el espacio imagen. Los segmentos en la imagen se definen como las componentes conectadas de los pixels que tienen la misma etiqueta de clase. El atributo más básico para este tipo de segmentación, en imágenes monocromas, es la amplitud o intensidad del nivel de gris. A

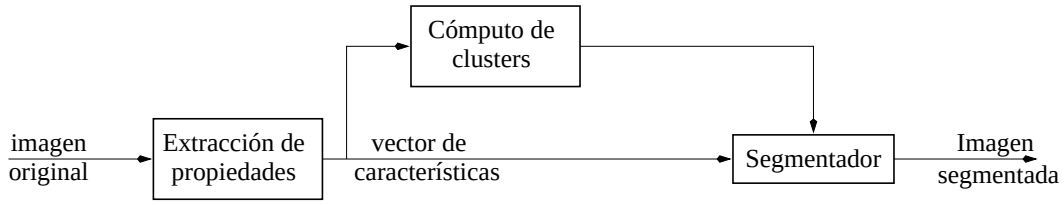


Figura 2.1: Esquema simplificado de segmentación por clasificación.

medida que el conocimiento sobre los objetos a identificar aumenta es posible disponer de un mayor número de características en base a las cuales clasificar los pixels como correspondientes a un objeto o a otro. En métodos de segmentación más elaborados, a parte de las consideraciones sobre características de pixels de la imagen, se tienen en cuenta relaciones de tipo espacial. Esto sucede por ejemplo, en los sistemas de crecimiento de regiones y/o detección de bordes, que se describirán más adelante.

La clasificación que sólo tiene en cuenta propiedades de los pixels sin considerar relaciones espaciales, recibe el nombre de clasificación o agrupación estadística (*clustering*). Este tipo de segmentación se basa en las ideas subyacentes en el reconocimiento estadístico de patrones, entendiendo por patrón una descripción cuantitativa de un objeto o de alguna otra entidad de interés en una imagen. En general, un patrón está representado por uno o más descriptores. El reconocimiento de patrones supone la utilización de técnicas que permitan asignar los patrones a sus respectivas clases, entendiendo por clase una familia de patrones que comparten algunas propiedades comunes.

En la segmentación basada en la clasificación de pixels la entidad de interés es el pixel. Consideremos un vector de medidas $x = [x_1, \ldots, x_N]$ para cada pixel (i, j) de una imagen. Si el conjunto de medidas va a ser efectivo para la segmentación, los datos recogidos en varios pixels dentro de un segmento de atributos comunes deben ser similares. Si se da esta condición la segmentación consiste en subdividir el espacio de medidas N-dimensional en compartimentos mutuamente exclusivos, cada uno de los cuales encierra un grupo (*cluster*) de datos típicos para cada segmento de la imagen. El diagrama de flujo de un algoritmo simple de clasificación se muestra en la figura 2.1 [Coleman y Andrews, 1979]. En la primera etapa se computan las características. En el siguiente paso se determina el número óptimo de agrupaciones (*clusters*) y el centro del espacio de características para

cada grupo. En el segmentador se asigna cada vector de propiedades con el centro del grupo más próximo. La diferencia entre otras técnicas de segmentación de imágenes y la agrupación estadística es que en esta última la agrupación se hace en el espacio de medidas, mientras que en la segmentación la agrupación se hace en el dominio espacial de la imagen. Los métodos de agrupación estadística se pueden dividir en dos grupos: supervisados y no supervisados. En la agrupación supervisada el conjunto de muestras de entrenamiento que se escogen para caracterizar las diferentes clases se etiquetan de acuerdo a clasificaciones conocidas a priori. Estos métodos necesitan una importante interacción por parte del usuario. Si las muestras no son etiquetadas a priori y la clasificación se establece estimando los parámetros a partir de las muestras, la segmentación se define entonces como no supervisada.

Los clasificadores más simples son aquellos basados en la mínima distancia entre un patrón y el centro de cada clase. Supongamos que cada clase de patrones w_j se caracteriza mediante un vector medio:

$$m_j = \frac{1}{N_j} \sum_{x \in w_j} x \quad j = 1, \ldots, M \tag{2.1}$$

donde N_j es el número de patrones vectoriales de la clase w_j y la suma se realiza para todos los vectores. Una forma de determinar la pertenencia a una clase de un patrón vectorial desconocido x consiste en asignarlo a la clase del prototipo más próximo. Si se utiliza la distancia euclídea para determinar el grado de proximidad, se reduce el problema del cálculo de las medias de distancia:

$$D_j(x) = ||x - m_j|| \quad j = 1, \ldots, M \tag{2.2}$$

Si $D_i(x)$ es la menor distancia, entonces se asigna x a la clase w_i

En una imagen monocroma, la forma más simple de clasificación consiste en analizar el histograma y sobre él establecer el límite entre las diferentes clases presentes en la imagen. En general la separación entre clases no resulta tan simple. En estos casos se puede obtener información de probabilidades a partir del histograma y a partir de ahí derivar un sistema de clasificación que atienda a requerimientos tales como minimizar el coste del error de clasificación. En la mayoría de los campos relacionados con la medición e interpretación, las consideraciones probabilísticas tienen su importancia en el reconocimiento de patrones

debido a la aleatoriedad a la que normalmente está sometida la generación de clases de patrones. El planteamiento Bayesiano es apropiado cuando no hay ambigüedad sobre el patrón a clasificar mismo [Pao,1989]. Se sabe exactamente lo que refleja el patrón, aunque se pueda desconocer la causa responsable de su manifestación. La tarea de clasificación consiste en decidir cuál de las posibles causas es de hecho la responsable de ese patrón particular. La cuestión principal en la clasificación estadística no es tanto si se cometerán errores o no, sino como minimizar el coste global de los errores.

La clasificación estadística empieza con unidades, tales como pixels, regiones o segmentos de una imagen, sobre los cuales puede efectuarse un conjunto de medidas. Cada unidad tiene asociado un vector de medidas. El propósito de la clasificación estadística es etiquetar cada unidad en base a ese vector de medidas. La clasificación *empareja* la unidad con el vector de características de la categoría más próxima, mediante una regla de decisión. Por tanto estas técnicas incluyen los siguientes procesos: (1) selección de las propiedades y técnicas de extracción, (2) construcción de la regla de decisión, y (3) técnicas de estimación de las reglas de decisión.

En todo proceso de clasificación se parte de un conjunto de medidas D, cuyos elementos pueden ser un único valor o una n-tupla de valores. A cada una de estas medidas, d_i, se le asociará una categoría c_i perteneciente a un conjunto C. Las estadísticas del proceso global de clasificación se pueden describir en términos de las siguientes probabilidades:

$P(c_i)$: probabilidad a priori de que un patrón pertenezca a la clase c_i, independientemente de la identidad del patrón.

$P(x_k)$: probabilidad de que un patrón sea x_k, sin considerar su clasificación.

$P(x_k|c_i)$: probabilidad condicional de que el patrón sea x_k, dado que la clase es c_i.

$P(c_i|x_k)$: probabilidad condicional a posteriori de que la clase del patrón sea c_i, dado que el patrón es x_k.

$P(c_i, x_k)$: probabilidad conjunta de que el patrón sea x_k y de que su clase sea c_i.

En estas definiciones tanto x_k como c_i son cantidades discretas y cumplen las siguientes condiciones de normalización:

$$\sum_i P(c_i) = 1 \tag{2.3}$$

$$\sum_k P(x_k) = 1 \tag{2.4}$$

La probabilidad conjunta se puede expresar de dos formas equivalentes:

$$P(c_i, x_k) = P(x_k|c_i)P(c_i) \tag{2.5}$$
$$P(c_i, x_k) = P(c_i|x_k)P(x_k) \tag{2.6}$$

Y a partir de estas expresiones se obtiene:

$$P(c_i|x_k) = \frac{P(x_k|c_i)P(c_i)}{P(x_k)} \tag{2.7}$$

Esta última relación se puede usar para estimar los valores de la probabilidad a posteriori $P(c_i|x_k)$ si estas estadísticas no son conocidas directamente y si las probabilidades condicionales de clase y las probabilidades a priori son conocidas.

En la clasificación probabilística, la probabilidad a posteriori se trata de la misma forma que todas las otras medidas de probabilidad. La probabilidad a posteriori es una *probabilidad objetiva* e indica la frecuencia relativa de ocurrencia en un experimento aleatorio. En el caso más simple, dado x_k, se evaluaría $P(c_i|x_k)$ para todas las clases en C, y se decidiría en favor de la clase i para la cual $P(c|x_k)$ es mayor. En el caso más general se construyen *funciones de decisión* $g_i(x_k)$, una para cada clase, y la regla de decisión sería como sigue:

x_k pertenece a la clase c_j si y sólo si:

$$g_j(x_k) > g_i(x_k) \quad \forall i = 1, \quad i \neq j \tag{2.8}$$

$g_i(x_k)$ podría ser cualquier función de $P(c_i|x_k)$. A partir de estas funciones de probabilidad condicional, se pueden obtener funciones de riesgo y tomar decisiones en base al mínimo riesgo o máxima ganancia,

$$R_i(x_k) = l_{ii}P(c_i|x_k) + \sum_{j \neq i} l_{ij}P(c_j|x_k), \tag{2.9}$$

donde l_{ij} representa la penalización asociada a la decisión de asociar la clase c_i cuando en realidad es la c_j. Cuando no hay penalización en el caso de la suposición correcta ($l_{ii} = 0$),

y cuando las suposiciones incorrectas se penalizan por igual ($l_{ij} = l, \forall j \neq i$), se recupera la regla de decisión de la probabilidad máxima a posteriori (MAP), la cual consiste en decidir en favor de c_i si y sólo si $P(c_i|x_k) > P(c_j|x_k), \forall j \neq i$.

Para poder realizar el acto de clasificación es preciso cierto conocimiento que lo apoye. Este conocimiento se adquiere durante la fase de aprendizaje. En esta fase se cuantiza el espacio de patrones (formados por un vector de propiedades y una pertenencia de clase) en un conjunto de celdas de tamaño finito, y se agrupan todos los patrones dentro de cada celda para dar un valor posicionado en el centro de dicha celda. A partir de esto se pueden obtener funciones continuas en la forma de funciones de distribución de densidades de probabilidad dependientes de clase.

Para las representaciones continuas, las distribuciones de densidad son $p(x|c_j)$ y $P(c_j)$, y la condición de normalización es

$$\sum_i \int_v p(x|c_i)P(c_i)dx = 1 \tag{2.10}$$

Se ha visto como la representación de información de patrones mediante el uso de distribuciones conjuntas o condicionales permite interpolar y extrapolar datos de patrones. La clasificación en base a la máxima probabilidad a posteriori constituye un procedimiento óptimo. Sin embargo esto no significa que no haya errores, sino que el error global está en un mínimo.

Las distribuciones de probabilidad conjuntas continuas son particularmente adecuadas para el origen y naturaleza del error del sistema. Por ejemplo, en una clasificación binaria utilizando la regla de MAP, para un x se decide que

$$x \in C_1 \iff P(c_i|x) > P(c_2|X) \quad \text{y} \quad x \in c_2 \quad \text{en otro caso} \tag{2.11}$$

Para cualquier x, la densidad de probabilidad de cometer un error viene dada por

$$P_e(x) = \min\{P(c_1|x), P(c_2|x)\} \tag{2.12}$$

En particular, para el caso unidimensional, las estadísticas se pueden describir en términos de dos distribuciones unimodales que representan las dos distribuciones de probabilidad conjuntas, como se ilustra en la figura 2.2.

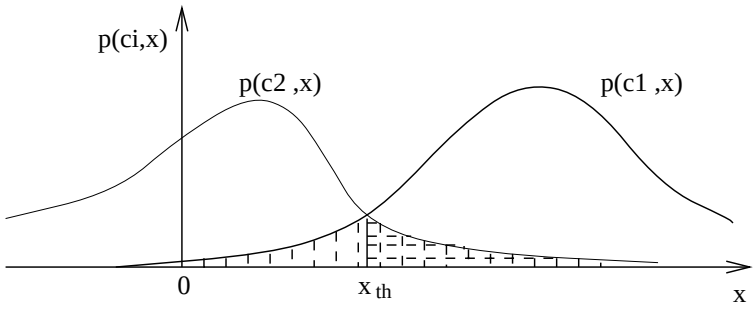


Figura 2.2: Ilustración de un discriminante de umbral.

Sea x_{th} el valor de x donde las dos distribuciones de probabilidad conjunta intersectan. Entonces, decidiremos que $x \in c_1$ si y sólo si $x > x_{th}$, sino $x \in c_2$. La probabilidad de error promediado sobre todos los valores de x es

$$\begin{aligned}
\langle P_e \rangle &= \int_{-\infty}^{x_{th}} P(c_1|x)p(x)dx + \int_{x_{th}}^{-\infty} P(c_2|x)p(x)dx \\
&= \int_{-\infty}^{x_{th}} \frac{p(x|c_1)P(c_1)}{p(x)}p(x)dx + \int_{x_{th}}^{\infty} \frac{p(x|c_2)P(c_2)}{p(x)}p(x)dx \\
&= \int_{-\infty}^{x_{th}} p(x,c_1)dx + \int_{x_{th}}^{\infty} p(x,c_2)dx
\end{aligned} \tag{2.13}$$

Este error promediado se corresponde con el área sombreada en la figura 2.2. En este caso el acto de clasificación no requiere un conocimiento detallado de las distribuciones de probabilidad. Basta con conocer el número x_{th}, el cual funciona como discriminante. Situaciones más complejas requieren del conocimiento de mayor número de discriminantes. Y para dimensiones superiores, los umbrales se transforman en curvas, superficies e hipersuperficies. Aunque la clasificación sea estadística tiene la apariencia de determinística.

En general, un discriminante es una función que aplicada a un patrón da una salida que es, o bien una estimación de la pertenencia de clase de ese patrón, o una estimación de los valores de uno o más de los atributos del patrón. Una vez que la regla de decisión de Bayes se transforma en la acción de un discriminante, la tarea de aprendizaje durante la sesión de entrenamiento se transforma en un aprendizaje de la métrica del espacio de patrones, o aprendizaje de la posición de prototipos, o aprendizaje de la localización de las hipersuperficies. Aparentemente, se olvidan todas las nociones estadísticas y todo es determinístico. Evidentemente, esto no es cierto. En la clasificación estadística el objetivo

no es eliminar los errores, sino asegurar el mínimo error global. Como ejemplos de sistemas de clasificación estadística en aplicaciones de segmentación de imágenes médicas se pueden citar los de Lei y Sewchand [1992] para la segmentación de imágenes de CT, y Liang [1993] para la segmentación de imágenes de MR.

Sin embargo, en la mayoría de los casos no se recurre al estudio probabilístico para determinar las reglas y umbrales de decisión, sino que se hace en base a procedimientos heurísticos más rápidos. A continuación describiremos algunos de los más representativos.

Umbralización

Si los valores de intensidad de un objeto están en un intervalo y los valores de intensidad de los pixels de fondo están fuera de este intervalo, se puede obtener una imagen binaria, que particiona la imagen original, usando una operación de umbralización sobre los valores de intensidad de los pixels de la imagen. Para que esta umbralización sea efectiva es preciso que exista suficiente contraste entre los objetos de interés y el fondo y se conozcan los niveles de intensidad asociados a los objetos o al fondo. Sea $F[i,j]$ la imagen original, la imagen binaria resultante de la segmentación por umbralización, $F_T[i,j]$, se define como:

$$F_T[i,j] = \begin{cases} 1 & \text{si } F[i,j] \in Z \\ 0 & \text{otro caso} \end{cases} \tag{2.14}$$

donde Z es un conjunto de valores de intensidad para las componentes de los objetos. Este proceso se puede ejecutar de forma paralela, ya que el resultado obtenido para cada pixel depende únicamente del valor de intensidad en ese punto. Por otra parte, se incluye conocimiento acerca del dominio de aplicación. El umbral utilizado proviene del conocimiento a cerca del tipo de escena, ya que no se determina a partir de ningún procesado sobre la imagen. En muchos casos, las primeras ejecuciones del sistema se pueden usar para analizar interactivamente una escena y determinar un valor apropiado para el umbral; en tal caso sería el operador humano el que aporte el conocimiento. En general interesa hacer que la segmentación sea independiente del operador humano, con el objetivo de hacerla más rápida y objetiva. Por eso se han desarrollado muchas técnicas para utilizar la distribución de intensidad en una imagen y el conocimiento sobre los objetos de interés para seleccionar automáticamente un valor de umbral adecuado.

Para hacer la segmentación más robusta, el sistema debería seleccionar automáticamente el umbral. Para ello se necesita disponer de conocimiento sobre los objetos de la escena, tal como: características de intensidad de los objetos, tamaños de los objetos, fracciones de una imagen ocupadas por los objetos, o número de objetos presentes en la imagen. Las técnicas de umbralización automática analizan el histograma usando el conocimiento para seleccionar el umbral más apropiado.

Supongamos que una imagen contiene n objetos $O_1, O_2, \ldots, O_n$, incluyendo al fondo, y valores de gris de diferentes poblaciones $\Pi_1, \Pi_2, \ldots, \Pi_n$, con distribuciones de probabilidad $p_1(z), \ldots, p_n(z)$. La mayoría de los métodos de selección automática del umbral usan el tamaño y la probabilidad de ocurrencia, y estiman las distribuciones de intensidad a partir de los histogramas. El método *P-Tile* [Jain, 1995] hace uso del conocimiento acerca del tamaño del objeto deseado para umbralizar la imagen. Si se conoce que un objeto ocupa el $p\%$ de una imagen, se pueden escoger uno o varios umbrales que asignen el $p\%$ del área de la imagen al objeto. Este método sólo es aplicable a casos muy simples. Si los objetos tienen diferentes niveles de gris entre si y a su vez diferentes del fondo, y los pixels están afectados por ruido Gaussiano de media cero, entonces se puede suponer que los valores de gris se obtienen a partir de distribuciones normales con parámetros $(\mu_1, \sigma_1), \ldots, (\mu_n, \sigma_n)$. El problema consiste, pues, en determinar los picos y valles del histograma. Para una umbralización automática, se debe efectuar una medida del grado de pico y grado de valle de cada punto del histograma. El grado de pico se determina basándose en la altura de los picos y la profundidad de los valles, y se ignorarán picos locales entre los que no hay una distancia mínima.

En general, la determinación de umbrales de decisión en el histograma no es una tarea simple. Así, existen diversos métodos para la selección iterativa de umbral. Estos método comienza con un umbral aproximado y sucesivamente va refinando esta estimación en base a alguna propiedad de las subimágenes resultantes de la umbralización. Otsu [1979] define el mejor umbral como aquél para el que la suma ponderada de las varianzas dentro de cada partición es mínima. Kittler y Illingworth [1985] sugieren un criterio diferente. Suponen que la escena obedece a una mezcla de dos distribuciones Gaussianas y buscan el umbral que minimice una función en la que se pondera el error de clasificación para todos los puntos de la imagen.

Cuando se trabaja con imágenes complejas, los métodos basados en histogramas se comportan pobremente. Sus limitaciones más básicas se deben al hecho de que no consid-

eran información espacial acerca de los valores de intensidad sobre una imagen. Distintas imágenes con muy diferentes distribuciones espaciales de niveles de gris pueden tener histogramas similares. No se explota el hecho importante de que puntos del mismo objeto están, generalmente, espacialmente próximos debido a la coherencia de la superficie.

Si la iluminación no es uniforme en toda la imagen o la distribución de gris es desigual en el fondo, es más adecuado considerar pequeñas regiones de una imagen y analizarlas por separado para determinar el umbral que le afectará de forma individual. La segmentación final será la unión de las regiones de las subimágenes. Estos métodos se conocen como *segmentación adaptativa*, y combinan información espacial con información de intensidad de gris para la determinación del umbral. Chow y Kaneko [1972] dividen la imagen en bloques mutuamente exclusivos, calculan el histograma de cada bloque y determinan el umbral para cada histograma. Consideran el umbral aplicado al centro de cada bloque y al resto de pixels se le asigna un umbral por interpolación del obtenido para el centro de los bloques. Weska y col. [1974] sugieren determinar un histograma sólo para pixels con magnitud de Laplaciano alta. La razón es que los pixels de los contornos corresponden a cruces por cero del operador Laplaciano, por lo que los valores altos de la convolución con este operador corresponderán a puntos del borde interno de las regiones. De esta forma se evita introducir los pixels de transición entre regiones haciendo el histograma más abrupto; al tiempo que se tiende a introducir un número semejante de puntos de los objetos y del fondo, por lo que los dos modos del histograma tendrán igual tamaño.

Otra técnica útil en el caso de iluminación desigual es aproximar los valores de intensidad por una función simple tal como un plano o una bicuadrática. Esta aproximación se hace en gran parte a partir de los valores del fondo. Puntos de la imagen por encima de la función serán parte del objeto y los demás fondo. Esta técnica recibe el nombre de *umbralización variable* [Jain, 1995] .

En muchas aplicaciones se conoce que ciertos valores de gris pertenecen a los objetos. Sin embargo, lo común es que existan valores adicionales que pertenezcan a los objetos y al fondo. Las técnicas de *doble umbralización* intentan resolver este problema. Panda y Rosenfeld [1978] sugirieron el uso de dos umbrales, uno para pixels de gradiente bajo y otro para pixels de gradiente alto, para la segmentación de manchas claras sobre fondo oscuro. Los pixels de gradiente bajo corresponderán al interior de los objetos y darán lugar por tanto un histograma claramente bimodal; mientras tanto, los pixels con gradiente alto serán mezcla de las dos clases y será unimodal. En este último caso un buen umbral de

decisión puede ser el valor medio de estos pixels. De esta forma se realiza la agrupación en un espacio de medidas bidimensional: gradiente e intensidad.

En otros casos se pueden usar umbrales conservativos para obtener el núcleo del objeto y a continuación emplear algún método de crecimiento de regiones. Estos métodos son dependientes de la aplicación específica. Un planteamiento común es el del usar otro umbral para aceptar pixels que tengan vecinos que son núcleo o, usando características de intensidad tales como un histograma, para determinar puntos que están debajo de un segundo umbral y están conectados al conjunto original de puntos.

Un ejemplo de sistema de segmentación que emplea la doble umbralización es el empleado por Brunie y col. [1991] para la segmentación 3D de los huesos de la muñeca. Aquí se aplica un procedimiento de segmentación independiente para cada corte (*slice*) 2D más que una técnica directa en 3D. Es después de la segmentación cuando se clasifican los objetos como los diferentes huesos. La idea del algoritmo de umbralización empleado es la siguiente: después de emplear un umbral bajo, los contornos externos de los huesos son bastante buenos, pero aparecen conexiones entre huesos distintos; mientras que tras una umbralización alta, los huesos están separados, pero los contornos exteriores empeoran. Para integrar ambas informaciones lo que hacen es partir de la umbralización alta, en la que las clases aparecen separadas, y dilatar dichos objetos hasta que alcancen los contornos de la umbralización baja. Así se obtienen los contornos de los diferentes objetos en cada corte. La reconstrucción 3D se completa asignando la misma etiqueta a los contornos que pertenecen al mismo hueso en los diferentes cortes. Para ello se construye un grafo de adyacencias en donde la relación de adyacencia se verifica si el centro de masas de uno de ellos caen en la elipse de aproximación del otro.

Algoritmo K-medias

Otro ejemplo tipo de clasificación por agrupación (*clustering*) es el algoritmo K-medias. El nombre de este algoritmo hace referencia a que se conoce que existen K clases o patrones, siendo por tanto necesario aportar este dato a priori. Este algoritmo se basa en la minimización de las sumas de las distancias al cuadrado de todos los puntos del dominio al respectivo centro de cluster:

$$\min \sum_{x \in S_j(k)} (x - Z_j)^2 \tag{2.15}$$

donde $S_j(k)$ es el dominio del cluster para el centro z_j en la n-ésima iteración.

El algoritmo K-medias actúa de forma iterativa sobre el conjunto de objetos a clasificar $x_1, x_2, \ldots, x_p$. Establecido previamente el número exacto de clases existentes, digamos K, se escogen aleatoriamente K vectores, $z_i^0, \quad i = 1, \ldots, K$ de entre los elementos a agrupar. Estos constituirán inicialmente los centros de las clases $\alpha_i^0, \quad i = 1, \ldots, K$. A partir de aquí el algoritmo entra en su parte iterativa, y en cada iteración r se hace una redistribución de las muestras $\{x_j\}, \quad 1 \leq j \leq P$, entre las K clases de acuerdo con la siguiente regla:

$$x \in \alpha_j^r \quad si \quad \| x - z_j^r \| < \| x - z_i^r \| \tag{2.16}$$

$$\forall i = 1, 2, \ldots, K \quad i \neq j \tag{2.17}$$

Una vez hechas las asignaciones de clase, se recalculan los centro de las clases. El objetivo en el cálculo de los nuevos centros es minimizar el índice de rendimiento siguiente:

$$J_i = \sum_{x \in \alpha_j^r} \| x - z_i^r \|^2, \quad i = 1, 2, \ldots, K \tag{2.18}$$

Este índice se minimiza utilizando la media aritmética de α_i^r:

$$z_i^{(r+1)} = \frac{1}{N_i^r} \sum_{x \in \alpha_i^r} x, \quad i = 1, 2, \ldots, K \tag{2.19}$$

siendo N_i^r el número de elementos de la clase α_i en la iteración r. Si se verifica que $||z_i^{(r+1)} - z_i^r|| < \epsilon, \quad \forall i = 1, 2, \ldots, K$, siendo $\epsilon > 0$ el factor de convergencia, la clasificación terminó. En caso contrario se realiza una nueva iteración.

Como ejemplo de la aplicación de este algoritmo en segmentación de imágenes médicas se puede citar el trabajo de Loncaric y col. [1995], para la detección de hemorragias en el cerebro a partir de imágenes CT.

Clusterización borrosa

La clusterización borrosa (*fuzzy*) es una generalización de la clusterización clásica, en la que cada elemento de un conjunto de muestras se asignaba a un único cluster. En los

métodos de clusterización borrosa, basados en el concepto de conjunto borroso introducido por Zadeh, cada muestra se asocia a varios clusters con distintos grados de pertenencia. Este grado de pertenencia se representa mediante los *coeficientes de pertenencia*, que pueden variar entre 0 y 1.

Consideremos el espacio R^s de características para la clasificación. Un clasificador para un conjunto de datos en c clases es una función de decisión $D : R^S \rightarrow I_c$, donde I_c es el conjunto de etiquetas de clase. Un clasificador borroso indicará el grado de pertenencia de un vector de propiedades a cada una de las c clases.

Un ejemplo de clasificador borroso K-NN es aquél que comienza con un conjunto de aprendizaje etiquetado (X, W), donde la matriz de etiquetas W constituye la c-partición de X, y cada uno de sus elementos w_{ik} representa el coeficiente de pertenencia del patrón k a la clase i. Estos coeficientes han de satisfacer las siguientes restricciones:

$$w_{ik} \in [0, 1] \tag{2.20}$$

$$\sum_{i=1}^{c} w_{ik} = 1 \tag{2.21}$$

$$0 < \sum_{k=1}^{n} w_{ik} < n \tag{2.22}$$

para $1 \leq i \leq c, \quad 1 \leq k \leq n$, donde c representa el número de clases y n el número de patrones en X. El vector columna $W_k = (w_{1k}, \ldots, w_{ck})^t$ es el vector de etiquetas del patrón X_k. La clasificación de una nueva muestra se hace mediante la asignación de coeficientes de pertenencia:

$$^K l(y) = \frac{1}{K} \sum_{j \in j_K(y)} W_j \tag{2.23}$$

donde $J_k(y)$ son los índices de los K vectores de datos en X más próximos a y de acuerdo con la función de distancia considerada en R^s. $^K l(y)$ es en general, un vector de etiquetado borroso para y, donde la i-ésima componente se corresponde con el coeficiente de pertenencia de y a la clase i.

Hay al menos tres circunstancias en las cuales los conceptos y las técnicas de la teoría de conjuntos borrosos son útiles en el reconocimiento de patrones [Jain, 1995] . Uno de los casos es aquél en el que el conjunto borroso sirve como interfaze entre las características

simbólicas, no numéricas, y las medidas cuantitativas. La segunda circunstancia es referente al nivel de pertenencia de clase, más que al nivel de característica. De esta forma se permite que un patrón pueda pertenecer a cada una de las clases con ciertos grados de pertenencia. La ventaja de este método aparece cuando las decisiones se propagan a través de una red con otras decisiones relacionadas. La tercera circunstancia de uso es cuando los valores de la función de pertenencia se usan para ayudar a proporcionar una estimación del conocimiento incompleto.

Como ejemplo reciente de la aplicación de técnicas borrosas en el análisis de imágenes está el trabajo de Richardt y col. [1996] que aborda el problema de identificar patrones geométricos que puedan ser descritos mediante ecuaciones matemáticas. Otro ejemplo de clusterización borrosa es el algoritmo K-medias borroso (*fuzzy k-means*) [Bezdek, 1981]. La función a minimizar por el algoritmo K-medias borroso es la suma de las distancias al cuadrado de todos los puntos x_k a su respectivo centro de cluster v_i, cuando las distancias están ponderadas por alguna potencia de los coeficientes de pertenencia W_{ik}:

$$J_m(W,X) = \sum_{k=1}^{n}\sum_{i=1}^{c}(w_{ik})^m(d_{ik})^2 \tag{2.24}$$

$$v_i = \frac{\sum_{k=1}^{n}(w_{ik})^m x_k}{\sum_{k=1}^{n}(w_{ik})^m} \tag{2.25}$$

$$w_{ik} = [\sum_{j=1}^{c}(\frac{d_{ik}}{d_{jk}})^{\frac{2}{(m-1)}}]^{-1} \tag{2.26}$$

El algoritmo de minimización, sería uno iterativo en el que mientras no se cumpla la condición de convergencia impuesta ($\| W_{ik}^r - W_{ik}^{(r-1)} \| \leq E, \quad \forall i,k$) se calcularán los nuevos centros de cluster v_i^r a partir de $W^{(r-1)}$, y los nuevos coeficientes de pertenencia W^r a partir de dichos centros de cluster, V^r. Este tipo de algoritmo ha sido utilizado con frecuencia, por ejemplo, en la segmentación de imágenes de Resonancia Magnética (MR) [Bezdek y col., 1993].

Clasificación mediante redes neuronales

Las redes neuronales son otro método frecuente de implementación de esquemas de clasificación. Su principal atractivo reside en su habilidad para particionar el espacio de propiedades usando superficies no lineales para la separación de las clases. Estas fronteras de separación se obtienen mediante procesos de entrenamiento. Además, sucede a

menudo que las propiedades estadísticas de las clases de patrones de un problema son desconocidas, o no es posible realizar una estimación de las mismas. En la práctica, estos problemas de decisión se gestionan mejor utilizando métodos que obtienen directamente, mediante el entrenamiento, las funciones de decisión requeridas. Así, no es necesario realizar suposiciones sobre las funciones de densidad de probablilidad subyacentes o sobre cualquier otra información estadística relativa a las clases de patrones consideradas. Las redes neuronales están formadas por multitud de elementos de cálculo no lineales y elementales, organizados como redes que se asemejan a la forma en que se piensa que están interconectadas las neuronas en el cerebro. Como ejemplo de aplicación en clasificación de redes neuronales se puede citar el trabajo de Dhawan y Arata [1993] para la segmentación de imágenes médicas. La principal limitación de las redes residen en su incapacidad de tomar en cuenta información estructural [Pun y col., 1994].

2.2.2 Crecimiento de regiones

Los métodos descritos con anterioridad se basan en propiedades de puntos de una imagen. Con ellos, la segmentación se obtiene al pasar al dominio imagen la información de clases obtenida en el espacio de propiedades. La extensión lógica consistiría en utilizar propiedades espaciales de la imagen en el proceso de segmentación.

Las técnicas de crecimiento de regiones emplean, además de las consideraciones de similaridad de propiedades, otras que tienen que ver con la información de localización espacial. En su forma más simple, el crecimiento de regiones comienza con un pixel y entonces examina sus vecinos con el fin de decidir si tienen intensidades similares. Si eso ocurre, se agrupan para formar una región mayor. Aquellos pixels del entorno que no se unen a ninguna de las regiones de crecimiento actualmente consideradas pasan a ser un nuevo núcleo de crecimiento. Las formas más avanzadas de crecimiento de regiones comienzan con una partición de la imagen en regiones, generalmente, pequeñas y con un grado de homogeneidad elevado. Estas regiones semilla se hacen crecer para formar regiones mayores que satisfagan un determinado conjunto de ligaduras. El predicado de homogeneidad se puede basar en cualquier característica de las regiones.

No es posible conocer por adelantado como se debería dividir la imagen en regiones elementales que respondan a determinados modelos, puesto que este es el problema de la segmentación misma. Por este motivo se suelen usar métodos conservadores para generar

las regiones semilla del crecimiento, generalmente basados en el conocimiento sobre el dominio o la naturaleza de las imágenes. A partir de estas regiones se van formando otras mayores añadiéndoles puntos compatibles. El crecimiento de regiones es, pues, un método secuencial de extracción de los objetos o regiones de una imagen. El objetivo ahora es extraer objetos de un determinado tipo, más que particionar completamente la imagen. Para extraer estos objetos se puede usar un amplio abanico de criterios.

Unión de regiones

La operación de unión (*merging*) combina regiones consideradas similares. La operación más importante es, precisamente, la de determinar el grado de la similaridad entre dos regiones. Los planteamientos que juzgan la similaridad se basan bien en los valores de gris de las regiones y sus estadísticas asociadas, bien en la debilidad de los contornos entre regiones, y pueden incluir la propiedad de proximidad espacial. Dos posibles formas para juzgar la similaridad entre regiones adyacentes son: (1) comparar los valores de sus características; si éstas no difieren en más que un valor predeterminado, las regiones se consideran similares y serán candidatas a ser unidas; (2) asumir que los valores de intensidad se obtienen de una distribución de probabilidad y considerar si conviene, o no, unir regiones adyacentes basándose en la probabilidad de que tengan la misma distribución estadística de valores de intensidad.

El método estadístico asume que las regiones en una imagen tienen un valor de gris constante corrompido por ruido Gaussiano de media cero, de tal forma que los valores de gris se obtienen a partir de distribuciones normales. Supóngase que dos regiones adyacentes R_1 y R_2 contienen m_1 y m_2 puntos, respectivamente. Hay dos posibles hipótesis. (H_0): ambas regiones pertenecen al mismo objeto; en este caso las intensidades se sacan todas de una única distribución Gaussiana con parámetros (μ_0, σ_0). (H_1): las regiones pertenecen a diferentes objetos; en este caso las intensidades de cada región se obtienen de distribuciones Gaussianas diferentes con parámetros (μ_1, σ_1) y (μ_2, σ_2).

En general, estos parámetros no son conocidos, sino que se estiman usando muestras. Por ejemplo, cuando una región contiene n pixels con niveles de gris $g_i, \quad i = 1, 2, \ldots, n$ obtenido a partir de una distribución normal dada por:

$$p(g_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(g_i-\mu)^2}{2\sigma^2}} \tag{2.27}$$

las ecuaciones de la estimación máxima verosimilitud (Maximum Likelihood, (ML)) para los parámetros vienen dadas por:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} g_i \tag{2.28}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(g_i - \hat{\mu})^2 \tag{2.29}$$

Bajo la suposición H_0, todos los pixels se obtienen independientemente a partir de una única distribución, $N(\mu_0, \sigma_0)$. La densidad de probabilidad conjunta bajo H_0 viene dada por:

$$p(g_1, \ldots, g_{m_1+m_2}|H_0) = \prod_{i=1}^{m_1+m_2} p(g_i|H_0) = \frac{1}{(\sqrt{2\pi}\sigma_0)^{m_1+m_2}} e^{-\frac{\sum_{i=1}^{m_1+m_2}(g_i-\mu_o)^2}{2\sigma_0^2}} \tag{2.30}$$

$$= \frac{1}{(\sqrt{2\pi}\sigma_0)^{m_1+m_2}} e^{-\frac{m_1+m_2}{2}} \tag{2.31}$$

Bajo la hipótesis H_1, m_1 pixels pertenecen a la región 1 con una distribución $N(\mu_1, \sigma_1)$ y m_2 pixels pertenecen a la región 2 con una distribución $N(\mu_2, \sigma_2)$. La densidad de probabilidad conjunta será ahora:

$$p(g_1, \ldots, g_{m_1+m_2}|H_1) = \frac{1}{(\sqrt{2\pi}\sigma_1)^{m_1}} e^{-\frac{m_1}{2}} \frac{1}{(\sqrt{2\pi}\sigma_2)^{m_2}} e^{-\frac{m_2}{2}} \tag{2.32}$$

La razón de semejanza L se define como:

$$L = \frac{p(g_1, g_2, \ldots|H_1)}{p(g_1, g_2, \ldots|H_0)} = \frac{\sigma_0^{m_1+m_2}}{\sigma_1^{m_1}\sigma_2^{m_2}} \tag{2.33}$$

Si L es menor que un umbral, hay fuerte evidencia de que sólo debe existir una región y de que se debe producir la unión de ambas.

Otra forma de abordar el proceso consiste en combinar dos regiones si el borde entre ellas es débil. Este enfoque intenta eliminar bordes entre regiones adyacentes considerando no sólo características de intensidad, sino también el tamaño de los perímetros de las regiones, así como el del perímetro compartido. El contorno común se elimina si el borde es débil según alguna medida y en el contorno resultante no se producen cambios de nivel

de gris demasiado rápidos. Además, habrá que tener en cuenta otro tipo de consideraciones que afecten a la congruencia de bordes.

La unión de regiones se puede implementar representando éstas y sus adyacencias mediante un grafo. La unión de regiones se logra mediante la contracción del grafo, recalculando las propiedades siempre que se unan dos regiones (nodos).

División de regiones

Las técnicas de división de regiones se basan en el hecho de que si alguna propiedad dentro de una región no es constante, ésta debería dividirse. La segmentación basada en la división (*splitting*) comienza con regiones grandes, en muchos casos comienza con la imagen completa, y la va dividiendo hasta obtener particiones homogéneas. El problema está, generalmente, en decidir cuando una propiedad no es constante. Estas cuestiones dependen de la aplicación y requieren del conocimiento de las características de las regiones en esa aplicación. Más difícil que decidir sobre la homogeneidad de las propiedades es el determinar por donde dividir una región. Por este motivo la división de regiones resulta más compleja que la unión.

Uno de los planteamientos más simples consiste en subdividir cada región en un número fijo de subregiones del mismo tamaño. Se analiza cada nueva subregión creada y se vuelve a dividir siempre que no se cumpla la condición de homogeneidad. Estos se llaman métodos de descomposición regular, entre los que se encuentra la representación *quadtree*, cuyo esquema de representación se ilustra con el ejemplo de la figura 2.3 [Samet, 1990]. Debido a las numerosas formas en las que se puede dividir una región, la división es, en general, más difícil que la unión.

División y unión

Las técnica de división y unión pueden verse como una extensión del método de crecimiento de regiones y testeo de homogeneidad. Esta técnica fue propuesta por primera vez por Pavlidis y Horowitz [1977] para adecuación a curvas (*curve fitting*), y más tarde la extendieron a regiones. Un ejemplo simple de este método comienza con la imagen completa como segmento inicial. Entonces cada segmento se divide en cuatro partes (*quadtree)* si no se verifica la condición de homogeneidad. Después de la división se aplica el test de

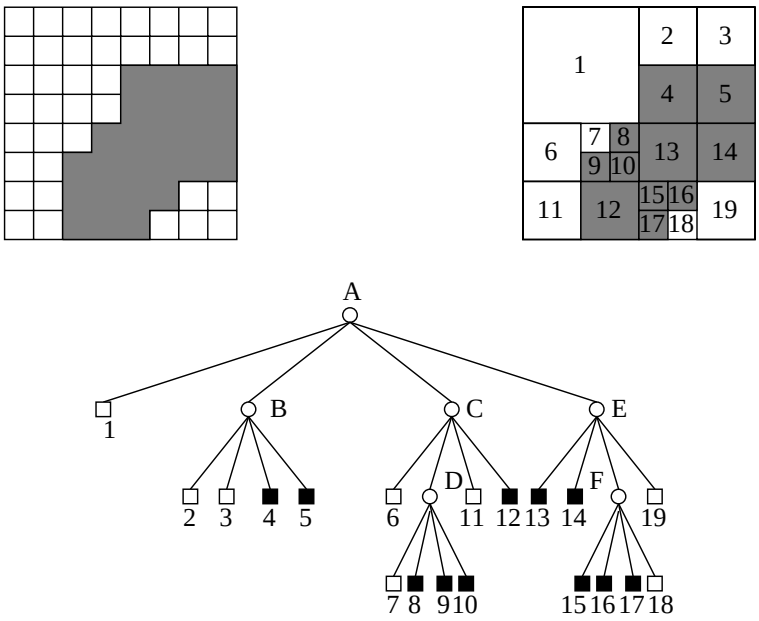


Figura 2.3: Ejemplo de quadtree: imagen original, cuadrantes homogéneos y árbol asociado.

uniformidad a grupos apropiados de cuatro regiones adyacentes. Si el test se satisface, entonces las cuatros partes se unen para formar una región mayor. El proceso termina cuando no son posibles más operaciones de división y unión. La segmentación final se obtiene, finalizadas las operaciones de división y unión, tras la agrupación de regiones adyacentes. Para ello es preciso pasar de la representación *quadtree* a la representación mediante un *grafo de adyacencias de regiones* [Strasters y Gerbrands, 1991]. Con ello se permite realizar uniones de regiones adyacentes que verifican en conjunto el criterio de homogeneidad, pero que no pueden unirse dentro de la representación quadtree a causa de un nodo padre inhomogéneo, o porque las regiones están en diferentes niveles de nodos. Esta técnica se puede extender de forma directa a 3D. En los trabajos de Jackins y Tanimoto [1980], Meagher [1982, 1987], Yau y Srihari [1983], Chen y Huang [1988], Hull y col. [1990], Ibaroudene y col. [1990], Samet [1990a, 1990b] o Hecquard y Acharya [1991] se pueden encontrar discusiones extensas acerca de métodos y algoritmos para extraer las relaciones de adyacencia implícitas en los quadtrees (2D), octrees (3D), y k–trees en general (kD).

Crecimiento guiado por modelos

Los esquemas de crecimiento de regiones vistos hasta ahora son procedimientos de propósito general, diseñados para extraer objetos homogéneos o conjuntos de regiones relacionadas de una imagen. Se empieza con un núcleo (punto, bloque pequeño de puntos, o una región uniforme) escogido por ser típico de una clase o cluster, y en etapas sucesivas se evalúan determinadas funciones de mérito para pares de fragmentos de regiones vecinos y se hace la mejor elección. Las funciones de mérito pueden depender de propiedades tales como la homogeneidad, simplicidad de forma, etc. El proceso termina cuando ya no quedan uniones aceptables que hacer.

En muchas ocasiones, una segmentación basada simplemente en estadísticas de la intensidad da como resultado un elevado número de regiones. Las razones de este problema están en el ruido de alta frecuencia y en las transiciones graduales entre valores de gris de diferentes regiones. A causa de esto, tras la segmentación inicial, las regiones pueden necesitar refinamiento o reforma. Este refinamiento puede hacerse de forma interactiva o automática. El refinamiento automático se hace usando una combinación de operaciones de división y unión. De esta forma se eliminarán falsos contornos y regiones espúreas mediante la unión de regiones adyacentes que pertenecen al mismo objeto, y añaden contornos perdidos mediante la división de regiones que contienen partes de diferentes objetos.

La combinación de algoritmos de división y unión guiada por el conocimiento sobre el dominio es adecuada para segmentar escenas complejas. Este conocimiento puede consistir en información relativa a forma, tamaño, posición espacial, etc., de los objetos de interés. Idealmente, el crecimiento de regiones debería basarse en un modelo para el objeto a extraer de la imagen, y las uniones en cada etapa deberían ser compatibles con ese modelo. Los criterios de unión de propósito general deberían usarse sólo en la ausencia de un modelo más específico.

En un crecimiento de regiones guiado por un modelo también se comienza por un núcleo altamente representativo de la clase de interés, y si se conoce algo acerca del tamaño o forma de los objetos, esto se podrá usar para guiar el crecimiento. Así, se van uniendo regiones siempre que con ello se produzca una mayor aproximación al modelo, y se finaliza el proceso cuando no hay más regiones que puedan mejorar la correspondencia (*matching*) entre el modelo y la región crecida. Dhawan y Arata [1991] describen una

técnica para segmentación de imágenes de cortes transversales de tórax obtenidas mediante CT. El sistema parte de una sobresegmentación de la imagen y una selección de núcleos de crecimiento. Estos núcleos van creciendo de forma iterativa, uniéndose a sus vecinos, hasta que una medida de distancia de momentos centrales entre los agregados de regiones y los momentos centrales del modelo empieza a crecer. De entre todos los agregados de regiones obtenidos al final, se seleccionará como reconocimiento del modelo aquella que presente una menor distancia con las características del modelo. En esta medida se ponderan distancias de segundos momentos centrales, centroide y área. La forma de describir e incorporar conocimiento en los procesos de segmentación se describe con mayor detalle en la sección 2.3.

2.2.3 Detección de bordes

Los métodos de segmentación basados en la detección de bordes utilizan también información de localización espacial, pero a diferencia de los métodos basados en regiones, no se evalúan condiciones de homogeneidad sino de cambios bruscos en la intensidad. Los bordes son características importantes en el análisis de imágenes, y su detección es, a menudo, el primer paso en la extracción de información de las imágenes. Un borde se puede definir como un cambio local significativo en la intensidad de la imagen, usualmente asociado con una discontinuidad, ya sea en la intensidad o en su primera derivada. El conjunto de bordes producidos por un detector puede particionarse en dos subconjuntos: bordes correctos, que corresponden a bordes reales presentes en la escena, y bordes falsos, los cuales no tienen correspondencia con bordes reales de la escena. Un tercer conjunto englobaría a aquellos bordes reales que no han sido detectados. Se habla de falsos positivos cuando se hace referencia a bordes obtenidos no presentes en la escena real y falsos negativos en referencia a aquellos que se han perdido y no han sido detectados.

La detección de bordes es la operación de detectar cambios de intensidad significativos en una imagen. En una dimensión, un borde se asocia con un pico en la primera derivada. El gradiente es la medida de cambio en una función, y una imagen se puede considerar como un array de muestras de alguna función de intensidad. Por analogía, los cambios significativos en los valores de gris se pueden detectar usando una aproximación discreta del gradiente. Los algoritmos para la detección de bordes constan generalmente de tres pasos: filtrado, realce y detección. Puesto que el gradiente se basa en valores de

intensidad, es muy susceptible al ruido y otros problemas asociados a la aproximación discreta. Por consiguiente, se suele usar un primer proceso de filtrado tendente a mejorar el rendimiento del detector de bordes frente al ruido. Sin embargo, existe un compromiso entre reducción del ruido y fortaleza del borde. Con el fin de facilitar la detección de los bordes, es esencial determinar los cambios de intensidad en el entorno de un punto. El realce enfatiza los pixels donde existe un cambio significativo en los valores de intensidad local y se realiza calculando magnitudes tales como el gradiente. Al final sólo interesan puntos que contienen bordes fuertes. Por tanto, se debe utilizar algún método para determinar qué puntos son borde; de entre estos métodos, el de la umbralización es el más simple. Es importante hacer notar que la detección indica, simplemente, que un borde está presente en la proximidad de un pixel, pero no proporciona, necesariamente, una adecuada estimación de la localización o de la orientación. Por esta razón muchos algoritmos para detección de bordes incluyen un cuarto paso, el de localización, que pretende ajustar la posición real del borde en la escena.

Hay dos clases principales de operadores para la detección diferencial de bordes: operadores de la primera derivada y operadores de la segunda derivada. Para los de la primera clase se efectúa de alguna forma una diferenciación espacial de primer orden, y el resultado se compara con un valor umbral. Se considera que existe un borde si el gradiente excede un umbral. Para la segunda clase, se considera la existencia de un borde cuando hay un cambio espacial significativo en la polaridad de la segunda derivada. Otros métodos tratan primero de aproximar la imagen mediante funciones simples, y es sobre estas funciones sobre las que determinan la presencia de transiciones significativas. En los siguientes apartados se discuten estos métodos, que, en general, determinarán puntos sobre al imagen en los cuales tienen lugar transiciones importantes, pero que no proporcionan una segmentación de la misma. Para ello es preciso realizar una agrupación de bordes que den lugar a contornos cerrados, que serán los que definan la segmentación de la imagen. Finalizaremos esta sección, precisamente, con una revisión de métodos para la obtención de contornos cerrados que particionan la imagen a partir de un mapa de bordes proporcionados por los operadores de gradiente.

Operadores de primera derivada

Hay dos métodos fundamentales para la generación de bordes por derivadas de primer orden [Pratt, 1991]. El primero implica la generación de gradientes en dos direcciones ortogonales en la imagen, mientras que el otro utiliza un conjunto de derivadas direccionales.

De entre los operadores de gradiente que utilizan dos direcciones ortogonales se pueden mencionar los de Roberts, Sobel, Prewit, Frei-Chen, etc. El operador de Roberts proporciona una aproximación simple a la magnitud del gradiente:

$$G[f[i,j]] = |f[i,j] - f[i+1,j+1]| + |f[i+1,j] - f[i,j+1]| \tag{2.34}$$

Usando máscaras de convolución, esto se convierte en:

$$G[f[i,j]] = |G_x| + |G_y| \tag{2.35}$$

donde G_x y G_y se calculan utilizando las siguientes máscaras:

$$G_x = \begin{array}{|c|c|}\hline 1 & 0 \\ \hline 0 & -1 \\ \hline \end{array} \qquad (2.36) \qquad G_y = \begin{array}{|c|c|}\hline 0 & -1 \\ \hline 1 & 0 \\ \hline \end{array} \qquad (2.37)$$

Aquí el gradiente se calcula para el punto interpolado $[i+\frac{1}{2}, j+\frac{1}{2}]$, y además presenta una alta sensibilidad al ruido. Una forma de evitar tener que calcular el gradiente para un punto interpolado entre pixels y de mitigar, en cierta medida, la influencia del ruido es usar una vencindad mayor para el cálculo del gradiente.

a_0	a_1	a_2
a_7	$[i,j]$	a_3
a_6	a_5	a_4

Tabla 2.1: Etiquetado de pixels en una vecindad 3×3 de un pixel $[i,j]$.

Considerando la distribución de pixels alrededor del pixel $[i,j]$ de la tabla 2.1, el operador de Sobel proporciona la magnitud del gradiente como:

$$M = \sqrt{s_x^2 + s_y^2} \simeq |s_x| + |s_y| \tag{2.38}$$

donde las derivadas parciales s_x y s_y se calculan como:

$$s_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6) \quad (2.39)$$

$$s_7 = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4) \quad (2.40)$$

con $c = 2$. Como cualquier operador de gradiente, s_x y s_y se pueden implementar usando máscaras de convolución:

$s_x =$

-1	0	1
-2	0	2
-1	0	1

(2.41)

$s_y =$

1	2	1
0	0	0
-1	-2	-1

(2.42)

Este operador pone énfasis en los pixels más próximos al centro de la máscara.

Una limitación común a estos métodos es su incapacidad de detectar adecuadamente bordes en entornos altamente ruidosos. El problema se puede aliviar aumentando el tamaño de las máscaras para incluir vecinos más alejados. Así surgen operadores como el de Prewitt 7×7, o los de Abdou y Macleod, que implementan funciones de peso Gaussianas en una vecindad mayor. Éstos se pueden considerar como operadores de gradiente diferencial de mayor tamaño, que efectúan una operación de suavización seguida por la operación de diferenciación. Un ejemplo bien conocido de este tipo de operadores es la derivada de la Gaussiana que describiremos más en detalle al final de esta subsección.

Con las técnicas diferenciales ortogonales discutidas previamente los gradientes se calculan en dos direcciones ortogonales, y la dirección del borde se infiere calculando el vector suma de los gradientes. Otro planteamiento consiste en calcular gradientes en un número mayor de direcciones convolucionando la imagen con la máscara correspondiente a cada dirección. En estos casos la dirección del gradiente se determina a partir de la dirección con gradiente mas alto. Prewitt, Kirsch, y Nevatia y Babu, entre otros, han desarrollado técnicas de este tipo.

Detección de bordes Gaussiana La idea esencial en la detección de bordes tipo escalón es encontrar puntos que presentan magnitudes de gradiente altas. En imágenes reales los escalones no son perfectamente agudos debido a la suavización efectuada por el filtrado pasa baja inherente a la óptica de las cámaras y al ancho de banda limitado

de la electrónica de la cámara. Las imágenes están, además, corrompidas por el ruido de la cámara. Cualquier aproximación al gradiente debe ser capaz de satisfacer dos requerimientos en conflicto: (1) la aproximación debe suprimir los efectos del ruido, y (2) la aproximación debe localizar los bordes de la forma más precisa posible. Como se ha comentado, existe un compromiso entre localización y supresión del ruido. El operador lineal que proporciona el mejor compromiso entre inmunidad al ruido y localización, al mismo tiempo que conserva las ventajas del filtrado Gaussiano, es la derivada de la Gaussiana.

Todos los operadores de gradiente mencionados previamente en esta subsección han sido derivados heurísticamente. Canny [1986] ha seguido un procedimiento analítico en el diseño de tales operadores. El detector de bordes de Canny es la primera derivada de la Gaussiana, y aproxima el operador que optimiza el producto de la razón señal ruido y la localización. El desarrollo de Canny se basa en un modelo unidimensional continuo de un borde escalón de amplitud h_E al que se le suma ruido blanco Gaussiano con desviación estándar σ_n. Se asume que la detección del borde se realiza mediante la convolución de una señal de borde unidimensional, ruidosa y continua con una función de respuesta al impulso antisimétrica $h(x)$, la cual tiene amplitud cero fuera del rango $[-W, W]$. El máximo local del gradiente convolucionado $f(x) * h(x)$ se marcará como borde. La función de respuesta al impulso se escoge para que satisfaga los siguientes tres criterios:

- *Buena detección.* Se maximiza la relación señal ruido (snr) del gradiente para obtener una baja probabilidad de fallo. La snr para el modelo es

$$snr = \frac{h_E}{\sigma_n} S(h) \tag{2.43}$$

con

$$S(h) = \frac{\int_{-W}^{0} h(x)dx}{\int_{-W}^{W} [h(x)]^2 dx} \tag{2.44}$$

- *Buena localización.* Los puntos marcados como borde deberían estar tan próximos como fuese posible al centro del borde. El factor de localización se define como:

$$LOC = \frac{h_E}{\sigma_n} L(h) \tag{2.45}$$

con

$$L(h) = \frac{h'(0)}{\int_{-W}^{W}[h'(x)]^2 dx} \tag{2.46}$$

donde h' es la derivada de $h(x)$.

- *Respuesta única.* Debería haber sólo una respuesta a cada borde real. La distancia x_m entre picos del gradiente cuando sólo está presente el ruido es una fracción k del ancho del operador W. Así

$$x_m = kW \tag{2.47}$$

Canny combinó estos tres criterios para maximizar el producto $S(h)L(h)$ sujeto a la ligadura de la ecuación 2.47. A medida que x_m se va haciendo mayor la función de Canny se aproxima a la derivada de la Gaussiana.

Una posible implementación del algoritmo de detección de bordes de Canny es el que se resume a continuación. Sea $I[i,j]$ la imagen. Sea

$$S[i,j] = G[i,j;\sigma] * I[i,j] \tag{2.48}$$

el resultado de convolucionar la imagen con un filtro de suavización Gaussiano, donde σ controla el grado de suavización.

El gradiente del array suavizado $S[i,j]$ se puede calcular usando las aproximaciones de la *las primeras diferencias* 2×2 para producir dos arrays $P[i,j]$ y $Q[i,j]$ para las derivadas parciales en x e y:

$$\begin{aligned} P[i,j] &\approx (S[i,j+1] - S[i,j] \\ &\quad + S[i+1,j+1] - S[i+1,j])/2 \end{aligned} \tag{2.49}$$

$$\begin{aligned} Q[i,j] &\approx (S[i,j] - S[i+1,j] \\ &\quad + S[i,j+1] - S[i+1,j+1])/2 \end{aligned} \tag{2.50}$$

La magnitud y orientación del gradiente se calculan como:

$$M[i,j] = \sqrt{P[i,j]^2 + Q[i,j]^2} \tag{2.51}$$

$$\theta[i,j] = \arctan(Q[i,j], P[i,j]) \tag{2.52}$$

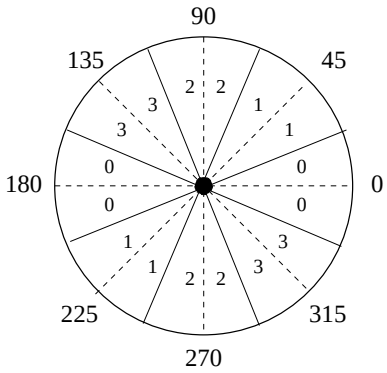


Figura 2.4: Partición de las posibles orientaciones del gradiente en sectores para supresión de no máximos.

Para identificar ahora los bordes hay que determinar los máximos locales de $M[i,j]$; para ello habrán de suprimirse todos los valores a lo largo de la línea de gradiente que no son valores de pico. El algoritmo empieza por reducir el ángulo del gradiente $\theta[i,j]$ a uno de los cuatro sectores mostrados en la figura 2.4,

$$\zeta[i,j] = Sector(\theta[i,j]). \tag{2.53}$$

Cada punto de $M[i,j]$ se compara con sus dos vecinos en un entorno 3×3 que están a lo largo de la línea de gradiente dada por el valor del sector $\zeta[i,j]$ en el centro de la ventana. Si el valor de $M[i,j]$ no es mayor que ambas magnitudes vecinas a lo largo de la línea de gradiente, entonces se pone $M[i,j]$ a cero. Este proceso proporciona bordes de anchura unidad. A pesar del suavizado efectuado como primer paso de la detección de bordes, la imagen de bordes con la supresión de los no-máximos contendrá falsos fragmentos de borde provocados por el ruido y la textura fina. Para eliminar estos falsos contornos se utiliza una doble umbralización, de tal forma que con un umbral más alto T_2 se eliminen los falsos contornos, y con un umbral menor T_1 se recuperen las discontinuidades producidas por la eliminación de contornos verdaderos en la primera umbralización.

Operadores de la segunda derivada

Los operadores de gradiente ponen puntos de contorno en aquellos pixels donde el valor del gradiente supera un umbral. En general esto resulta en la detección de un

elevado número de puntos de borde. Una solución mejor sería encontrar sólo aquellos puntos que son máximos locales en el operador gradiente. Es decir, aquellos puntos que presentan un pico en la primera derivada y un cruce por cero en la segunda derivada. Hay dos operadores en dos dimensiones que se corresponden con la segunda derivada: la Laplaciana y la segunda derivada direccional [Rosenfeld y Kak, 1982; Gonzalez y Wintz, 1987].

Operador Laplaciano La segunda derivada de un borde de escalón suavizado es una función que cruza por cero en la localización del borde, tal como ilustra la figura 2.5. La

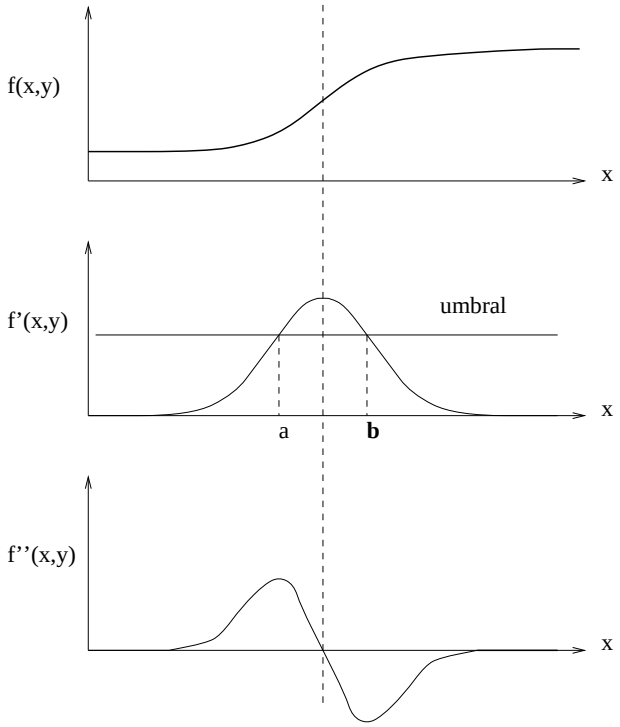


Figura 2.5: Si se indica un umbral, todos los puntos entre **a** y **b** serán marcados como pixels de borde. Mediante la segunda derivada se detecta el máximo de forma más precisa.

Laplaciana es el equivalente bidimensional de la segunda derivada. La fórmula para la Laplaciana de una función $f(x, y)$ es

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{2.54}$$

Aproximando estas derivadas parciales mediante ecuaciones de diferencias se obtiene:

$$\frac{\partial^2 f}{\partial x^2} \simeq f[i, j+1] - 2f[i, j] + f[i, j-1] \tag{2.55}$$

$$\frac{\partial^2 f}{\partial y^2} \simeq f[i+1, j] - 2f[i, j] + f[i-1, j] \tag{2.56}$$

Estas dos ecuaciones se pueden combinar en un operador simple y así obtener la siguiente máscara que aproxima la Laplaciana:

$$\nabla^2 \approx \begin{array}{|c|c|c|}\hline 0 & 1 & 0 \\ \hline 1 & -4 & 1 \\ \hline 0 & 1 & 0 \\ \hline \end{array} \tag{2.57}$$

Algunas veces es deseable dar más peso a los pixels del centro:

$$\nabla^2 \approx \begin{array}{|c|c|c|}\hline 1 & 4 & 1 \\ \hline 4 & -20 & 4 \\ \hline 1 & 4 & 1 \\ \hline \end{array} \tag{2.58}$$

El operador Laplaciano señala la presencia de un borde cuando la salida del operador realiza una transición a través de cero.

Segunda derivada direccional La segunda derivada direccional es la segunda derivada calculada en la dirección del gradiente. Este operador se implementa usando la fórmula:

$$\frac{\partial^2 f}{\partial n^2} = \frac{f_x^2 + 2 f_x f_y f_{xy} + f_y^2 f_{xy}}{f_x^2 f_y^2} \tag{2.59}$$

donde los subíndices x e y representan la derivada a lo largo de las direcciones x e y.

La Laplaciana y la segunda derivada direccional son poco usados ya que la segunda derivada se ve más afectada por el ruido que la primera derivada. Para evitar este efecto de ruido se deben usar métodos de filtrado potentes.

Laplaciana de la Gaussiana Para reducir la sensibilidad al ruido de los operadores de la 2ª derivada será necesario filtrar el ruido previo al realce de los bordes. Para hacer esto, la Laplaciana de la Gaussiana (LoG), debida a Marr y Hildredth [1980], combina el filtrado Gaussiano con la Laplaciana para la detección de bordes. En primer lugar se convoluciona la imagen con un filtro Gaussiano. Este paso suaviza la imagen y reduce el ruido. Los puntos de ruido aislado y las pequeñas estructuras se filtran. Puesto que la suavización *dispersa* los bordes, el detector de bordes considera como bordes sólo aquellos pixels que presentan máximos locales en el gradiente. Esto se realiza mediante el uso de los cruces por cero de la segunda derivada. Se usa el Laplaciano como la aproximación a la segunda derivada en 2-D porque es un operador isótropo. Para evitar la detección de bordes no significativos, sólo se seleccionan como puntos de borde aquellos cruces por cero cuya primera derivada supere un umbral.

La salida del operador LoG, $h(x, y)$, se obtiene mediante la operación de convolución

$$h(x, y) = \nabla^2[g(x, y, \sigma) * f(x, y)]. \tag{2.60}$$

Usando la regla de la derivada, para la convolución

$$h(x, y) = [\nabla^2 g(x, y, \sigma)] * f(x, y), \tag{2.61}$$

donde

$$\nabla^2 g(x, y, \sigma) = (\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4}) e^{-\frac{(x^2+y^2)}{2\sigma^2}} \tag{2.62}$$

Los filtros LoG son en general de tamaño grande, por lo que muchas implementaciones aproximan este operador por el operador diferencia de dos Gaussianas (DOG) con una relación entre sus parámetros σ de 1 a 6. Otra razón del uso del operador DOG reside en implicaciones de tipo fisiológico [Marr y Hildredth, 1980].

Las pendientes de los cruces por cero dependen del contraste del cambio en la intensidad a través del borde, por lo que hay que establecer un umbral de decisión para evitar el ruido. Además, estos cruces por cero no siempre se corresponden con posiciones enteras de pixels dentro de la imagen. Huertas y Medioni [1986] han desarrollado un método sistemático para la localización de bordes y determinación de la orientación del mismo a partir de los cruces por cero. Lu y Ramesh [1989, 1992] hicieron un estudio del

comportamiento de este tipo de detectores según el tamaño del operador y proponen un sistema de integración de los resultados obtenidos por operadores de diferentes tamaños. La conveniencia de esta integración ya había sido propuesta por Marr y Hildredth [1980] y Canny [1986]. El problema de combinar bordes obtenidos de aplicar operadores de diferentes tamaños aún está abierto.

La operación de suavización de la Gaussiana origina una pérdida de definición de los bordes y otras discontinuidades agudas en una imagen. Esta pérdida de definición depende del valor de σ. Cuanto mayor sea σ mejor es el filtrado del ruido, pero al mismo tiempo se pierde más información de bordes importantes, lo cual afecta al rendimiento del detector de bordes. Cuanto más pequeño es el filtro menor será la cantidad de ruido filtrado debido al insuficiente promediado. Con filtros grandes, los bordes que estén suficientemente próximos se unirán por la suavización y serán detectados como un único borde. En general, los filtros pequeños darán como resultado demasiados puntos ruidosos y los filtros grandes provocarán la dislocación de bordes e incluso podrán generar bordes falsos. El tamaño exacto del filtro no puede determinarse sin conocimiento del tamaño y localización de los objetos en una imagen. En general, un único filtro no puede recoger toda la información de la imagen. Para ello es necesario integrar información de filtros de diferentes tamaños. Otro ejemplo de sistema para la integración de esta información es el de Söberg y Bergholm [1988].

Se han desarrollado muchas técnicas que aplican máscaras para el filtrado de múltiples tamaños y después analizan el comportamiento de los bordes en las diferentes escalas [Lu y Ramesh, 1989, 1992]. La idea básica es explotar el hecho de que con escalas grandes se obtienen bordes robustos pero desplazados. La localización de estos bordes se puede determinar con escalas más pequeñas.

Detección de contornos en base a modelos de imagen

Una imagen es un array de muestras de una función continua. Si se puede estimar la función continua a partir de la cual se tomaron las muestras de las imágenes, entonces se pueden extraer propiedades de la imagen a partir de la función estimada. Esto podría permitir el determinar la localización de los bordes con aproximación de subpixel. Para imágenes complejas, la función de intensidad continua podría contener potencias muy altas de x e y. Esto haría extremadamente difícil la reconstrucción de la imagen original, si no

imposible. Por tanto, se podría intentar modelar la imagen como una función analítica simple a trozos. Ahora la tarea consiste en reconstruir las funciones a trozos individuales, o *facets*. Esta aproximación se llama *modelo facet*. La función de intensidad se aproxima localmente en cada pixel de la imagen. Estas funciones y no los valores de los pixels, son las que se usan para localizar bordes en la imagen.

Dependiendo de la complejidad de la imagen, se usan diferentes funciones para aproximar las intensidades. Para imágenes simples son adecuadas funciones constantes a trozos o bilineales. Sin embargo, para imágenes con regiones más complejas se usan funciones bicuadráticas, bicúbicas e incluso de mayor orden. El objetivo, ahora, es calcular los coeficientes k_i de las funciones polinómicas para cada función de aproximación, para lo cual se pueden emplear una serie de máscaras que nos permiten determinar el valor de dichos coeficientes.

Para detectar los bordes se usa el hecho de que los puntos de borde ocurren en extremos relativos de la primera derivada direccional de la función de aproximación en la vecindad del pixel. La presencia de extremos relativos en la primera derivada resultará en cruces por cero de la segunda derivada en la dirección de la primera derivada. La primera derivada en la dirección θ viene dada por

$$f_\theta^{'}(x,y) = \frac{\partial f}{\partial x}\cos\theta + 2\frac{\partial^2 f}{\partial x \partial y}\cos\theta\sin\theta + \frac{\partial f}{\partial y}\sin\theta \tag{2.63}$$

La segunda derivada direccional en la dirección θ viene dada por

$$f_\theta^{''}(x,y) = \frac{\partial^2 f}{\partial x^2}\cos^2\theta + \frac{\partial^2 f}{\partial y^2}\sin^2\theta \tag{2.64}$$

Puesto que se están considerando únicamente puntos en la dirección θ, las coordenadas de los puntos dentro de cada ventana se pueden expresar como $x_0 = \rho\cos\theta$ y $y_0 = \rho\sin\theta$, con respecto al centro de la ventana. De esta forma tendremos

$$f_\theta^{''}(x_0,y_0) = A\rho + B \tag{2.65}$$

Así, habrá un borde en la imagen en (x_0, y_0) si para algún ρ, $|\rho| < \rho_0$, donde ρ_0 es la mitad de la longitud del lado de la ventana,

$$f_\theta^{''}(x_0,y_0;\rho) = 0 \quad \text{y} \tag{2.66}$$

$$f_\theta^{'}(x_0,y_0;\rho) \neq 0 \tag{2.67}$$

En otras palabras, sólo se marca el pixel como borde si la localización del borde cae dentro de la vecindad del pixel.

Detección de contornos

Una imagen con sólo elementos de borde desconectados está relativamente sin *caracterizar*; se requiere un posterior proceso de agrupación de elementos de borde en estructuras mejor adecuadas para el proceso de interpretación. El objetivo de las técnicas que se relatan a continuación es realizar un nivel de segmentación, es decir, obtener una característica unidimensional coherente, *contorno*, a partir de muchos elementos de borde locales individuales. Esta característica podría corresponder al contorno de un objeto o a cualquier contorno significativo entre entidades de la escena. Se han desarrollado muchos algoritmos para la aproximación por curvas simples para objetos simples y estructurados [Duda y Hurt, 1973; Pavlidis, 1977]. Sin embargo, surgen dificultades cuando se trata de aplicarlos a imágenes con mapas de bordes muy densos. Los métodos indicados a continuación se ordenan de acuerdo con la cantidad de conocimiento incorporado a la operación de agrupación. Conocimiento significa de nuevo ligaduras implícitas o explícitas sobre la probabilidad de realizar una agrupación dada. Tales ligaduras pueden surgir de argumentos físicos generales o de restricciones más fuertes provenientes de consideraciones dependientes del dominio. Mucho conocimiento implica que la forma global del contorno y su relación con otras estructuras de la imagen es muy restringida. Poco conocimiento a priori significa que la segmentación debe proceder más en la base de pistas locales y suposiciones generales, con pocas expectativas y ligaduras sobre el contorno final resultante.

Búsqueda alrededor de una localización aproximada Si existe conocimiento a priori sobre una localización aproximada del contorno, éste puede usarse para guiar el esfuerzo para refinar ese contorno. En este caso las búsquedas locales se realizan a intervalos regulares a lo largo de direcciones perpendiculares al contorno aproximado. Se aplica un operador de borde a cada uno de los puntos discretos sobre estas direcciones perpendiculares. Para cada dirección se selecciona el borde con mayor magnitud de entre aquellos cuyas orientaciones sean casi paralelas a la tangente al punto en la proximidad de un contorno a priori. Si se encuentran los suficientes elementos, sus localizaciones pueden ser aproximadas por una curva analítica, tal como un polinomio de grado bajo, que se

convierta en la representación del contorno.

Método de Hough Este método [Hough, 1962] es aplicable a casos donde se sabe poco a cerca de la localización del contorno, pero su forma puede ser descrita como una forma paramétrica. Sus principales ventajas son que se ve relativamente poco afectado por las discontinuidades (*gaps*) en la curva y por el ruido. Esta técnica se aplica a cualquier curva $f(\vec{x}, \vec{a})$ donde $\vec{a}$ es un vector de parámetros. Para introducir el método se puede considerar el problema de detectar líneas rectas $y = mx + c$ en una imagen [Duda y Hart, 1972]. Para este caso el algoritmo procedería de la siguiente forma:

1. Se cuantiza el espacio de parámetros entre valores de máximo y mínimo apropiados para cada uno de ellos, en este caso $\vec{a} = (c, m)$.
2. Se crea un array acumulador, $A(c, m)$ inicializado a cero.
3. Para cada punto $\vec{x} = (x, y)$ de una imagen de gradientes, tal que su magnitud supere un umbral, se incrementan todos los puntos del array acumulador que satisfacen $f(\vec{x}, \vec{a}) = y - mx + c = 0$.
4. Los máximos locales del acumulador se corresponden con los puntos sobre la misma curva. Los valores del acumulador proporcionan una medida del número de puntos sobre la curva.

El tiempo de cálculo y el tamaño del acumulador crecen exponencialmente con el número de parámetros, por lo que esta técnica sólo es práctica para curvas con un número de parámetros pequeño.

La transformada de Hough se puede generalizar a casos de curvas de forma analítica complicada, pero con silueta particular [Ballard y Brown, 1982]. Para ello se toma un punto de referencia sobre la silueta y se trazan radios hacia el contorno. En estos puntos del contorno (x, y, α) se determina la dirección del gradiente y se almacena el punto de referencia como función de esta orientación. Así, es posible precalcular la localización de los puntos de referencia a partir de los puntos del contorno, dado el ángulo del gradiente. El conjunto de tales localizaciones, indexada por el ángulo del gradiente, da lugar a una tabla. Las coordenadas del punto de referencia son, por tanto, los únicos parámetros necesarios (asumiendo rotación y escalado fijos). Así, un punto de borde

(x,y) con orientación de gradiente ϕ restringe los posibles puntos de referencia a estar en $\{x + r_i(\phi)\cos[\alpha_i(\phi), y + r_i(\phi)\sin[\alpha_i(\phi)]\}$, donde el subíndice i hace referencia a las distintas localizaciones sobre el borde que presentan el mismo valor de ϕ. Determinados los parámetros y la función a ajustar, el algoritmo es análogo al anterior.

Ejemplos de la aplicación del método de Hough lo encontramos en el trabajo de Lai y Chin [1994] como etapa inicial de localización y aproximación de contornos previa a la aplicación de modelos deformables, y en el trabajo de Kübler y Gerig [1990] para la agrupación de segmentos de contornos en el siguimiento de secuencias temporales 2D.

Búsqueda en grafos Un grafo es un objeto general que consiste en un conjunto de nodos $\{n_i\}$ y arcos entre dichos nodos $< n_i, n_j >$. Estos arcos tienen un peso o coste asociado. La búsqueda del contorno de un objeto es la búsqueda del camino con el coste más bajo entre dos nodos del grafo ponderado. Asumiendo que se aplica un operador de gradiente a la imagen de niveles de gris para obtener una imagen de magnitudes $s(x)$ y otra de direcciones $\theta(x)$, se interpretarán los elementos de $\theta(x)$ como nodos de un grafo, cada uno con factor de peso $s(x)$. Los nodos x_i y x_j tienen arcos entre ellos si las direcciones del contorno $\theta(x_i)$ y $\theta(x_j)$ están adecuadamente alineados con el arco dirigido en el mismo sentido como dirección del contorno. Para que un arco se conecte de x_i a x_j, x_j debe ser uno de los 3 8-vecinos en frente a la dirección del contorno $\theta(x_i)$, y además $s(x_i) > T, \quad s(x_j) > T$, donde T es un valor umbral de magnitud de gradiente y $|\{[\theta(x_i) - \theta(x_j)]mod2\pi\}| < \pi/2$. Para generar un camino en el grafo de x_A a x_B se puede aplicar la técnica de búsqueda heurística. Para hacer esta búsqueda es preciso disponer de una adecuada función de evaluación $f(x_i)$ que sea una estimación del coste del camino óptimo de x_A a x_B restringido a pasar a través de x_j. Un ejemplo de aplicación de la búsqueda en grafos a la segmentación de imágenes lo constituye el trabajo de Wu y Leahy [1993].

Programación dinámica La programación dinámica [Bellman y Dreyfus, 1962] es una técnica para resolver problemas de optimización cuando no todas las variables de la función de evaluación están interrelacionadas simultáneamente. Considérese el problema de obtener el máximo de una función h de varias variables, por ejemplo 4, $\max_{x_i} h(x_1, x_2, x_3, x_4)$. Si no se conoce nada sobre h, la única técnica que garantiza un máximo global es la exhaustiva enumeración de todas las combinaciones de valores discretos de $x_1, \ldots, x_4$.

Supongamos que la función h se puede descomponer de la siguiente forma

$$h(\cdot) = h_1(x_1, x_2) + h_2(x_2, x_3) + h_3(x_3, x_4), \tag{2.68}$$

de tal modo que x_1 sólo depende de x_2 en h_1. La programación dinámica plantea el problema de obtención del máximo del modo que se indica a continuación. Maximícese sobre x_1 en h_1 y tabúlese el mejor valor de $h_1(x_1, x_2)$ para cada x_2:

$$f_1(x_2) = max_{x_1} h_1(x_1, x_2) \tag{2.69}$$

Puesto que los valores de h_2 y h_3 no dependen de x_1, no se necesita considerar a estos en este punto. Continuando de esta forma y eliminando x_2 mediante el cálculo de $f_2(x_3)$ como:

$$f_2(x_3) = \max_{x_2}[f_1(x_2) + h_2(x_2, x_3)], \quad y \tag{2.70}$$

$$f_3(x_4) = \max_{x_3}[f_2(x_3) + h_3(x_3, x_4)] \tag{2.71}$$

tal que finalmente $\max_{x_i} h = \max_{x_4} f_3(x_4)$. Generalizando el problema para N variables, donde $f_0(x_1) = 0$,

$$f_{n-1}(x_n) = \max_{x_{n-1}}[f_{n-2}(x_{n-1}) + h_{n-1}(x_{n-1}, x_n)] \tag{2.72}$$

$$\max_{x_i} h(x_1, \ldots, x_N) = max_{x_N} f_{N-1}(x_N) \tag{2.73}$$

Esta fórmula es válida si el problema se puede descomponer en la forma de la ecuación 2.68. Para formular el procedimiento de seguimiento de contorno como programación dinámica es preciso definir una función de evaluación que involucre el concepto de *mejor contorno*. Esta función ha de definirse, por lo tanto, en base al conocimiento sobre el tipo de escenas concreto con el que se trate. Un posible criterio de *buen contorno* es la suma ponderada de fuerza de borde alta y baja curvatura acumulada [Ballard y Brown, 1982]. Para una curva de n segmentos,

$$h(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{n} s(x_k) + \alpha \sum_{k=1}^{n-1} q(x_k, x_{k-1}) \tag{2.74}$$

donde la ligadura implícita es que los sucesivos x_k's deben ser vecinos adyacentes:

$$||x - K - x_{k+1}|| \leq \sqrt{2} \tag{2.75}$$

$$q(x_k, x_{k+1}) = diff[\phi(x_k), \phi(x_{k+1})] \tag{2.76}$$

donde α es negativo. Un ejemplo de la aplicación de estas técnicas a la segmentación se encuentra en el trabajo de Geiger y col [1995].

Seguimiento de contornos Si no se conoce nada a cerca de la forma del contorno, una forma de recuperarlo es mediante el seguimiento del contorno a partir de un punto inicial sobre el que hay certeza de su pertenencia a dicho contorno. Los detalles de la operación de seguimiento varían de tarea a tarea, pero se puede presentar una generalización del método que emplee gradientes locales de magnitud $s(x)$ y dirección $\theta(x)$ asociados a cada punto x. θ apunta en la dirección del máximo cambio. Si x está sobre el contorno de un objeto de la imagen, los puntos vecinos sobre el contorno deberán estar dentro de las direcciones comprendidas en $\theta(x) \pm \pi/2$. Un procedimiento representativo podría ser el siguiente:

1. Asumir que se ha detectado un punto de borde en x_i. Moverse al punto x_j adyacente a x_i en la dirección perpendicular al gradiente de x_i. Si la magnitud del gradiente en x_j es mayor que un umbral, añadir este punto al contorno.

2. En otro caso, calcular el nivel de gris promedio en el entorno 3×3 de x_j, compararlo con un umbral adecuado, y determinar si x_j está dentro o fuera del objeto.

3. Hacer otro intento con un punto x_k adyacente a x_i en la dirección perpendicular al gradiente en x_i más o menos $\pi/4$, de acuerdo con la salida del test anterior.

Los métodos de seguimiento de contornos funcionan bien cuando la imagen presenta poco ruido y no existen bordes espúreos. Métodos de este tipo son los desarrollados por Seitz y Rüegsegger [1983] y por Roberts y Robinson [Pratt, 1991].

2.2.4 Modelos deformables

Los modelos deformables son una alternativa a la segmentación clásica basada en bordes en la que se conjuga la detección de transiciones de niveles de gris en la imagen, utilizando información local, con la necesidad de obtener pixels de borde conectados formando contornos cerrados, para lo cual se hace uso también de información de tipo global como, por ejemplo, la forma esperada del contorno. Los contornos deformables o *snakes* son contornos planos deformables que se mueven bajo la influencia de las fuerzas actuantes en una imagen [Kass, Witkin y Terzopoulos, 1988]. Una *snake* es un contorno deformable que tiende a posicionarse sobre la imagen de forma que se minimice su energía, dada mediante la expresión

$$E(v) = E_{int}(v) + \mathcal{P}(v) \tag{2.77}$$

donde $E_{int}(v)$ es una energía de deformación, v representa el contorno como una proyección del dominio paramétrico unitario $s \in [0,1]$ en el plano de la imagen $\mathcal{R}^2$. Las componentes de la proyección $v(s) = (x(s), y(s))$ son funciones de las coordenadas del contorno. Por otra parte, es conveniente pensar en las fuerzas externas sobre el contorno como derivadas de un potencial $\mathcal{P}$. Para una simple *snake*, la energía interna de deformación, en su expresión clásica [Kass y col, 1988; Cohen y Cohen, 1993; Radeva y Serrat, 1993; Zhu y Yang, 1997] viene dada por:

$$E_{int}(v(s)) = \int_0^1 \frac{\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2}{2} ds \tag{2.78}$$

Con esta expresión, la energía interna, E_{int}, viene dada por la suma de la energía de membrana, que expresa el *estiramiento* de la *snake*, y la energía de deformación plana, que expresa la flexión de la *snake*. Como hemos indicado, $v(s) = (x(s), y(s))$ es la curva de la *snake*, s es la longitud de arco de la curva, y $v_s(s)$ y $v_{ss}(s)$ indican la primera y la segunda derivada respectivamente. Las funciones paramétricas de elasticidad $\alpha(s)$ y $\beta(s)$ controlan la suavidad de la curva. Estas funciones son útiles para manipular el comportamiento físico y la continuidad local del modelo. En particular, haciendo $\alpha(s_0) = \beta(s_0) = 0$ se permite una discontinuidad en la posición y haciendo $\beta(s_0) = 0$ se permite la ocurrencia de una discontinuidad tangente en s_0.

Sobre una imagen concreta $I(x, y)$, la *snake* también está sometido a los potenciales externos $\mathcal{P}$, los cuales atraen a la *snake* hacia extremos de intensidad, bordes, u otras

características de interés más complejas, dependiendo de la aplicación. Se define la energía externa como

$$E_{ext}(v(s)) = \mathcal{P}(v(s)) = \int_0^1 P(u(s))ds \tag{2.79}$$

donde la integral representa la evaluación del potencial a lo largo de la curva que define el contorno. Las fuerzas externas se definen a partir de funciones potencial cuyo dominio se corresponde con todo el plano de la imagen, y son calculadas típicamente mediante el procesado de imágenes. Es natural interpretar los mínimos locales de $\mathcal{P}$ como atractores de la *snake*. La *snake* tendrá afinidad por intensidades bajas o altas si $\mathcal{P}(x,y) = \pm\gamma I(x,y)$, dependiendo del signo, y por puntos de fuerte gradiente de intensidad si se incluye en ella el término $-\zeta|\nabla(G_\sigma(v(s)) * I(v(s)))|$. La operación $G_\sigma * I$ denota la imagen convolucionada con un filtro de suavización Gaussiano cuyo ancho característico σ controla la extensión espacial de la *depresión* atractiva de $\mathcal{P}$ [Blake y Yuille, 1993]. Otro tipo de fuerza externa común en las implementaciones de *snake*s es la provocada por los puntos de borde $-\eta e^{-d(v(s))^2}$, donde $d(v(s))$ denota la distancia entre el punto $v(s)$ y el borde más próximo [Cohen y Cohen, 1993]. γ, η, ζ son constantes positivas que ponderan las contribuciones de los diferentes términos en la expresión total de la energía externa.

Una vez que hemos planteado la idea general de los contornos deformables, expresaremos en los siguientes apartados los fundamentos matemáticos de los modelos deformables como una confluencia de geometría, física y teoría de la aproximación. La geometría sirve para representar la forma del objeto, la física impone ligaduras sobre como puede variar la forma sobre el espacio y el tiempo, y la teoría de la aproximación proporciona los mecanismos formales para ajustar los modelos a los datos medidos.

Dinámica Lagrangiana

Es natural ver el proceso de minimización de la energía como un problema estático. Sin embargo, de forma más general, se puede construir un sistema dinámico que alcance un estado de mínima energía como estado de equilibrio. Se pueden construir sistemas dinámicos con el atractivo de la semejanza con un comportamiento físico aplicando los principios de la mecánica Lagrangiana. Este planteamiento conduce a una variedad de posibles comportamientos interesantes que no son evidentes desde el punto de vista de la minimización estática. Por ejemplo, un modelo dinámico puede guiarse mediante un

sistema de control adaptativo o interactuar con el usuario cuando minimiza su energía [Terzopoulos, 1987; Terzopoulos, Witkin y Kasss, 1988; Metaxas y Terzopoulos, 1993; Metaxas y Kakadiaris, 1996; McInerney y Terzopoulos, 1996].

Para representar el comportamiento dinámico de la *snake* se puede introducir una proyección dependiente del tiempo $v(s,t)$ y una energía cinética $\int_0^1 \mu|v_t|^2 ds$, donde $\mu(s)$ es la densidad de masa, y el subíndice t denota la derivada temporal. Combinando energía cinética y el funcional de energía potencial de deformación se define el Lagrangiano

$$\mathcal{L}(v) = \frac{1}{2}\int_0^1 \mu|v_t|^2 ds - \frac{1}{2}E(v) \tag{2.80}$$

Si las configuraciones inicial y final son $v(s,t_0)$ y $v(s,t_1)$, el principio de Hamilton nos dice que el movimiento del modelo deformable $v(s,t)$ desde $t=t_0$ a $t=t_1$ es tal que la integral $\int_{t_0}^{t_1} \mathcal{L}(v)dt$ es estacionaria, lo cual implica que su variación con respecto a v es nula:

$$\delta_v\Big(\frac{1}{2}\int_{t_0}^{t_1}\int_0^1 (\mu|v_t|^2 - \alpha(s)|v_s(s)|^2 - \beta(s)|v_{ss}(s)|^2 - \mathcal{P}(v))\mathrm{ds\ dt}\Big) = 0 \tag{2.81}$$

La condición conduce a las ecuaciones de Lagrange del movimiento para el modelo. Una vez puesta en movimiento, una *snake* con una distribución de masa se movería perpetuamente, a menos que la energía cinética se disipase. Para amortiguar la *snake* de tal forma que se pueda alcanzar el equilibrio estático se incorpora el funcional de disipación de Rayleigh $\mathcal{D}(v_t) = \frac{1}{2}\int_0^1 \gamma|v_t|^2 ds$, donde $\gamma(s)$ es la densidad de amortiguamiento. Evaluando las apropiadas derivadas direccionales de los integrados $\mathcal{L}$ y $\mathcal{D}$, las ecuaciones de Lagrange son

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial v_t}\Big) + \frac{\partial D}{\partial v_t} - \frac{\partial L}{\partial v} + \frac{\partial}{\partial s}\Big(\frac{\partial L}{\partial v_s}\Big) - \frac{\partial}{\partial s^2}\Big(\frac{\partial L}{\partial v_{ss}}\Big) = 0 \tag{2.82}$$

Asumiendo densidad de masa constante $\mu(s) = \mu$ y disipación constante $\gamma(s) = \gamma$, las ecuaciones de movimiento de la *snake* se pueden escribir como

$$\mu v_{tt} + \gamma v_t + \frac{\partial}{\partial s}(\alpha v_s) + \frac{\partial^2}{\partial s^2}(\beta v_{ss}) = -\nabla P(v(s,t)) \tag{2.83}$$

con las apropiadas condiciones iniciales y de contorno. Esto puede interpretarse como una relación de balance de fuerzas. En la parte izquierda están las fuerzas inerciales, de

amortiguamiento, de tensión y de flexión. Estas fuerzas balancean el gradiente negativo del potencial del lado derecho de la ecuación, el cual puede interpretarse físicamente como las fuerzas externas generalizadas que acoplan la *snake* a los datos de la imagen.

Discretización y simulación numérica

Para calcular de forma numérica la solución correspondiente al mínimo de energía es necesario discretizar la energía $E(v)$. Un esquema general para abordar la discretización de la energía $E(v)$ es representar la función de interés v de una forma aproximada como superposición lineal de funciones base ponderadas mediante *variables nodales* u_i. Las funciones base polinómicas de soporte local consideradas en el método de los elementos finitos [Zienkiewicz y Tailor, 1989], son convenientes para la mayoría de las aplicaciones [Cohen y Cohen, 1993]. Una alternativa al método de los elementos finitos es aplicar el método de las diferencias finitas a las ecuaciones de Euler continuas, tales como las representadas en la ec. 2.83, asociadas con el modelo, [Press y col., 1992]. Los splines geométricos, [Farin, 1993], se utilizan también con frecuencia, pricipalmente en aplicaciones 3D [Sandor y Leahy, 1997; Radeva y col., 1997]. Estos métodos de representación utilizan funciones base de soporte local. Se utilizan también funciones base de soporte global como las bases de Fourier, [Ballard y Brown, 1982], y las representaciones paramétricas de curvas y superficies cuadráticas, sobre todo cuando la forma buscada está bien definida y admite pocas variaciones, [Staib y Duncan, 1992; Storvik, 1994; Bardinet y col., 1996]. Otra forma de representación global de la forma es mediante el modelo de forma activa (ASM) desarrollado por Cootes y col. [1992], donde el modelo es entrenado estadísticamente sobre un conjunto de ejemplos. En el ASM, la forma se representa como una superposición de modos de deformación ortogonales.

La forma discreta de energías cuadráticas tales como las ecs. (2.77) puede escribirse como

$$E(u) = \frac{1}{2} u^T K u + P(u), \tag{2.84}$$

donde K es la llamada *función de rigidez*, y $P(u)$ es la versión discreta del potencial externo. La solución de mínima energía (equilibrio) se puede alcanzar haciendo que el gradiente de la ecuación de la energía discreta sea cero, lo cual es equivalente a resolver

el conjunto de ecuaciones algebraicas

$$KU = \nabla P = f \tag{2.85}$$

donde f es el vector de fuerzas generalizadas externas.

Tanto los elementos finitos como las diferencias finitas generan discretizaciones locales del modelo continuo de las *snake*s, por lo que la matriz de rigidez tendrá una estructura en banda y dispersa. Para ilustrar el proceso de discretización supongamos que aplicamos el método de las diferencias finitas a la discretización de la energía de la ec. 2.78 sobre un conjunto de nodos $u_i = v(ih)$ para $i = 0, \ldots, N-1$, donde $h = 1/(N-1)$. Supongamos que usamos las diferencias finitas $v_s \approx (u_{i+1} - u_i)/h$ y $v_{ss} \approx (u_{i+1} - 2u_i + u_{i-1})/h^2$. Para condiciones de contornos cíclicas (contorno cerrado), se obtiene la siguiente matriz pentadiagonal simétrica

$$K = \begin{bmatrix} a_0 & b_0 & c_0 & & & & c_{N-2} & b_{N-1} \\ b_0 & a_1 & b_1 & c_1 & & & & C_{N-1} \\ c_0 & b_1 & a_2 & b_2 & c_2 & & & \\ & c_1 & b_2 & a_3 & b_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & c_{N-5} & b_{N-4} & a_{N-3} & b_{N-3} & c_{N-3} \\ c_{N-2} & & & & c_{N-4} & b_{N-3} & a_{N-2} & b_{N-2} \\ b_{N-1} & c_{N-1} & & & & c_{N-3} & b_{N-2} & a_{N-1} \end{bmatrix} \tag{2.86}$$

donde:

$$a_i = (\alpha_{i-1} + \alpha_i)/h^2 + (\beta_{i-1} + 4\beta_i + \beta_{i+1})/h^4, \tag{2.87}$$

$$b_i = \alpha_i/h^2 - 2(\beta_i + \beta_{i+1})/h^4, \tag{2.88}$$

$$c_i = \beta_{i+1}/h^4, \tag{2.89}$$

La versión discretizada de la ec. 2.83 puede escribirse como un conjunto de ecuaciones diferenciales de segundo orden para $u(t)$:

$$M\ddot{u} + C\dot{u} + Ku = f, \tag{2.90}$$

donde M es la matriz de masa, C es la matriz de amortiguamiento; ambas son típicamente matrices en banda y dispersas. En el simple caso de una discretización mediante diferencias finitas, donde se asume que la masa está distribuida en los nodos, M y C serán matrices diagonales. Para f estática, la condición de equilibrio dinámico $\ddot{u} = \dot{u} = 0$ conduce a la solución estática de la ec. 2.85.

Para simuilar el comportamiento dinámico de la *snake* se debe integrar este sistema de ecuaciones diferenciales ordinarias a lo largo del tiempo. La bibliografía sobre elementos finitos ofrece varios métodos de integración [Blake y Yuille, 1993]. Se puede ilustrar la idea básica con un método de Euler semi-implícito que toma intervalos temporales Δt. Se reemplazan las derivadas temporalesde u con las diferencias finitas $\ddot{u} \approx (u^{(t+\Delta t)} - 2u^{(t)} + u^{(t-\Delta t)})/(\Delta t)^2$, y $\dot{u} \approx (u^{(t+\Delta t)} - u^{(t-\Delta t)})/(2\Delta t)$, donde los superíndices denotan la cantidad evaluada en los tiempos indicados entre paréntesis. Esto da la fórmula de actualización:

$$Au^{(t+\Delta t)} = b^{(t)}, \tag{2.91}$$

donde $A = M/(\Delta t)^2 + C/2\Delta t + K$ es una matriz pentadiagonal y $b = (2M/(\Delta t)^2)u^{(t)} - (M/(\Delta t)^2 + C/2\Delta t)u^{(t-1)} + f^{(t)}$.

Programación dinámica y Temple Simulado

Existen métodos de minimización alternativos para la simulación numérica de los modelos de *snakes*. Ente ellos están la programación dinámica y el Temple Simulado (*Simulated Annealing*).

Los métodos variacionales de los que hemos tratado hasta ahora no garantizan la optimación global de la solución, requieren estimaciones de derivadas de orden elevado de los datos discretos y no permiten la imposición directa y natural de ligaduras. En el caso de planteamientos variacionales, los métodos basados en la Lagrangiana pueden tornar un problema determinado en uno indeterminado. Además, la imposición de ligaduras requiere espacios de mayores dimensiones, puesto que ahora hay más incógnitas con las que se debe tratar y las ligaduras mismas deben ser diferenciables. Con la programación dinámica las ligaduras simplemente limitan el conjunto de soluciones admisibles y de hecho reducen la complejidad computacional. Por otra parte, el uso de esquemas variacionales precisa de la estimación de derivadas de mayor orden de los datos discretos. El cálculo

de derivadas de orden elevado de datos discretos y ruidosos conduce a la inestabilidad numérica debido a la amplificación de componentes ruidosas de alta frecuencia. En la programación dinámica el orden de las derivadas es generalmente más bajo puesto que se optimizan los funcionales directamente y no se usan las condiciones necesarias para la convergencia de la solución.

En la programación dinámica, la minimización de la energía discretizada, dada por la siguiente expresión

$$E(u) = \sum_{i=0}^{N-1} E_{int}(u_i) + E_{ext}(u_i), \tag{2.92}$$

se puede ver como un proceso de decisión multietapa discreto, [Amini y col., 1990]. Empezando por el punto inicial en el contorno, se puede tratar el problema de minimización como uno en el que en cada una de las etapas de un conjunto finito $(i_0, i_1, \ldots, i_{N-1})$, se toma una decisión de un conjunto finito de posibles decisiones. Se puede establecer una correspondencia entre la minimización de la medida total de energía (con el término de orden dos en la medida de energía interna) y el problema de minimizar una función de la forma

$$\begin{aligned} E(u_1, u_2, \ldots, u_N) &= E_1(u_1, u_2, u_3) + E_2(u_2, u_3, u_4) + \\ &\quad \cdots + E_{N-2}(u_{N-2}, u_{N-1}, u_N) \end{aligned} \tag{2.93}$$

donde

$$E_{i-1}(u_{i-1}, u_{i-2}, u_{i-3}) = E_{ext}(u_i) + E_{int}(u_{i-1}, u_i, u_{i+1}) \tag{2.94}$$

Con el fin de aplicar la programación dinámica es preciso fijar un vector de dos elementos de variables de estado, (u_{i-1}, u_i). Ahora la función de valor óptimo es una función de dos puntos adyacentes sobre el contorno y se puede aplicar la forma estándar de la programación dinámica en la forma usual:

$$\begin{aligned} S_i(u_{i+1}, u_i) &= \min_{u_{i-1}} S_{i-1}(u_i, u_{i-1}) + \alpha|u_i - u_{i-1}|^2 \\ &\quad +\beta|u_{i-1} - 2u_i + u_{i+1}|^2 + E_{ext}(u_i) \end{aligned} \tag{2.95}$$

Inicializar temperatura T y contorno C;

Repetir

Seleccionar un nodo V_i de C.

Mover V_i a uno de sus pixels vecinos: $C \rightarrow C'$.

Repetir

Estimar la cantidad de cambio en E;

Si $\Delta E \leftarrow E(C') - E(C) < 0$ aceptar cambio para V_i;

en otro caso $P = \min\{1, e^{-\Delta E/T}\}$;

si $Randon[0,1) < P$ entonces $C \leftarrow C'$.

Hasta alcanzar equilibrio.

$T = f(T)$, f monótona decreciente.

Hasta $(T \rightarrow 0)$.

Tabla 2.2: Algoritmo de Temple Simulado para minimización del funcional de energía de un contorno activo.

Aunque el método de la programación dinámica presenta ventajas importantes frentes a los métodos variacionales, aún sigue presentando alta sensibilidad a los mínimos locales del funcional de energía. Por consiguiente, en general, se necesita partir de un contorno inicial próximo al final deseado para obtener buenos resultados. Como ejemplos de la aplicación de la programación dinámica a la minimización de energía en *snake*s se pueden citar, entre otros, los trabajos de Geiger y col. [1995] y Tagare [1997].

Entre los algoritmos de minimización que buscan mínimos globales está el de Temple Simulado (tabla 2.2), que es más robusto frente a los mínimos locales, aunque como contrapartida tenga asociado un mayor conste computacional. El uso de técnicas de muestreo para la simulación y la optimización en el análisis de imágenes fue introducido en Geman y Geman [1984]. El muestreo directo no es generalmente posible debido a la naturaleza compleja de los modelos involucrados, por lo que, algoritmos iterativos tales como el de Metropolis y el muestreador de Gibbs han sido ampliamente considerados en la bibliografía de análisis de imágenes. Storvik [1994] propuso un método para la

detección de curvas basado en un planteamiento completamente Bayesiano. El algoritmo de Temple Simulado fue usado para encontrar el mínimo global. Este algoritmo hace que el rendimiento del sistema sea menos sensible a las condiciones iniciales.

El algoritmo de Temple Simulado permite que el modelo deformable no quede atrapado en mínimos locales. Cuando se trabaja con modelos deformables no se partirá, en general, de un contorno próximo al deseado. Además, en las imágenes que presentan el problema de baja relación señal/ruido existirá un elevado número de mínimos locales. Estos inconvenientes pueden mitigarse mediante el uso de algoritmos que permitan cierto número de evoluciones aleatorias del contorno que aumenten la energía del sistema. De esta forma si se trata de un mínimo local poco pronunciado, se superará, mientras que si se trata de un mínimo de energía significativo, se volverá a caer en él.

El inconveniente relacionado con el algoritmo de Temple Simulado es su coste computacional elevado. En principio, se podría pensar en parametrizar la superficie utilizando modelos deformables paramétricos dentro de un planteamiento Bayesiano. Sin embargo, cuando se trata de contornos muy irregulares, como las que se pueden encontrar en imágenes médicas, se necesitaría tal número de parámetros que resultará preferible utilizar la descripción mediante puntos de superficie, la cual es capaz de manejar un amplio espectro de formas con un número limitado de parámetros. El coste computacional se puede reducir dividiendo el proceso de minimización en dos fases: en la primera sólo permitiendo desplazamientos que provoquen deflación o inflación del contorno, y en la segunda etapa sólo se permiten desplazamientos que disminuyen la energía. La restricción impuesta en la primera fase tiene el efecto positivo de impedir que el contorno quede *atrapado* por otra estructura próxima, haciendo más robusto el algoritmo a los mínimos locales. La restricción de la segunda fase impide además que el modelo deformable pueda entrar dentro del objeto y quedar atrapado en su interior, debido a un salto aleatorio de vértices de la aproximación poligonal. Para conseguir esto se hace que la temperatura T del sistema decrezca rápidamente al entrar el sistema en la segunda etapa. De esta forma el sistema evolucionará en las proximidades del punto de congelación.

Modelos deformables probabilísticos

Una visión alternativa de los modelos deformables consiste en considerar el proceso de ajuste del modelo como un problema de estimación probabilística. Esto permite la incor-

poración de las características de un modelo a priori y de un modelo sensor en términos de distribuciones de probabilidad. El esquema probabilístico también proporciona una medida de la incertidumbre en los parámetros de la forma estimada después de que el modelo se ajuste a los datos de la imagen.

Sea u el vector de parámetros de forma del modelo deformable con una probabilidad a priori $p(u)$ sobre los parámetros. Sea $p(I|u)$ el modelo de imagen (sensor), la probabilidad de producir una imagen I dado un modelo u. De acuerdo con el teorema de Bayes,

$$p(u|I) = \frac{p(I|u)p(u)}{p(I)} \tag{2.96}$$

la probabilidad a posteriori $p(u|I)$ de un modelo, dada la imagen, se puede expresar en términos del modelo de imagen y las probabilidades a priori de modelo y de imagen.

Es fácil convertir la medida de energía interna, ec. 2.78 del modelo defomable, en una distribución a priori sobre las formas esperadas, siendo las formas con menor energía las más probables. Esto se consigue empleando una distribución de Gibbs de la forma

$$p(u) = \frac{1}{Z_s} exp(-E_{int}(u)), \tag{2.97}$$

donde $E_{int}(u)$ es la versión discretizada de $E_{int}(v)$ en la ec. 2.78 y Z_s es una constante de normalización llamada constante de partición. Este modelo a priori se combina con un modelo sensor basado en medidas lineales con ruido Gaussiano

$$p(I|u) = \frac{1}{Z_s} exp(-P(u)), \tag{2.98}$$

donde $P(u)$ es una versión discreta del potencial $\mathcal{P}(v)$ en la ec. 2.78, el cual es una función de la imagen $I(x, y)$. Los modelos se ajustan buscando el vector u que maximice localmente $p(u|I)$. Ésta se conoce como la solución de *máximo a posteriori* [Geman y Geman, 1984; Chakraborty y col., 1994, 1996]. Con este planteamiento se obtiene la misma solución que con la minimización de la ec. 2.77.

2.2.5 Integración de homogeneidad con gradientes

Tanto las técnicas de segmentación basadas en regiones como las basadas en la detección de bordes presentan un éxito limitado cuando la calidad de las imágenes es pobre.

La integración de estos métodos puede resultar potencialmente en una salida mejorada del proceso de segmentación. Sin embargo, la implementación de dicha integración necesita de sistemas grandes y complicados. Probablemente, la forma más fácil de hacerlo es dividir el sistema en módulos funcionales y analizarlos por separado. De esta forma, el desarrollo del complejo sistema de integración consistiría en el ensamblado de módulos pequeños bien conocidos.

Los métodos basados en regiones y los métodos basados en gradientes tienen sus ventajas y sus inconvenientes. Los métodos basados en regiones dependen de estadísticas de los pixels sobre áreas localizadas de la imagen; las estadísticas se basan en puntos y no tienen características de localidad. Las técnicas de división y unión generan una segmentación, en general, dependiente de la forma en que se crean y escogen los segmentos iniciales, y del algoritmo de crecimiento concreto. Es posible que se unan diferentes regiones debido a que el cambio de las características de los pixels en el contorno queda debilitado ante el carácter global de las regiones. Las técnicas basadas en bordes tienen frecuentemente dificultad en establecer conectividad entre diferentes segmentos de borde. Las discontinuidades ocurren debido al ruido o a cambios agudos en la dirección de los bordes. Dependiendo de la escala escogida pueden surgir interferencias entre partes de un mismo contorno o entre contornos distintos próximos, provocando la desaparición de puntos de borde verdaderos y generando otros espúreos. Mientras que los métodos basados en regiones manejan mejor los problemas de ruido y de pérdida de información de alta frecuencia, los métodos basados en gradientes son menos proclives a errores debidos al cambio en los valores de la escala de grises y pueden tratar de forma más adecuada los cambios de forma. El objetivo de las técnicas de integración es combinar lo mejor de ambos esquemas en un método integrado. En las situaciones en las que se producen conflictos se suelen invocar reglas de producción para decidir sobre la solución más adecuada.

La mayoría de las técnicas de integración de regiones y bordes se pueden clasificar en cuatro categorías: (1) técnicas que usan conocimiento de alto nivel, (2) algoritmos basados en operaciones de lógica Booleana a nivel de pixel, (3) algoritmos para modalidades de imágenes o técnicas de segmentación específicas, y (4) técnicas que usan información *a priori* y modelos probabilísticos.

Citaremos a continuación ejemplos de técnicas de integración que caen dentro de algunas de las categorías anteriores. Pavlidis y Liow [1990] partieron de una imagen sobre-segmentada mediante un método de división y unión de regiones. Fuerzan la so-

bresegmentación para evitar los problemas de bordes no detectados y así centrarse en los problemas de mejora de localización de contornos verdaderos y eliminación de los falsos. Para corregir estos últimos calculan una función de mérito de la forma

$$f_1(\text{contraste}) + \beta f_2(\text{artefactos de segmentación}). \tag{2.99}$$

Valores bajos de esta función indican candidatos a falsos contornos. El significado del primer término es obvio; el segundo se introduce para tener en cuenta las propias limitaciones del método basado en regiones. Para la mejora en la localización de los contornos verdaderos introducen la siguiente función de mérito

$$f_3(\text{contraste}) + \alpha f_4(\text{suavidad}). \tag{2.100}$$

Esta expresión se evalúa a lo largo de los puntos del contorno y en su vecindad. La curva se modificará para situarse en el máximo de la función.

Chu y Aggarwal [1993] presentaron un algoritmo para la integración de múltiples mapas de segmentación basados en regiones y en bordes. El método, a diferencia del anterior, es independiente de los algoritmos concretos. A partir de cada mapa de segmentación genera el correspondiente mapa de bordes. Sólo utiliza información aportada por los mapas de segmentación relativa a tamaño de contorno y posición y las relaciones de adyacencia de las regiones. El resultado es *negociado* más que seleccionado entre las diferentes opciones, y finalmente suavizado. La solución al problema se plantea como la minimización de una función objetivo del tipo

$$\sum_{\forall T_i \in I} C_e(T_i) + \sum_{\forall Q_j \in O} C_s(Q_j), \tag{2.101}$$

sujeto a ligaduras de conectividad de contorno, tamaño de regiones, etc. El primer término hace referencia a la negociación sobre todos los mapas de entrada y el segundo término hace referencia a la suavidad en el mapa de salida.

Haddon y Boyce [1990] unifican la información de regiones y bordes mediante un planteamiento basado en matrices de coocurrencia. Suponen que la imagen obedece a una distribución normal. Los picos en la diagonal de la matriz caracterizan a las regiones, mientras que los picos fuera de la diagonal caracterizan a los contornos. El algoritmo de

segmentación se divide en dos etapas. La segmentación inicial se basa en en la localización de cada pixel y los de su entorno dentro de la matriz de coocurrencia. En una segunda etapa se refina esta segmentación mediante etiquetado con relajación. Los coeficientes del etiquetado se derivan de la matriz de coocurrencia, asegurando así la compatibilidad entre las dos etapas. En la primera etapa se consideran propiedades globales de la imagen, determinadas por la matriz de coocurrencia, mientras que en la segunda etapa se impone la consistencia local, minimizando la entropía de la información local. La salida del algoritmo es una asignación única de cada pixel a una región, junto con la indicación de si es interior o contorno en relación con cada una de las cuatro direcciones de coocurrencia. Los bordes se definen en localizaciones interpixel siempre y cuando un par de pixels sean contornos respecto de direcciones de coocurrencia opuestas.

Chakraborty y col. [1994] unifican información de regiones y bordes en un sistema integrado en el que la optimización se realiza al mismo tiempo sobre la localización de los procesos de línea, así como en la clasificación de los pixels. El sistema se implementa como dos módulos que operan en paralelo e interactúan de modo que en cada etapa la salida de cada modelo se actualiza usando información de las salidas de ambos modelos en la iteración previa. El sistema intenta integrar la búsqueda de contornos y la segmentación basada en regiones, más que la detección de bordes y el crecimiento de regiones. De esta forma se hace uso de información de carácter global y el sistema se basa más en la forma. En el sistema de segmentación basado en regiones modela la imagen como un Campo Aleatorio de Markov (MRF) y usan un esquema de máxima probabilidad a posteriori (MAP) para hacer la clasificación [Leahy y col., 1989]. En cada iteración se calcula la probabilidad de que un pixel particular pertenezca a las diferentes clases, y al final se clasifica de acuerdo con la mayor probabilidad. El proceso termina cuando no hay cambios entre iteraciones. El sistema de segmentación basado en contornos utiliza el método de modelos deformables de Staib y col. [1992]. La curva se representa mediante una parametrización de Fourier, p, y el problema se convierte en uno de optimización para la estimación de los parámetros. Las entradas al módulo de integración las constituyen la imagen de gradientes, I_g, y la imagen clasificada de regiones, I_s. El problema de estimación de contornos se plantea usando la información de gradientes y de homogeneidad de regiones en un esquema de MAP. De esta forma se pretende maximizar $P(p, I_g, I_s)$. Se incorpora información de forma a priori mediante p, de bordes mediante I_g, y de regiones mediante I_s.

2.3 Interpretación de imágenes. Sistemas basados en conocimiento

Todos los métodos de segmentación descritos anteriormente permiten agrupar partes de una imagen formando unidades homogéneas con respecto a una o más características. Sin embargo, un sistema automático debe identificar o asignar etiquetas distintivas a los objetos existentes en la escena. Los esquemas de segmentación de bajo nivel permiten la interpretación de imágenes para casos muy simples. Este no suele ser el caso cuando se trata con imágenes reales. En general va a ser precisa una etapa de postprocesado de la información de bajo nivel, basado en el conocimiento específico sobre el domino de la aplicación. La imagen segmentada proporcionará los principios de la interpretación dependiente del dominio. La necesidad de inyectar conocimiento específico del dominio concreto de aplicación es lo que conduce a sistemas basados en conocimiento.

Los sistemas de visión basados en conocimiento usan conocimiento sobre el dominio para analizar imágenes. El conocimiento puede ser general o extremadamente específico, pasando por todos los niveles intermedios. Los sistemas de visión hacen uso de una gran variedad de representaciones del conocimiento, desde simples vectores a estructuras complejas. Comúnmente, el conocimiento experto se considera constituido por dos elementos: *conocimiento del dominio* y *conocimiento de control* [Garbay, 1995]. El primero hace referencia a aquellos elementos (hechos, propiedades, relaciones) usados para representar un caso, mientras que el segundo representa la forma de *razonar* sobre el caso. El conocimiento se adquiere generalmente a través de la interacción directa con el experto humano, pero también a través de otras fuentes, como son las constituidas por estudios previos.

En estos sistemas existen diversos niveles de procesado. El procesado de bajo nivel integra algoritmos de segmentación tales como los descritos hasta ahora. Sobre el resultado de esta segmentación se aplican procesos de medio y/o alto nivel que aplican conocimiento específico sobre el dominio para el análisis e identificación de los objetos de interés en la escena. En Rao y Jain [1988] y Crevier y Lepage [1997] se pueden encontrar dos interesantes estudios a cerca del uso y representación del conocimiento, así como de las diferentes estrategias de control empleadas en sistemas de visión.

A continuación analizaremos las características básicas de estos sistemas, como son la

representación del conocimiento y las estrategias de control. Posteriormente, tras analizar los problemas involucrados en la reconstrucción 3D, describiremos una serie de sistemas de análisis de imágenes basados en conocimiento que podemos encontrar en la bibliografía.

2.3.1 Representación del conocimiento

El *conocimiento sobre el dominio* se puede representar de dos formas diferentes dependiendo de si el énfasis se pone sobre las propiedades y la estructura o en su relaciones mutuas.

La forma más simple de representar conocimiento es mediante un *vector de características*, usado comúnmente en el reconocimiento estadístico de patrones. Un vector de características es simplemente una tupla de medidas de propiedades que tienen un valor numérico y que se puede usar para el reconocimiento de un patrón o un objeto. Así, por ejemplo, se utilizan características de momentos para el reconocimiento de formas. Las características usadas en este tipo de representación son propiedades globales del objeto que caracterizan al conjunto del mismo y no a sus partes. Este tipo de representación no permite el reconocimiento de objetos complejos. Para ello es preciso recurrir a estructuras jerárquicas y relacionales [Haralick y Shapiro, 1992].

Las escenas y los objetos complejos están compuestos, generalmente, de partes diferenciadas. Por tanto, una descripción completa de una entidad compleja contendrá propiedades globales, propiedades locales de cada una de sus partes y las relaciones entre las mismas. Este tipo de representación corresponde a las estructuras relacionales. Aunque en general es suficiente describir un objeto por sus partes y sus relaciones, las escenas y los objetos más complejos pueden tener tantas partes que la descripción se haga difícil de manejar. Si una entidad puede dividirse de forma natural en partes, cada una de estas en subpartes, y así sucesivamente, entonces las estructuras jerárquicas pueden ofrecer una mejor representación.

Cualquier método de representación cuyo principal formato para almacenar información pueda compararse con los nodos y los vínculos de un grafo, recibe el nombre de *objetos estructurados*. Los objetos estructurados mantienen información sobre los objetos y las relaciones entre ellos. Las *redes semánticas* fueron los primeros de estos métodos en desarrollarse. En el modelo de redes semánticas se representan modelos, objetos, conceptos,

situaciones y acciones mediante nodos. Las relaciones entre estos nodos se representan mediante arcos etiquetados. El principal problema de las redes semánticas es la carencia de una teoría de la semántica. No se hace distinción entre diferentes tipos de vínculos, ni entre clases de objetos e individuos.

Los marcos (*frames*) son otro esquema de representación de objetos estructurados con una agrupación más semántica de la información. Con ellos se pretende almacenar representaciones estereotípicas de diferentes situaciones, objetos y eventos. Marcos individuales corresponde a nodos en la red, y las relaciones entre marcoss son los arcos de la red. Cada marco individual almacena información sobre un objeto o clase de objetos particular. Áquel se puede ver como una estructura de datos con un conjunto de características y sus valores asociados. Los marcos se disponen en jerarquías de clases mediante relaciones definidas. La jerarquía de tipos proporciona mecanismos para la herencia. Una de las características de los sistemas de marcos es que la información en el tope de la jerarquía de clases está fija, y como tal puede proporcionar expectativas sobre ocurrencias específicas de valores en marcos individuales. Ninguna característica de un marco se deja sin asignación de valor; en todo caso, toman valores por defecto que heredarán de sus ascendentes. Hay cuatro métodos para la extracción de información de una característica de un marco. En primer lugar, si tiene asociado un valor específico, entonces se devuelve éste, asumiendo que es el valor más fiable y actualizado. Si falla ésto, el sistema puede o bien mirar en ascendentes sucesivos del marco en cuestión e intentar heredar valores, o bien el sistema puede buscar en el marco mismo, o en sus ascendentes, por procedimientos que especifiquen como calcular el valor requerido. Si nada de lo anterior funciona, entonces el sistema puede extraer el valor por defecto almacenado para el objeto prototipo en cuestión.

Los marcos son muy atractivos como método para almacenar modelos de alto nivel de estructuras anatómicas en sistemas de visión computacional. Permiten representar propiedades genéricas, relaciones entre ellas, y permiten almacenar instancias específicas de dichas estructuras dentro del mismo esquema.

No todo el conocimiento complejo es jerárquico o relacional. Los sistemas basados en conocimiento también pueden codificar éste en forma de reglas (conocimiento heurístico). Este es un conocimiento empírico y con un mecanismo de representación formalizado.

El etiquetado de una imagen implica el reconocimiento (*matching*) de características

extraidas de la imagen con modelos de los objetos a identificar en la escena. El desarrollo y uso de los modelos es crítico para la efectividad del sistema de visión. La producción de un modelo simbólico de alto nivel requiere la representación del conocimiento sobre los objetos a modelar, sus relaciones, y como y cuando usar la información almacenada dentro del modelo. Una de las formas de representar el conocimiento es mediante *reglas de producción* las cuales tienen cinco propiedades principales: (1) la incorporación de la experiencia humana en las reglas condicionales *"if ... then"*; (2) un incremento en la habilidad proporcional al aumento de la base de conocimiento; (3) la posibilidad de resolver problemas complejos mediante la selección de reglas y combinando los resultados; (4) la determinación adaptativa de la mejor secuencia de reglas a ejecutar; y (5) la explicación de sus resultados invirtiendo la línea de razonamiento. Las reglas son una buena forma de representar el conocimiento proporcionado por un número grande de hechos discretos.

2.3.2 Sistemas de control

La estrategia de control de un sistema determina como va a ser usado el conocimiento en ese sistema. Las dos formas principales de control son el jerárquico y el no-jerárquico. El *control jerárquico* es la estructura de control más común en programación. Se define por una jerarquía u ordenación predefinida de los procedimientos que ejecuta la tarea de análisis requerida. El orden más común es aquél en que un procedimiento principal llama a un conjunto de rutinas de preprocesado que convierten la imagen original en una forma más adecuada para la extracción de características primitivas. A continuación, una rutina para la extracción de características localiza las características de interés, para las que construye una descripción simbólica. Finalmente, un procedimiento de toma de decisiones realiza algún tipo de reconocimiento haciendo uso de esta descripción. Cada uno de estos procedimientos tiene su propia tarea y su propio lugar en la jerarquía. La comunicación se realiza através de los resultados intermedios.

Hay dos tipos principales de control jerárquico: ascendente (*bottom-up)* y descendente (*top-down*). En el control ascendente se extraen las características de la imagen y se agrupan de algún modo, sin conocimiento de la estructura del objeto o de la escena. Sólo usa el procedimiento de reconooimiento (*matching*) de los objetos una vez que se ha construido la descripción simbólica completa. La estrategia descendente comienza con una hipótesis de que la imagen contiene un objeto particular. Esto conduce a posteriores

hipótesis acerca de las partes del objeto. Estas hipótesis, a su vez, conducirán a ciertos tipos de primitivas en ciertas relaciones. Al final, las rutinas de extracción de propiedades se llaman o bien para verificar la existencia de una primitiva e indicar su localización o para indicar fallo.

Ni el control ascendente ni el descendente estrictos son lo suficientemente flexibles para el análisis de imágenes complejas. Una variante potente es el control híbrido con realimentación. Esta estrategia comienza con una segmentación inicial de la imagen y la extracción de conjuntos preliminares de propiedades y relaciones. Sobre la base de esta descripción prelimiar se hacen hipótesis sobre la identidad de uno o más objetos. A continuación se verifica la existencia de esos objetos mediante una estrategia descendente. El conocimiento de estos objetos verificados permite la formación de nuevas hipótesis. Además, a medida que se reconocen objetos o partes, se deduce más información que ayudará en el proceso de reconocimiento global.

Mientras el control jerárquico dicta el orden en el que operan los diferentes procesos, el control no-jerárquico permite que los datos mismos dicten el orden en el que se hacen las cosas. En el modelo de control no jerárquico, el conocimiento está embuido en un conjunto de procedimientos llamados *fuentes de conocimiento*. Cada fuente de conocimiento se puede comunicar con alguna de las restantes fuentes, de tal forma que trabajan de forma cooperativa en el análisis de la imagen. Una forma simple de implementar esta idea es el planteamiento de pizarra (*blackboard*), en el cual un conjunto de fuentes de conocimiento independientes comparten una base de datos global. Cada fuente de conocimiento chequea la pizarra para ver si se cumplen sus precondiciones. Si es así, se ejecuta y coloca sus resultados en la pizarra, donde podrán ser usados por otras fuentes de conocimiento. Un planificador de la pizarra controlará el acceso de las fuentes de conocimiento que compitan por dicho acceso. Este controlador elige la mejor acción basándose en lo que se conoce hasta el momento. La figura 2.6 muestra los esquemas de los mecanismos de control citados.

2.3.3 Mecanismos de inferencia

La inferencia es el proceso de deducir hechos a poartir de otros hechos conocidos. De entre todos los mecanismos de inferencia, los *sistemas de producción*, de *etiquetado,* y de *conocimiento activo* han sido los de mayor uso en visión.

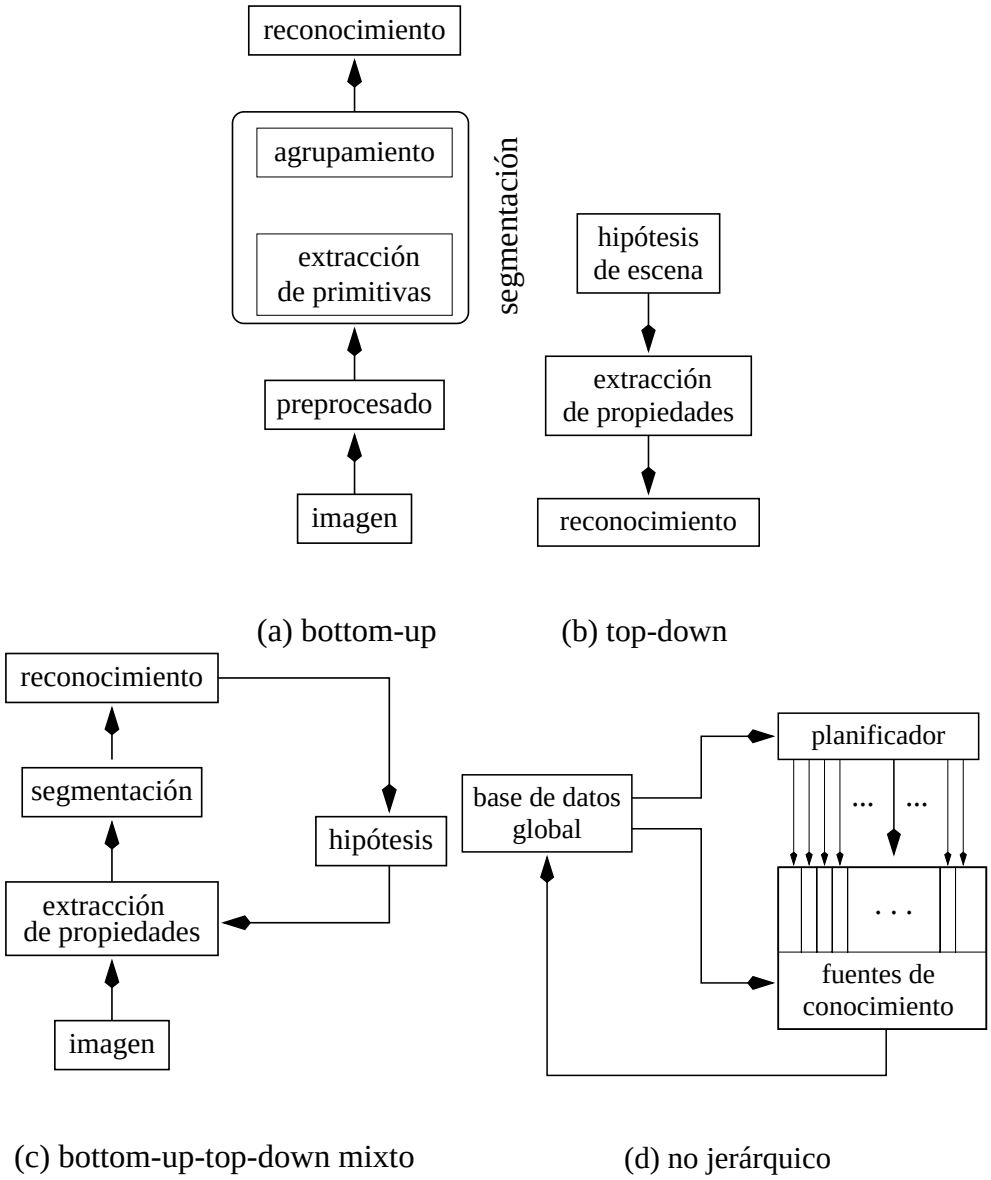


Figura 2.6: Esquemas de control.

Un **sistema de producción** consiste de un conjunto explícito de pares situación–acción, que se aplican a la base de datos de situaciones. Por ejemplo la simple regla:

```
(AltaDensidad(Region X))→ (HuesoCortical(RegionX))
```

puede inferir la interpretación de una región dada. También se pueden desarrollar reglas de segmentación; el siguiente ejemplo une dos regiones adyacentes en una única región:

```
(AltaDensidad(Región X)) AND (AltaDensidad(Región Y)) AND
(Adyacentes(Región X, Región Y))
→ (Unión(Región X, Región Y))
```

En su forma más simple un sistema de producción tiene tres componentes básicos:

1. Una base de datos.
2. Un conjunto de reglas.
3. Un intérprete para las reglas.

La base de datos de visión contiene un conjunto de hechos conocidos (valores de propiedades, relaciones) a cerca del entorno visual. Una regla es un par ordenado de patrones con una parte derecha y una parte izquierda. Las reglas se pueden aplicar de dos formas diferentes: (1) la parte izquierda identifica características en la base de datos, y la parte derecha añade las consecuencias a la base de datos; (2) la parte derecha identifica características y los antecedentes de la parte izquierda se añaden a la base de datos. En el primer caso (encadenamiento directo) el objetivo es ver si se puede derivar un conjunto de consecuencias a partir de un conjunto de hechos iniciales. En el segundo caso (encadenamiento hacia atrás), el objetivo es determinar un conjunto de hechos que hayan podido producir una consecuencia particular.

El control en un sistema de producción es simple: se aplican las reglas mientras no se alcance una determinada condición en la base de datos.

Los **esquemas de etiquetado** se diferencian de la mayoría de los mecanismos de inferencia en que involucran a procesos de optimización matemática en espacios continuos y pueden implementarse con procesamiento paralelo. La necesidad de asignar etiquetas a objetos de forma consistente se presenta en muchos contextos tales como los asociados a grafos. Todos los problemas de etiquetado tienen los siguientes elementos:

1. Un conjunto de objetos.

2. Un conjunto de relaciones entre los objetos: unarias (propiedades de los objetos) y de orden superior (relaciones geométricas y topológicas).

3. Un conjunto finito de etiquetas cada uno asociado con un significado concreto. El etiquetado asignará una o más etiquetas a cada objeto de la estructura relacional. En el último caso a cada objeto se le asignan pares (etiqueta, peso).

4. Ligaduras, las cuales determinan que etiquetas pueden ser asignadas a un objeto, y que conjuntos de etiquetas pueden se asignadas a objetos en una estructura relacional.

El problema básico consiste entonces en dada una escena de entrada, un conjunto de etiquetas y un conjunto de ligaduras, encontrar un etiquetado consistente. En problemas reales, más que buscar soluciones consistentes, lo que se hace es buscar la solución óptima.

Los **sistemas de conocimiento activo** se caracterizan por el uso de procedimientos como unidades elementales de conocimiento. Estos sistemas son paralelos y no jerárquicos. Se pueden activar simultáneamente varios procedimientos dependiendo de la entrada. Cada procedimiento maneja un trozo de conocimiento. El control está totalmente distribuido. La determinación de lo que se va a hacer a continuación está totalmente contenida en los módulos y no depende de elementos exteriores.

2.4 Reconstrucción 3D

Como resultado de la aplicación de cualquier técnica de segmentación a imágenes 3D se obtiene un conjunto de contornos definidos en diferentes planos 2D y que delimitan las estructuras de interés en ese corte del conjunto de datos 3D. El siguiente paso que hay que dar para disponer de verdadera información 3D es la creación de modelos tridimensionales de las estructuras segmentadas. La disposición de estas reconstrucciones 3D en el campo de las imágenes médicas permitirá realizar medidas volumétricas, planificación quirúrgica, diseño de prótesis, etc.

El problema de la reconstrucción del contorno de un objeto sólido a partir de un conjunto de secciones planas paralelas ha atraido la atención de muchos investigadores en las últimas décadas, motivada, fundamentalmente, por las aplicaciones en medicina, donde

se obtienen secciones transversales de órganos humanos, tales como huesos, mediante exploraciones de CT o MR. Estas secciones transversales (*slices*) son la base para la interpolación de la superficie del órgano. El objeto interpolado podrá entonces ser visualizado en aplicaciones gráficas, o usarse como modelo de entrada a otro tipo de aplicaciones tales como el análisis mediante elementos finitos, manufactura mediante máquinas controladas numéricamente, etc.

Las bases de datos 3D se reconstruyen a partir de un conjunto de secciones transversales 2D. La visualización se puede realizar de dos formas principales: basada en superficies o basada en volúmenes. La reconstrucción basada en superficies requiere un preprocesado consistente en la detección de contornos y posterior aproximación poligonal de los mismos. Las representaciones volumétricas consideran la base de datos de entrada como un array de elementos de volumen donde cada objeto se especifica como un conjunto conectado de voxels.

2.4.1 Representación de superficies

La reconstrucción de superficies se suele abordar como el apilamiento de contornos en el espacio según la disposicion longitudinal de las imágenes CT. La definición de la superficie se puede hacer mediante parches triangulares *anclados* en vértices de las aproximaciones poligonales de los diferentes contornos, o mediante definición de parcelas de superficie de orden de representación mayor.

Interpolación por triangulación

Estas técnicas se han aplicado especialmente a la representación 3D de estructuras anatómicas a partir de cortes paralelos [Barillot, 1993]. El planteamiento clásico comienza con la secuencia binaria de los datos fuente y extrae un conjunto de contornos que representan la superficie. El recubrimiento de ésta se hace juntando los diferentes puntos pertenecientes a los contornos con triángulos, con el fin de optimizar un criterio dentro de un grafo planar. Se han propuesto muchas soluciones para la interpolación *raster* pura. Éstas manejan, generalmente, dos imágenes raster donde cada pixel es o bien blanco o negro, o se le asigna un nivel de gris dentro de una escala fija. La interpolación produce una o más imágenes raster que van transformando, de forma suave, la primera imagen en

la segunda. Entonces, la superficie se determina mediante técnicas de detección de bordes que identifiquen lugares de transición del interior al exterior del objeto. Lorensen y Cline [1987] intentaron convertir directamente los datos de voxel en superficie poliédrica sugiriendo la técnica de *marching cubes*, la cual produce triángulos muy pequeños cuyo tamaño es semejante al de los voxels de entrada.

Muchas otras técnicas asumen que la interpolación viene precedida de un proceso de detección de bordes ejecutado sobre cada uno de los cortes. De esta forma, cada corte está representado por una jerarquía de contornos no intersectantes, que delimitan las áreas de interés. Cada contorno viene dado por una secuencia circular discreta de puntos, que se puede considerar como una curva cerrada poligonal cuyos vértices son dichos puntos.

Por tanto, el problema planteado es el siguiente: dada una serie de cortes paralelos planos, cada uno consistente en una colección de curvas poligonales simples, cerradas y no intesectantes, se desea construir el modelo sólido poliédrico cuyas secciones transversales a lo largo de los planos dados coincidan con los cortes dados. Una simplificación natural del problema es considerar sólo un par de cortes paralelos consecutivos y construir el modelo sólido dentro de la capa delimitada por los planos de los cortes. La unión, o concatenación, de estos modelos dará una solución al problema completo del modelado. Aunque se podría perder suavidad en este tipo de interpolación, ésto no parecer ser un problema serio con las imágenes médicas, donde dos cortes adyacentes no suelen diferir de modo considerable, por lo que las consideraciones de suavidad no parecen tener mucha influencia sobre la solución. Aquí hay que remarcar que la solución no es única, y que la medida de *bondad* de la solución propuesta es más bien subjetiva e intuitiva [Barequet y Sharir, 1996].

La mayoría de los primeros trabajos sólo estudian la variante donde cada corte sólo contiene un contorno, la cual se denota con el nombre de caso *uno-a-uno*, en oposición a los casos *uno-a-muchos* y *muchos-a-muchos*. Estos estudios buscaban la optimización de alguna función objetivo. Las primeras soluciones fueron propuestas por Keppel [1975] y por Fuchs y col. [1977]. Keppel trató el caso simple en el que cada corte contenía sólo un contorno y asumió implícitamente un grado elevado de semejanza entre ellos. Después de representar el problema usando la terminología de grafos, buscó el camino mínimo en un grafo toroidal ponderado que diese el recubrimiento de la superficie con máximo volumen. Fuchs y col. también trataron el caso simple de uno-a-uno, y usando una reducción similar al problema de grafos, obtuvieron el recubrimiento con la superficie de área más pequeña.

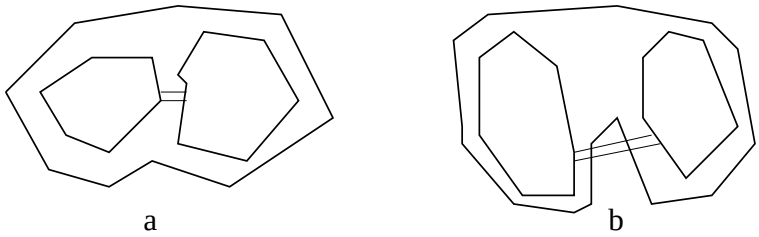


Figura 2.7: Puenteado en casos de bifurcación simple: (a) puenteado aceptable, (b) puenteado conflictivo.

Un planteamiento similar fue el adoptado por Sloan y Painter [1988], quienes sugirieron una heurística mejorada (*pessimal guesses*) para la búsqueda en grafos.

Christiansen y Sederberg [1978] empezaron con el caso simple de uno-a-uno. A diferencia de los dos primeros trabajos, desarrollaron una regla de avance local. Con el fin de manejar contornos sin semejanza aparente, sugirieron mapearlos en el cuadrado unitario primero y después construir el recubrimiento entre ellos, para, finalmente, mapearlos en sus escalas y localizaciones iniciales. Intentaron además extender el método a casos más complejos. En los casos simples de bifurcación, donde un corte contiene un contorno y el otro contiene dos, propusieron conectar los dos contornos del segundo corte por un puente de mínima longitud, de modo que se reducía el problema al caso simple de uno-a-uno. Este procedimiento falla cuando el puente es inconsistente con la geometría del otro corte, tal como ilustra la figura 2.7, por lo cual decidieron hacer el puenteado manual.

Wang y Aggarwald [1986] también trataron el caso uno-a-uno. Basándose en la cantidad de contorno solapante, deciden si proyectar un contorno de un corte sobre su contrapartida en el otro corte, o no. La triangulación que define la superficie reconstruida entre los dos contornos se obtuvo mediante un algoritmo de búsqueda heurística.

Boissonnat [1988] construyó la triangulación de Delaunay para cada corte, proyectó una triangulación sobre la otra y obtuvo una colección de tetraedros para la que intentaba maximizar la suma de sus volúmenes. Éste fue un paso considerable hacia el manejo de casos donde cada corte tiene múltiples contornos. Boissonnat mencionó tres casos típicos de fallo de su método estándar que necesitaban un tratamiento especial. El primer ejemplo contenía dos contornos solapantes, pero con diferencias considerables en su geometría. El segundo caso consistía en dos contornos similares, donde uno de ellos contenía un agujero.

El último ejemplo correspondía a un problema de bifurcación, pero sin solape de contornos. Boissonnat sugirió un esquema de corrección en el cual o se cambiaba la geometría de uno de los cortes o bien se construían cortes intermedios entre los originales.

Barequet y Sharir [1996] proponen un nuevo método de interpolación para el problema general de cortes con cualquier número de contornos. El algoritmo se basa en un análisis de semejanza entre los dos cortes. Así, se tratan de forma separada porciones similares entre los dos cortes, y después se consideran las partes no *emparejadas*. Entre cada dos porciones de contorno emparejadas se hace un recubrimiento de la superficie mediante triangulaciones que minimicen el área total de la superficie. Para el propósito de identificar las porciones emparejables, usan una técnica de emparejamiento parcial de curvas. Una vez hechos estos emparejamientos, se hace una triangulación sobre esos fragmentos de contorno. Después de esto, resta por determinar la superficie que conecta los tramos no emparejados. Estos tramos, junto con los extremos de los recubrimientos anteriores, forman colecciones de polígonos cerrados 3D. Cuando se proyectan en el plano x-y, estos ciclos no intersectan, ya que con anterioridad se detectaron los trozos de contorno semejantes entre los contornos de los dos cortes, salvo en casos excepcionales. A continuación se triangulan los ciclos 3D usando una técnica de programación dinámica para obtener la triangulación de área mínima en 3D.

Representación de superficies mediante B-splines

Otra forma de reconstrucción de superficies implica el trabajar con curvas de órdenes superiores al de la representación poligonal anterior. Consiste en estimar la función continua para la superficie a partir de puntos muestra, lo cual puede ser implementado por interpolación o aproximación.

Los polinomios son una elección natural para la representación de curvas y superficies. De entre las técnicas de representación que utilizan polinomios, las más populares son los *B-splines*. Los *B-splines* son curvas polinómicas a trozos que están relacionadas con un polígono guía. Los polinomios cúbicos son los más utilizados para los *splines* debido a que son los polinomios de menor orden que permiten cambio de signo de la curvatura. Los *splines* tienen buenas propiedades desde los puntos de vista computacional y de representación, tales como el que se garantiza que la curva varía menos que el polinomio guía, y la interpolación es local; es decir, si un punto del polígono guía se mueve, los efectos se

limitan a los puntos cercanos de la *spline* [Ballard y Brown, 1982].

Los *splines* cúbicos son una secuencia de curvas polinómicas cúbicas que se unen extremo a extremo para representar una curva compleja. Cada segmento del spline cúbico es una curva paramétrica:

$$x(t) = a_x t^3 + b_x t^2 + c_x t + d_x \tag{2.102}$$
$$y(t) = a_y t^3 + b_y t^2 + c_y t + d_y \tag{2.103}$$
$$z(t) = a_z t^3 + b_z t^2 + c_z t + d_z, \tag{2.104}$$

para $0 \leq t \leq 1$. Los polinomios cúbicos permiten que las curvas pasen a través de puntos con una tangente específica y son los polinomios de menor orden que permiten representar curvas no planas.

Estas ecuaciones se pueden representar de forma más compacta organizando los coeficientes en vectores. Si los coeficientes se escriben como

$$\overline{a} = (a_x, a_y, a_z) \tag{2.105}$$
$$\overline{b} = (b_x, b_y, b_z) \tag{2.106}$$
$$\overline{c} = (c_x, c_y, c_z) \tag{2.107}$$
$$\overline{d} = (d_x, d_y, d_z) \tag{2.108}$$

y $\overline{p}(t) = (x(t), y(t), z(t))$, entonces una curva polinómica cúbica se puede escribir como

$$\overline{p}(t) = \overline{a}t^3 + \overline{b}t^2 + \overline{c}t + \overline{d}. \tag{2.109}$$

Para curvas más complejas se puede utilizar una secuencia de polinomios cúbicos unidos por sus extremos:

$$\overline{p}_i = \overline{a}_i t^3 + \overline{b}_i t^2 + \overline{c}_i t + \overline{d}_i, \quad i = 1, \ldots, n \tag{2.110}$$

Esta secuencia se llama spline cúbico y es una forma común de representar curvas arbitrarias en visión.

Los polígonos planos pueden usarse para modelar objetos complejos. Una malla poligonal se representa mediante una lista de coordenadas de vértices que definen los polígonos

planos de la malla. Puesto que un vértice puede estar compartido por muchos polígonos, se usan representaciones indirectas que permiten que cada vértice sea listado una vez, al representar cada cara mediante una lista de vértices. Para asegurar la consistencia necesaria para usar la información de vértices se sigue el convenio de listar los vértices en el orden en que son encontrados, moviéndose en sentido antihorario sobre la cara vista desde fuera del objeto. La representación mediante lista de vértices no representa de forma explícita las aristas entre caras adyacentes y no proporciona una forma eficiente de encontrar todas las caras que incluyen un vértice dado. Estos problemas se resuelven usando la *estructura de datos de bordes con alas*. Ésta es una red con tres tipos de entradas: vértices, bordes y caras. La estructura de datos incluye punteros que pueden seguirse para encontrar todos los elementos vecinos sin buscar en la malla completa ni almacenar una lista de vecinos para cada entrada. Hay una entrada de vértice para cada vértice en la malla poligonal, una entrada de borde para cada borde de la malla poligonal, y una entrada de cara para cada cara de la malla poligonal. El tamaño de cada entrada no cambia con el tamaño de la malla o el número de vecinos. Los vértices a los extremos de un borde y las caras a ambos lados de un borde pueden encontrarse de forma directa; todas las caras (bordes) que usan un vértice se pueden encontrar en un tiempo proporcional al número de caras (bordes) que incluyen el vértice; todos los vértices (bordes) alrededor de una cara se pueden encontrar en un tiempo proporcional al número de vértices (bordes) alrededor de la cara.

Cada entrada de una cara apunta a la entrada de uno de sus lados, y cada entrada de un vértice apunta a la entrada de uno los lados que terminan en dicho vértice. Las entradas de los lados contienen los punteros que conectan las caras y los vértices en una malla poligonal y permiten que la malla sea atravesada de forma eficiente. Cada entrada de borde contiene un puntero a cada uno de sus vértices, un puntero a cada una de las caras que separa, y punteros a los cuatro bordes de ala que son vecinos en la malla poligonal. La figura 2.8 ilustra la información contenida en una estructura de datos de borde.

Como ejemplos del uso de aproximaciones poligonales en el campo de las imágenes médicas se pueden citar, entre otros, los trabajos de Münch y Rüegsegger [1990], Ayache y col. [1990], Kang y col. [1993] y Müller y Rüegsegger [1995].

La representación mediante curvas paramétricas puede extenderse para obtener una representación paramétrica de superficies complejas. La representación de superficies

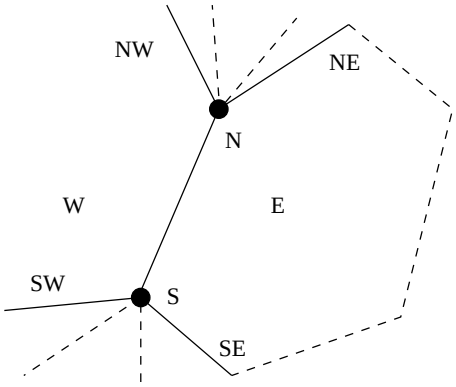


Figura 2.8: Cada entrada de borde contiene punteros a sus dos vértices, a las dos caras que separa, y a los cuatro bordes de ala.

mediante producto de tensores se forma mediante la combinación de dos representaciones de curvas, una en cada coordenada paramétrica,

$$\overline{p}(u,v) = [u^3 u^2 u 1] \begin{bmatrix} \overline{a}_1 \\ \overline{a}_2 \\ \overline{a}_3 \\ \overline{a}_4 \end{bmatrix} [\overline{b}_1 \overline{b}_2 \overline{b}_3 \overline{b}_4] \begin{bmatrix} v^3 \\ v^2 \\ v \\ 1 \end{bmatrix} \tag{2.111}$$

La superficie paramétrica puede escribirse como

$$\overline{p}(u,v) = U^t M V \tag{2.112}$$

donde la matriz M tiene como coeficientes los vectores de coeficientes para cada coordenada de la superficie paramétrica.

Muchas veces interesa más encontrar la superficie que aproxima los datos que interpolar éstos. Muchas representaciones de superficies se pueden expresar como combinaciones de funciones base:

$$f(x,y;a_0,a_1,\ldots,a_m) = \sum_{i=0}^{m} a_i B_i(x,y) \tag{2.113}$$

donde a_i son los coeficientes escalares de las funciones base B_i. Con splines producto de tensores, las funciones base se organizan en una cuadrícula:

$$f(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_{ij} B_{ij}(x,y). \tag{2.114}$$

Consideremos inicialmente el caso unidimensional, asumiendo que las funciones base están uniformemente espaciadas en posiciones enteras. La función base $B_i(x)$ será distinta de cero en el intervalo $[i, i+4)$. Las funciones base están localizadas en las posiciones de 0 a m, de tal forma que hay $m+1$ coeficientes que definen la forma de la curva base B–spline, como se ilustra en la figura 2.9.

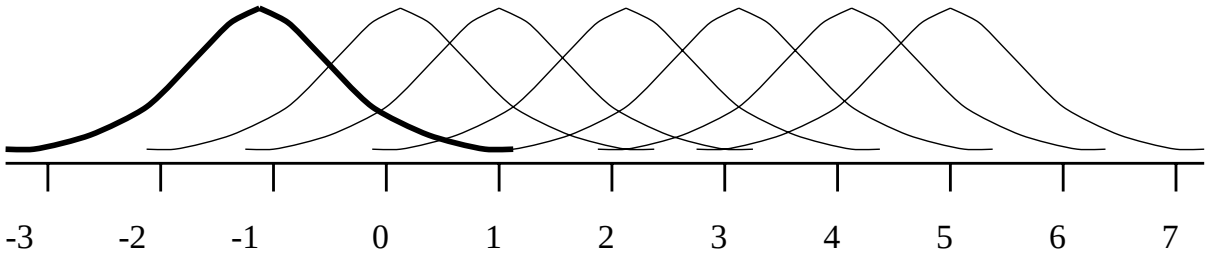


Figura 2.9: La curva B–spline en una dimensión es una combinación lineal de funciones base localizadas en posiciones enteras.

La curva B–spline se define en el intervalo de $x = 3$ a $x = m + 1$. La función base en $x = 0$ se extiende tres intervalos a la izquierda de la curva y la función base en $x = m + 1$ se extiende tres intervalos a la derecha. Estos intervalos extra establecen las condiciones de contorno y se deben incluir para la correcta definición de la curva. Cada función base es distinta de cero sobre cuatro intervalos. Puesto que las funciones base se solapan, cada intervalo está cubierto por cuatro funciones base. Cada segmento $b_i(x)$ de la función $B_i(x)$ es un polinomio cúbico que se define sólo en su intervalo entero, como se representa en la figura 2.10.

Los polinomios cúbicos para los segmentos individuales de los B–splines son:

$$b_0(x) = \frac{x^3}{6} \tag{2.115}$$

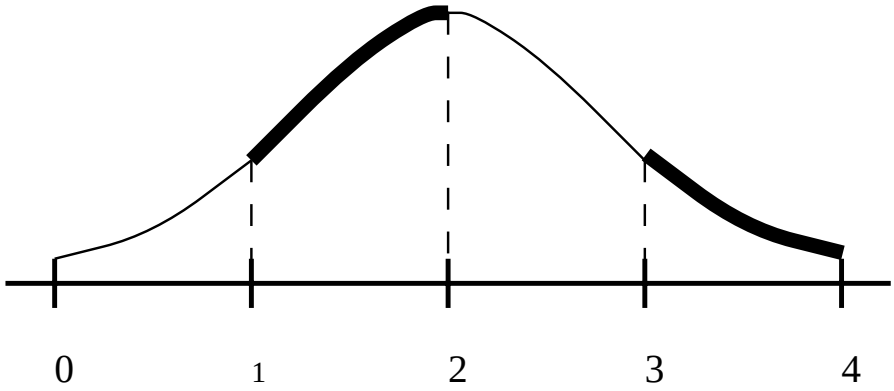


Figura 2.10: Cada función base B–spline consiste de cuatro polinomios cúbicos que son distintos de cero en intervalos adyacentes.

$$b_1(x) = \frac{1+3x+3x^2-3x^3}{6} \tag{2.116}$$

$$b_2(x) = \frac{4-6x^2+3x^3}{6} \tag{2.117}$$

$$b_3(x) = \frac{1-3x+3x^2-x^3}{6} \tag{2.118}$$

Para la aproximación de superficies en 3D, el modelo es una superficie B–spline producto de tensores:

$$f(x,y) = \sum_{i=0}^{n}\sum_{j=0}^{m} a_{ij}B_{ij}(x,y) = \sum_{i=0}^{n}\sum_{j=0}^{m} a_{ij}B_i(x)B_j(y). \tag{2.119}$$

El problema de aproximación de superficie se resuelve minimizando la norma

$$\chi^2 = \sum_{k=1}^{N}[z(k) - (\sum_{i=0}^{n}\sum_{j=0}^{m} a_{ij}B_{ij}(x,y))]^2. \tag{2.120}$$

Cada función base cubre 16 cuadrículas $[i, i+4) \times [j, j+4)$. Puesto que las funciones base se solapan, cada cuadrícula rectangular está cubierta por 16 funciones base. Cada función base se puede separar en el producto de dos funciones base unidimensionales. Las 16 *parcelas* polinómicas en una función base se forman a partir del producto a pares de los polinomios unidimensionales.

Como ejemplos de la aplicación de splines en la representación de curvas y superficies de estructuras anatómicas podemos citar los trabajos de Breau y col. [1991], Brunie y col. [1991], Bae y col. [1992], Chen y col. [1992], Kankanhalli y col. [1993], Steele y col. [1994], Declerck y col. [1995], Radeva y col. [1997].

2.4.2 Representaciones volumétricas

Las representaciones de objetos en forma de primitivas sólidas es, con frecuencia, útil debido a sus propiedades de fácil computación y concisión. De entre estas técnicas, la más simple es la de *ocupación espacial*, en la que un volumen se representa como una matriz tridimensional de celdas marcadas como pertenecientes, o no, al objeto a representar. Para obtener una buena resolución se necesita, en este caso, de una cantidad de memoria alta. Cuando se trabaja con objetos muy irregulares tal como ocurre en la tomografía computerizada, esta representación es muy común [Raya, 1990; Dhawan y Arata, 1991; Keyak y Skinner, 1992; Loncaric y col., 1995; Sonka y col., 1996].

Otra técnica para representación de volúmenes mediante primitivas es la de *descomposición de celdas*. Aquí las celdas son más complejas en forma, pero todavía no comparten volumen, por lo que la única operación posible es la de *pegar*. Estas técnicas no son especialmente concisas. Entre estas técnicas está la del *oct-tree* [Jackins y Tanimoto, 1980; Yau y Srihari, 1983; Hull y col., 1990]. Esta técnica produce una subdivisión recursiva del volumen.

Otro esquema de representación volumétrico, menos usado en el campo de las imágenes médicas, es el de la *geometría sólida constructiva* (CSG). Los sólidos se representan como composiciones, mediante operaciones de conjuntos, de otros sólidos que pudieran haber sido sometidos a movimientos rígidos. En el nivel más bajo están las primitivas sólidas, que son intersecciones acotadas de semiespacios cerrados definidas por alguna función analítica $F(x, y, z) \geq 0$, que pueden estar arbitrariamente escaladas y posicionadas en el espacio. Una representación CSG es una expresión formada mediante primitivas sólidas y un conjunto de operadores para combinación y movimiento. Los operadores de combinación son versiones regularizadas de la unión, intersección y diferencia de conjuntos. La regularización hace referencia a la propiedad de conservación de la dimensión. Como ejemplo de este tipo de representación se puede citar el sistema de representación 3D de imágenes médicas de Ney y Fisman [1991].

2.5 Sistemas automáticos de análisis de imágenes

El rápido desarrollo y proliferación de tecnologías para obtención y tratamiento de imágenes médicas está revolucionando la medicina. Las imágenes médicas permiten al médico obtener información del interior del cuerpo humano sin necesidad de utilizar métodos invasivos. El papel de las imágenes médicas ha ido más allá de la simple visualización e inspección de las estructuras anatómicas. Se ha convertido en una herramienta para planificación quirúrgica y simulación, navegación intraoperativa, planificación de radioterapia, y para seguimiento de la evolución de la terapia.

Aunque los modernos dispositivos de obtención de imágenes proporcionan vistas excepcionales de la anatomía interna, el uso de métodos computacionales para la cuantificación y análisis de las estructuras implicadas es todavía de una eficiencia limitada. La segmentación de estructuras a partir de imágenes médicas y la reconstrucción de representaciones geométricas es difícil debido a la complejidad y variabilidad de las formas anatómicas. Además, los inconvenientes típicos del muestreo de datos tales como artefactos, solape espacial y ruido, pueden causar que los contornos de las estructuras sean indistinguibles y desconectados en muchas de sus partes. El objetivo es extraer los elementos de contorno pertenecientes a la misma estructura e integrarlos en un modelo de la estructura coherente y consistente.

Las técnicas tradicionales de procesado de bajo nivel descritas con anterioridad, basadas únicamente en criterios de homogeneidad, pueden hacer suposiciones incorrectas durante el proceso de integración y generar contornos incorrectos. En general, no pueden proporcionar la descomposición de una imagen en regiones que se correspondan con partes de objetos del mundo real [Cho y Meer, 1997]. Como resultado estas técnicas, libres de modelo, requieren por lo general gran cantidad de interacción por parte del experto. Para evitar esta interacción en la medida de lo posible, es preciso recurrir a técnicas que incorporen conocimiento sobre el dominio que pueda restringir el, de otra forma libre, proceso de segmentación.

Los pasos de procesado requeridos para efectuar el análisis se aplican en secuencia dando la idea de un *pipeline* de análisis a través del cual circulan los datos de entrada [Pun y col., 1994]. El *pipeline* genérico de análisis de imágenes y visión computacional es el que se representa en la figura 2.5.

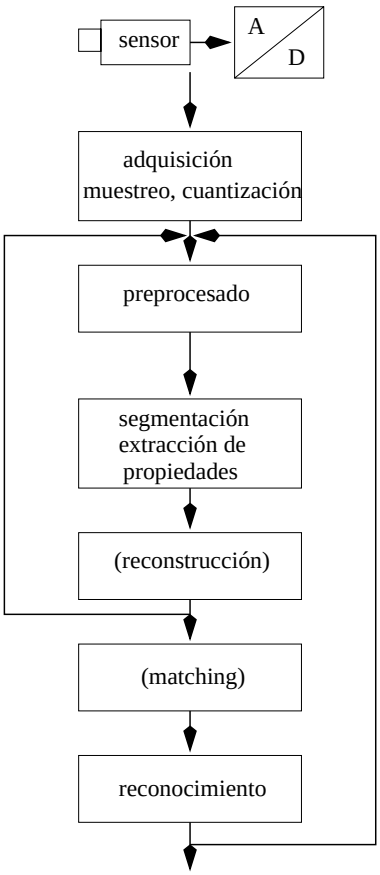


Figura 2.11: *pipeline* genérico de análisis de imágenes médicas.

En una primera etapa se adquieren las imágenes haciendo una transformación de analógico a digital necesaria para el manejo computacional de la información. Tras ésta, viene una etapa de preprocesado que incluye una serie de operaciones realizadas de forma sistemática sobre todas las imágenes (reducción de ruido, realce, ...). La siguiente etapa corresponde a la segmentación, en la cual se divide la imagen en sus partes constitutivas, con el fin de obtener las características significativas. En esta etapa operan tanto procesos de bajo nivel como procesos que integran conocimiento concreto sobre el domino de aplicación, de los que presentamos una revisión en la primera parte de este capítulo. Esta etapa puede ir seguida de otra de reconstrucción en la que se pretende proporcionar información n-dimensional a partir de las series de datos m-dimensionales, con $m < n$. Un ejemplo clásico lo constituye el modelado geométrico 3D a partir de información extraida de cortes 2D. Opcionalmente se pueden combinar series de datos n-dimensionales, como la

correspondiente a información de diferentes fuentes, o de la misma fuente pero en distintos intervalos temporales. Finalmente, está la etapa de reconocimiento en la que se identifican los elementos presentes en la escena. Existen variantes a este flujo unidimensional en las que la información ya extraida puede permitir refinar el proceso de análisis. Este uso de la información descendente (*top-down*) aún no es demasiado común en imágenes médicas, pero si se encuentra cada vez más a nivel teórico y en aplicaciones industriales.

Las imágenes biomédicas se caracterizan por su baja relación señal/ruido, debida tanto a los métodos de formación de la imagen como al proceso de digitalización y adquisición de las mismas para su tratamiento computacional. Esto hace que el tratamiento mediante los métodos clásicos basados en regiones o bordes no de resultados satisfactorios. A su vez, la natural variación de las estructuras anatómicas hace difícil el manejo de modelos a priori que puedan guiar con absoluta confianza el proceso de segmentación. En un intento de mejorar la calidad de la segmentación se han propuesto sistemas que con alguna de estas tres características [Chu y Aggarwal, 1993]:

- Integración de información obtenida a partir de diferentes técnicas de segmentación.
- Utilización de información de continuidad 3D.
- Aportación de conocimiento específico sobre el dominio de aplicación.

En general, los sistemas para análisis de imágenes van a constar de un módulo de segmentación de bajo nivel, de cuyos métodos hemos realizado una revisión en las secciones 2.2.1, 2.2.2, 2.2.3 y 2.2.4, otro módulo para incorporación y/o integración de conocimiento, cuyos aspectos mas importantes describimos en las secciones 2.2.5 y 2.3, y un módulo para la obtención del modelo geométrico 3D. Aspectos relacionados con esto último los tratamos en la sección 2.4. A continuación comentaremos algunos de los sistemas propuestos en la bibliografía para el análisis de imágenes procedentes de distintos dominios, terminando con sistemas que analizan imágenes biomédicas.

Un ejemplo clásico del uso de un sistema experto basado en reglas para la segmentación es el desarrollado por Nazif y Levine [1984]. El conocimiento empleado en su sistema es independiente del dominio, incluyendo conocimiento independiente de la escena y de propósito general sobre imágenes y criterios de agrupamiento. El sistema de Nazif y Levine, cuyo esquema se muestra en la figura 2.5, contiene un conjunto de procesos (ini-

Figura 2.12: Diagrama de bloques del sistema de Nazif y Levine.

cializador, analizador de líneas, analizador de regiones, analizador de áreas, foco de atención y el planificador), más dos memorias asociativas (*short-term* (STM) y *long-term* (LTM)). La memoria STM mantiene la imagen de entrada, los datos de la segmentación y la salida. La memoria LTM contiene el modelo que representa el sistema de conocimiento sobre la segmentación de bajo nivel y las estrategias de control. El sistema procesa reglas de la LTM chequeando sus condiciones sobre los datos de la STM. Cuando la parte de la condición de la regla se cumple, se ejecuta la parte de acción de la regla, lo cual implica en general la modificación de los datos.

El modelo almacenado en la LTM tiene tres niveles de reglas. En el nivel 1 están las reglas que codifican información a cerca de las propiedades de las regiones, de las líneas y de las áreas, en la forma de pares situación-acción. Las acciones específicas incluyen dividir una región, unir dos regiones, añadir, borrar o extender una línea, unir dos líneas y crear o modificar una área de foco de atención. Las reglas de conocimiento se clasifican por sus acciones. En el nivel 2 están las reglas de control, las cuales se dividen en dos categorías: reglas de foco de atención y reglas de inferencia. Las reglas de foco de atención encuentran la siguiente entrada de datos a considerar: una región, una línea o un área entera. Las reglas de inferencia son metareglas en las que las acciones no modifican los datos de la STM, sino que alteran el orden de chequeo de los diferentes conjuntos de reglas de conocimiento. De esta forma controlan que procesos se van a activar a continuación. En el nivel 3, el nivel de reglas más alto, están las reglas de estrategia que seleccionan el conjunto de reglas de control que ejecutan la estrategia de

control más apropiada para un conjunto de datos dado. Las condiciones de las reglas estan constituidas por un cualificador simbólico que representa una operación lógica a ejecutar sobre los datos, un símbolo que denota la entrada de datos sobre la que se va a chequear la condición, una característica de esta entrada de datos, un cualificador opcional NOT, y un cualificador DIFFERENCE opcional que aplica la operación a diferencias en los valores de las propiedades.

Brooks [1979, 1983] desarrolló un sistema basado en modelos e independiente del dominio para tareas de visión. El sistema permite el reconocimiento de estructuras 3D a partir de imágenes 2D. Este sistema, llamado ACRONYM, se basa en modelos geométricos tridimensionales detallados que el usuario ha de introducir para dirigir la interpretación de la escena concreta. Originariamente usaba modelos de objetos específicos para dirigir un sistema de razonamiento geométrico cualitativo en el etiquetado de imágenes. El ámbito de ACRONYM se incrementó para incluir reconocimiento de clases de objetos y extracción de información tridimensional de las imágenes. El sistema funciona de la forma que se indica a continuación. Parte de un modelo *a priori* del mundo. Los objetos se modelan como jerarquías de subpartes de conos generalizados, cada uno con un sistema de coordenadas local relacionado con un sistema de coordenadas global. Estos conos generalizados describen un volumen tridimensional. En ACRONYM se representan los conos generalizados, la jerarquía de subpartes de los objetos y el árbol que representa las relaciones espaciales de los objetos. Esta representación es suficientemente rica para objetos completamente especificados. Sin embargo, es común en tareas de intepretación que no se conozcan todos los detalles de los objetos *a priori*, por lo que es deseable representar clases de objetos, a lo que llaman descripciones genéricas. La generalización consiste en permitir que los números sean reemplazados por variables formales, *cuantificadores*, y expresiones algebraicas sobre ellas. Adicionalmente se colocan ligaduras sobre dichas variables formales. Las entradas al sistema son imágenes preprocesadas mediante el detector de líneas de Nevatia y Babu. A partir de los bordes se extraen descripciones de las imágenes como *cintas*, representaciones 2D análogas a los conos generalizados en 3D, y *elipses*, para posteriormente identificar realizaciones de las clases de modelos de objetos. Mediante un sistema de razonamiento geométrico se identifican características y relaciones invariantes. Tales predicciones guiarán los procesos de descripción de bajo nivel, *filtrando* de esta forma las características que se utilizarán para el chequeo (*matching*) de propiedades locales. Estos chequeos locales se combinan sujetos a las restricciones

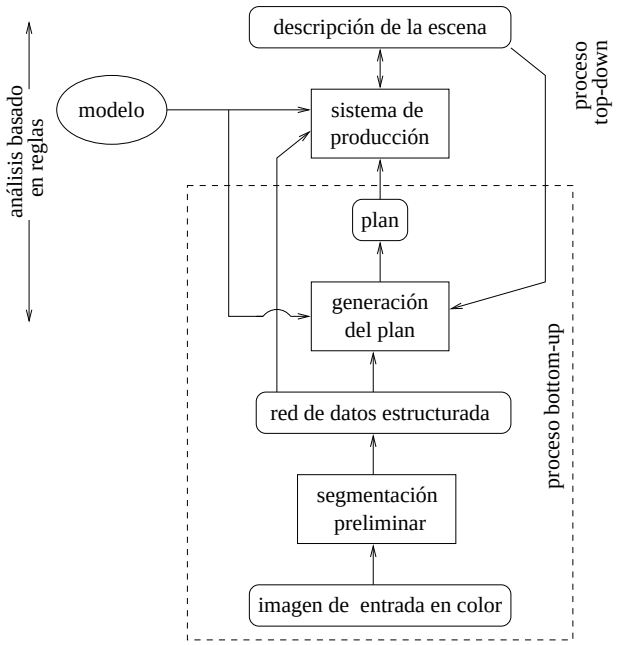


Figura 2.13: Esquema del analizador de regiones de Ohta.

impuestas por las ligaduras de relación y algebraicas. El resultado es la interpretación tridimensional de la imagen.

El sistema de control que implementa es el de encadenamiento hacia atrás (*chaining backward*), en el que las reglas establecen una serie de submetas, y, recursivamente, se va invocando el mecanismo de reglas para satisfacer esas submetas (estrategia de control descendente). El orden en que se aplican las reglas y el flujo de control no se pueden caracterizar a nivel local y son completamente dependientes de los modelos dados al sistema. Las limitaciones de ACRONYM se deben principalmente a la debilidad del proceso de segmentación y a los problemas que produce una estrategia descendente en la interpretación de escenas complejas.

Ohta [1985] desarrolló un sistema para la interpretación de escenas de exteriores en color, cuyo esquema se muestra en la figura 2.13. El sistema recibe una imagen en color que inicialmente sobresegmenta originando un conjunto coherente de regiones. Esta imagen segmentada se organiza en una red de datos estructurada sobre la que se aplica conocimiento mediante un conjunto de reglas que deciden sobre la unión de regiones. En

esta etapa de crecimiento de regiones se utiliza un control mixto ascendente y descendente. Mediante un proceso ascendente se genera un plan de imagen como una representación de la estructura de entrada. El plan es un conjunto de etiquetas y su grado de corrección que se establecen mediante razonamiento aproximado basándose en grados de certeza borrosos. En este proceso no se usa información semántica. El proceso descendente se organiza como un sistema de producción, donde cada regla es un par condición–acción que representa una unidad de representación del conocimiento.

Otro ejemplo del uso de conocimiento es el del sistema de segmentación basado en reglas desarrollado por Ehricke [1991] para el análisis automático de imágenes de Resonancia Magnética 3D. Aquí se justifica el empleo de conocimiento específico sobre la tarea específica a ejecutar en base al elevado grado de complejidad de las estructuras anatómicas que hacen que los operadores basados en regiones o bordes necesiten una gran interacción por parte del usuario para control y corrección del resultado de la segmentación. Ehricke desarrolló un sistema jerárquico para la segmentación de conjuntos de datos 3D de la cabeza. El sistema comienza con una subdivisión de la escena en objetos y fondo, y se van reconociendo estructuras con nivel de detalle mayor a medida que progresa la segmentación. La división inicial se realiza mediante la delineación de contornos en la cabeza proporcionando una separación entre objetos y fondo (background). El sistema consta de tres componentes: (1) el sistema de control, (2) un conjunto de métodos y (3) la base de conocimiento. El proceso de análisis es supervisado por la componente de control, la cual selecciona el método más adecuado en cada momento durante el proceso de segmentación, ajusta parámetros, inicia su aplicación y chequea los resultados. Si éstos no son satisfactorios se cambian parámetros o se selecciona otro método. Para este tipo de control se hace uso de conocimiento almacenado como un conjunto de reglas y hechos en la base de conocimiento.

Los métodos de procesado de bajo nivel utilizados se agrupan en las categorías de segmentación, verificación y postprocesado. Los operadores de segmentación incluyen clasificadores, operadores tipo gradiente y técnicas de contornos. Los métodos de verificación chequean los resultados de la aplicación del operador para proporcionar conocimiento al sistema de control sobre el estado del análisis. Los métodos de postprocesado se usan para refinar los resultados de la segmentación, entre los cuales se encuentran los operadores morfológicos.

Un ejemplo de sistema jerárquico es el presentado por Bae y col. [1992], para la

segmentación del área correspondiente al hígado a partir de imágenes CT, y el de Michael y Nelson [1989] para la segmentación de huesos en radiografías de manos. Estos métodos presentan con frecuencia errores en la segmentación final debido al poco conocimiento a priori utilizado y a la ausencia de realimentación en el proceso de segmentación.

Stansfield [1986] propuso un sistema basado en reglas para la segmentación automática de vasos coronarios a partir de angiogramas de pecho. El sistema global consta de tres módulos: uno de preprocesado y otros dos en los que está embebido el sistema experto. En la etapa de preprocesado se aplican técnicas de procesado de bajo nivel, una basada en regiones y otra basada en bordes. Se consideran las regiones obtenidas por el analizador de regiones como indicadores de los objetos reales, por su menor sensibilidad al ruido, y después se combinan ambas representaciones para el refinamiento del análisis de regiones. El sistema experto se encarga, a continuación, de etiquetar los vasos y eliminar estructuras ruidosas. Este es un sistema basado en reglas con control ascendnte o conducido por los datos, justificado por la natural diversidad de la anatomía cardíaca. En la etapa de bajo nivel del sistema experto se aplica conocimiento independiente del dominio para la agrupación y establecimiento de relaciones geométricas, tanto sobre el mapa de regiones como el de bordes. El objetivo es unir segmentos de líneas y agrupar regiones. En la etapa de alto nivel se aplica conocimiento dependiente del dominio para interpretrar el resultado de la segmentación anterior. Stansfield apunta finalmente que la separación entre conocimiento de bajo y alto nivel da lugar a una serie de problemas: la segmentación de bajo nivel sin ningún conocimiento a priori del dominio conduce, con frecuencia, este proceso por mal camino; además, el uso de estructuras de control conducidas por datos parece inadecuado para la implementación, incluso, de los sistemas basados en conocimiento simples.

Dhawan [1990] propuso un sistema automático para el análisis de imágenes CT de tórax para etiquetar las diferentes estructuras de interés que consta de cuatro bloques básicos: un bloque de preprocesado a nivel de entrada, para realzar las características y eliminar ruido; un bloque de segmentación de bajo nivel, donde la imagen se segmenta en regiones y se extrae un conjunto de características globales para ayudar al análisis de la segmentación para obtener un número razonable de regiones segmentadas significativas; un bloque para la extracción de características de nivel intermedio, donde se extraen las características específicas y se transforman en la forma apropiada, simbólica o cuantitativa; y un bloque de interpretación de alto nivel con una base de conocimiento para el

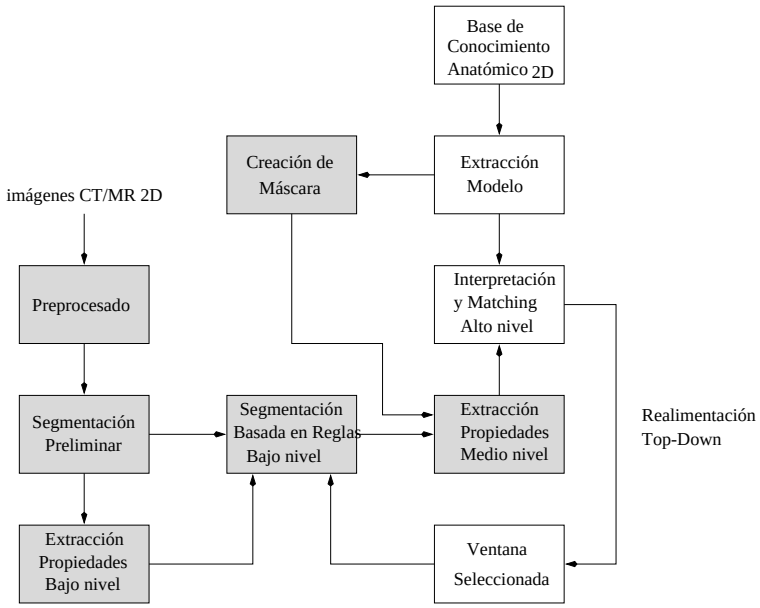


Figura 2.14: Diagrama de bloques del sistema de análisis 2D de Dhawan.

etiquetado e interpretación de las regiones segmentadas, de acuerdo con sus características específicas, usando la base de conocimiento. Un esquema del diagrama de bloques es el que se presenta en la figura 2.14.

Para realce de características junto con la eliminación de ruido, que se lleva a cabo en el bloque de preprocesado a nivel de entrada, propone una técnica de procesado con una ventana adaptativa de la característica. El algoritmo calcula el contraste de la característica con su fondo y a continuación lo realza usando una función de realce de contraste que puede diseñarse y sintonizarse de acuerdo con el conocimiento derivado del procesado del entorno. Tras esta etapa se realiza la segmentación de bajo nivel usando una técnica de procesado basado en pirámides multi-resolución variante de la propuesta por Hong y Rosenfeld [1984]. Esta segmentación es analizada posteriormente por un sistema basado en reglas que utiliza las propiedades de las regiones y el conocimiento de una máscara que representa las principales áreas anatómicas, para obtener una segmentación más significativa. El sistema basado en reglas se basa en el de Nazif y Levine [1984] pero es más efectivo y menos complejo debido a que usa datos presegmentados en vez de pixels y las reglas utilizan conocimiento específico y son menos numerosas. El sistema permite el análisis utilizando tanto una estrategia ascendente como una estrategia descendente. En el nivel

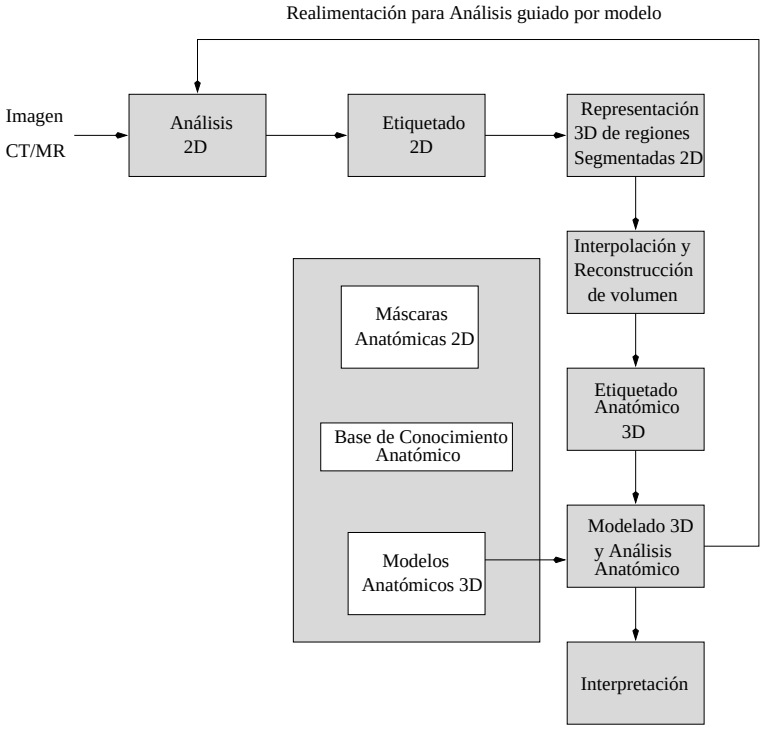


Figura 2.15: Diagrama de bloques del sistema de análisis 3D de Dhawan.

intermedio se extraen todas las características de las regiones útiles para el sistema de interpretación de alto nivel (área, centroide, orientación, adyacencias, elongación, y nivel medio). En el sistema de alto nivel se hace la comparación (*matching*) simbólica entre las características anteriores y las de los modelos. Dhawan [1991] extendió este método al análisis 3D a partir de imágenes 2D. El diagrama de bloques del sistema de análisis 3D es el que se muestra en la figura 2.15.

Nanning y Jianqin [1992] propusieron un nuevo esquema para la representación del conocimiento, que denominaron cuaderno (*Notebook*), y un nuevo método de segmentación para la segmentación inteligente de imágenes de sonar. Según ellos, una buena representación del conocimiento debe disponer el conocimiento de control, el procedural y el declarativo de forma coordenada y expresar explícitamente la descripción, el almacenamiento y la recuperación. Distintos tipos de representación del conocimiento tales como redes semánticas, modelo pizarra y dependencia conceptual enfatizan más la de-

scripción del problema en el nivel conceptual, pero el procesado semántico de bajo nivel es más bien difícil e ineficiente, puesto que la mayoría de las representaciones se proponen bajo la consideración de significado de concepto e inferencia de símbolos. El esquema *cuaderno* consiste de dos partes: una estructura basada en marcos (*frames*) y un procedimiento de interpretación. El primero contiene conocimiento declarativo acerca de la descripción del modelo. Es el núcleo del esquema y se conecta con otras partes. Consta de una serie de tablas para la descripción de objetos y regiones, y para el control de acceso. El procedimiento de interpretación es una integración de reglas de producción e inferencia gramatical. Representa el conocimiento de control y la estructura de datos para el reconocimiento de patrones sintáctico basado en conocimiento. Las reglas basadas en conocimiento se aplican para procesar la información donde la inferencia gramatical no es eficiente, y el análisis sintáctico se aplica cuando las reglas correspondientes son tan complejas que el correspondiente árbol de análisis no podría ser buscado por el algoritmo rápido. La inferencia gramatical automática se construye mediante un modelo heurístico. La interacción entre las dos partes se completa con un módulo de activación. Este método de representación del conocimiento, cuyo esquema se muestra en la figura 2.16, fue usado en el procesado de señal de bajo nivel dentro de un sistema experto de diagnosis para imágenes de sonar (MDESS). Su arquitectura se describe en la figura 2.17. La segmentación se implementa mediante un esquema jerárquico en tres niveles: en un primer paso se hace una segmentación de bajo nivel mediante un algoritmo de crecimiento de regiones; en un segundo paso se realiza la detección de objetos mediante conocimiento heurístico y en la tercera etapa se realiza el etiquetado de objeto y fondo. En el primer paso no es adecuado introducir conocimiento a causa del elevado número de regiones primarias. En esta fase es más adecuado utilizar heurísticas y modelos borrosos (*fuzzy*). Cuando el número de regiones se reduce, y en el nivel intermedio, es cuando se debe usar conocimiento de alto nivel para interpretar los resultados intermedios y guiar la segmentación de la forma más segura. Los dos últimos niveles son relativos y pueden interactuar.

Sonka [1996] presentó un sistema basado en reglas para la detección de las vías respiratorias intratorácicas a partir de conjuntos de imágenes de CT. El método usa conocimiento anatómico *a priori* y se basa en la combinación de crecimiento de regiones 3D, segmentación basada en reglas en 2D de cada uno de los cortes de CT, y la unión (*merging*) a través del conjunto 3D de cortes. El sistema realiza una segmentación mediante umbral-

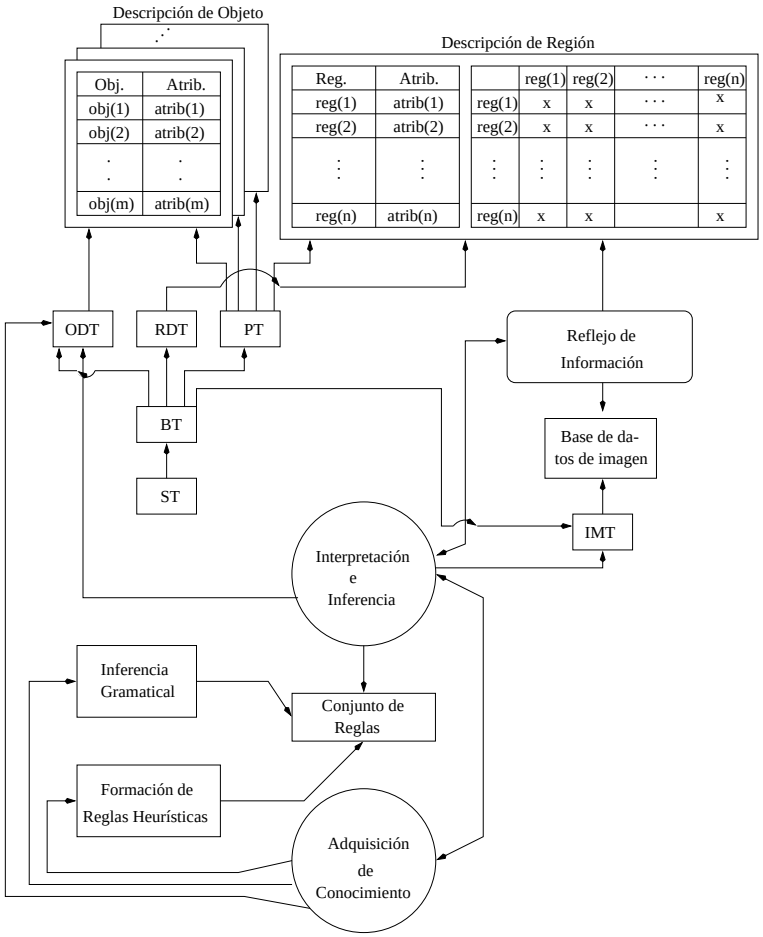


Figura 2.16: Esquema *Notebook*

ización para determinar el área pulmonar. A continuación se hace un procesado individual de cada corte mediante un operador de bordes, un proceso de umbralización, y otro de crecimiento de regiones que permiten detectar regiones de fondo fijas, y candidatos a vías respiratorias, a vasos sanguíneos y a fondo. El siguiente paso consiste en la aportación de conocimiento anatómico mediante un conjunto de reglas. Éstas se aplican en tres pasos secuenciales, de tal forma que al final del primer paso quedarán fijadas las etiquetas de fondo, al final del segundo quedarán fijadas las regiones de vasos sanguíneos y en el tercero se identificarán las vías respiratorias. En los diferentes pasos se van modificando los grados de confianza del etiquetado basándose en el conocimiento a priori y un grafo de adyacencias. Finalmente, las regiones identificadas para cada corte se apilan en el espacio

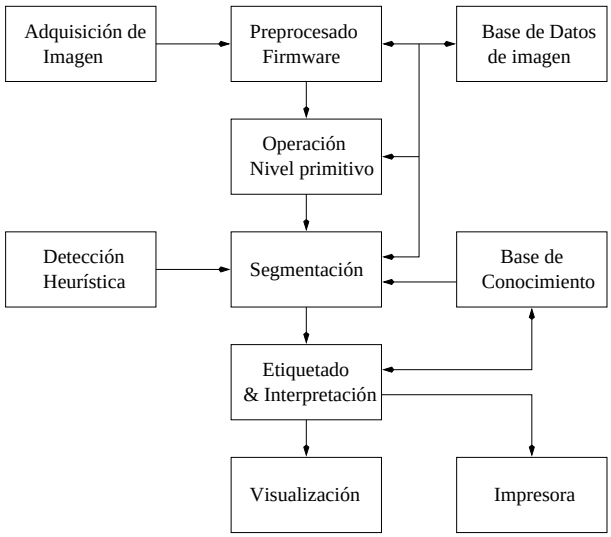


Figura 2.17: Arquitectura del sistema MDESS

3D para representar el árbol de vías respiratorias en 3D. En general, la identificación de regiones en los cortes individuales no es totalmente correcta, por lo que el apilamiento 3D no forma un único objeto, por lo que se necesita de la intervención humana para la correcta identificación.

Brunie y col. [1991] presentaron una técnica para la representación segmentada tridimensional de huesos de la muñeca a partir de un conjunto de cortes de CT. El objetivo del sistema era separar las estructuras óseas para permitir la emulación de la operación quirúrgica de forma individualizada. Para ello utiliza un procedimiento de segmentación corte a corte, para a continuación clasificar el conjunto de contornos como pertenecientes a uno de los diferentes tipos de hueso. El esquema del sistema se muestra en la figura 2.18. Como técnica de segmentación se utiliza una umbralización doble con agregación de regiones. Y para el etiquetado de contornos se construye un grafo de adyacencias con representación simbólica de contornos mediante centros de masas y elipses de mejor aproximación. Finalmente se hace una representación de las superficies mediante triangulaciones. Estas superficies se someten a suavización e interpolación para dar apariencia más realista a la visualización.

Münch y Rüegsegger [1990] describen un sistema semiautomático para la reconstruc-

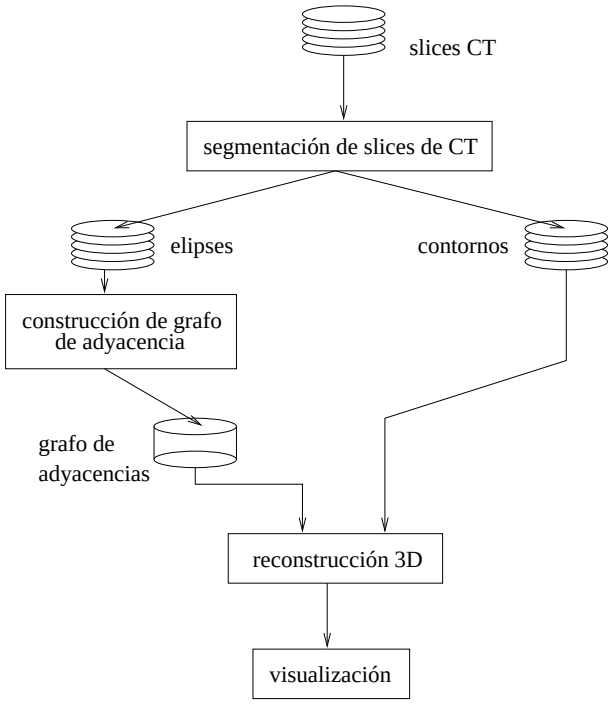


Figura 2.18: Arquitectura general para segmentación de muñeca

ción 3D de rodilla a partir de imágenes de QCT (tomografía computerizada cuantitativa) de dosis baja. Se hace una extracción de contornos de fémur y tibia corte a corte. Emplean un algoritmo interactivo para la búsqueda de contornos, en el que a partir de una suposición inicial, se va aproximando paso a paso al contorno más probable. El criterio de probabilidad contiene los siguientes aspectos: (1) los valores de intensidad fuera del contorno no son inferiores a los del músculo, (2) los valores de intensidad dentro de los contornos están próximos a los de hueso cortical, (3) los puntos del contorno no se desviarán mucho de las posiciones en la iteración previa, y (4) se favorecerán los puntos de contorno en la vecindad de otros puntos de contorno. Los pesos de cada aspecto se optimizan en base a una segmentación manual, y los contornos necesitan de corrección cuando hay destrucción ósea severa. A continuación se hace una evaluación cuantitativa, no interactiva, de los objetos óseos para determinar regiones de pérdida ósea y lesiones. Y finalmente se reconstruye la superficie por triangulación. Una descripción del sistema global aparece en la figura 2.19.

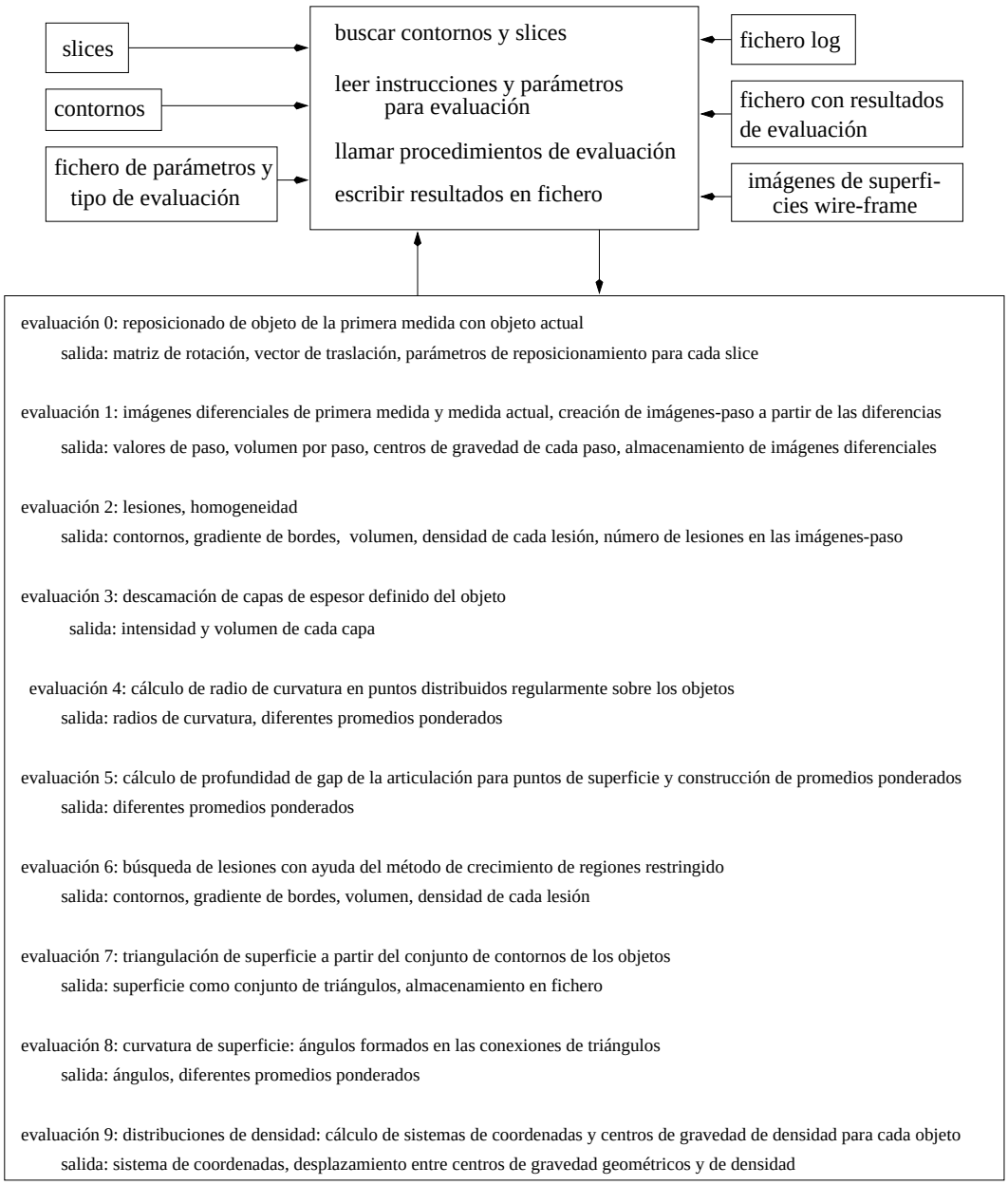


Figura 2.19: Evaluación 3D de hueso de Münch y Rüegsegger.

Zhu y col. [1991] presentan un sistema para la identificación y cuantificación de estructuras anatómicas y fisiológicas mediante segmentación 3D. El sistema parte de un conjunto de secciones de tomografía computerizada que inicialmente procesa una a una

para suavizarlas y obtener una presegmentación. La suavización la ejecuta mediante un filtrado espacial adaptativo, y la presegmentación de las imágenes 2D se realiza mediante umbralización seguida de filtrado morfológico, y se asigna a cada pixel de la imagen una etiqueta de acuerdo con su nivel de gris. El proceso de crecimiento de volúmenes utilizará la conectividad espacial entre pixels de cortes vecinos. La segmentación 3D se realiza mediante crecimiento de volúmenes. Para inicializar este proceso se define una región semilla para cada región volumétrica. El sistema se aplica a segmentación de troncos de madera para la localización e identificación de defectos internos.

Taratorin y Sideman [1993] presentaron un sistema para la detección automática de los contornos ventriculares empleando técnicas de teoría de conjuntos a partir de imágenes de tomografía computerizada rápida (cine CT). El sistema se basa en la incorporación de conocimiento a priori sobre la geometría del corazón, sus niveles de intensidad, estructura espacial y dinámica temporal, dentro del algoritmo de detección de contornos. El conocimiento es interpretado como conjuntos de ligaduras en el espacio funcional. Los contornos consistentes se considera que pertenecen a la intersección de todos los conjuntos introducidos, satisfaciendo así la información a priori.

Finalmente, Ayache y col. [1990] describen un sistema para la interpretación automática de imágenes 3D. Pretenden con el sistema, mejorar la diagnosis, así como la planificación, simulación y control de la acción terapéutica. Para las tareas de segmentación proponen el uso de modelos deformables combinados con la detección de puntos de borde. La segmentación 3D se realiza propagando los resultados de la segmentación 2D de un corte al siguiente. El sistema lo aplican a imágenes cardíacas, para lo que definen las expresiones de energía más adecuadas y resulta casi automático.

A modo de conclusión, podemos decir que el problema de segmentación e interpretación de imágenes sigue todavía abierto. Cuando se trata con imágenes reales no es posible identificar los objetos de interés en la imagen a no ser que inyectemos conocimiento explícito sobre el dominio. Esto en lo referente al alto nivel, pero incluso en el bajo nivel se observa como unas técnicas de segmentación son mas efectivas que otras dependiendo del tipo y modalidad de imagen. Así pues, la tarea de interpretación de una imagen concreta requiere el diseño de sistemas *ad hoc*. No es posible construir sistemas de visión *universales*, sino que se debe abordar el desarrollo de sistemas para tareas particulares y dentro de entornos controlados. Por otra parte, aunque se debe mostrar sensibilidad a las consideraciones prácticas de velocidad y coste, el enorme volumen de datos y la

naturaleza de los cálculos requeridos hacen difícil el alcanzar un compromiso satisfactorio entre estos factores [Klaus y Horn, 1997]. En los siguientes capítulos describiremos nuestro sistema para la reconstrucción automática 3D de tibia a partir de una secuencia de imágenes de Tomografía Computacional (CT). Se trata de un sistema que integra los resultados obtenidos mediante dos esquemas de segmentación diferentes y complementarios con el fin de dotar al proceso de la mayor robustez posible. Uno de los esquemas está basado en regiones y lleva a cabo un crecimiento de las mismas basado en reglas a partir de una presegmentación de bajo nivel. Este sistema implementa una estrategia mixta ascendente-descendente (*bottom-up/top-down*) guiado por datos y por modelos. El otro es un esquema basado en contornos deformables que hace uso de características relacionadas con las transiciones en los niveles de gris dentro de la imagen. Una vez identificada la estructura de interés en cada uno de los cortes de la secuencia CT, como resultado del proceso de segmentación previo, se generará la superficie mediante un proceso de triangulación entre puntos de contorno que definen la posición de la superficie en los diferentes cortes.

Capítulo 3

Reconstrucción de la geometría del hueso

3.1 Introducción

Gran parte de los métodos que podemos encontrar en la bibliografía para la segmentación 3D de estructuras anatómicas óseas a partir de imágenes emplean bien operadores basados en detección de bordes, bien operadores basados en regiones. Sin embargo, cualquiera de los métodos usados requiere una gran cantidad de interacción con el usuario para el control y corrección de los resultados de la segmentación [Ma y Overton, 1991; Pepino y col., 1993; Kankanhalli y col., 1993; Müller y Rüegsegger, 1995; Loncaric y col., 1995; Sonka y col., 1996]. Nuestra propuesta consiste en implementar un sistema automático para el modelado sólido de estructuras óseas, en concreto de la parte proximal de la tibia, a partir de imágenes de Tomografía Computacional (CT). El sistema pretende eliminar o minimizar la interacción del usuario experto con el propio proceso de segmentación y reconstrucción de la forma 3D de la estructura buscada. El objetivo perseguido es dotar al experto de una herramienta que le facilite tanto la evaluación de las lesiones y patologías presentes en dichas estructuras, lo que redundará en el proceso de diagnóstico, como la realización de una adecuada planificación quirúrgica.

El sistema global consta de tres partes, en las que discurren otras tantas etapas del proceso de reconstrucción. En un primera etapa se realizará la subdivisión de la escena en

estructuras óseas y fondo (*background*), eliminando las estructuras que no son de interés; su objetivo es obtener la superficie exterior de las estructuras óseas. Una segunda etapa distinguirá entre tejido óseo normal y zona de lesión. Tras una tercera etapa, y como resultado final, el sistema proporcionará un modelo sólido 3D de la tibia específico de cada paciente, aislada de los tejidos circundantes. Éste, con posterioridad, servirá de base para la definición de un modelo 3D de elementos finitos. Como ya hemos comentado, un aspecto importante en el diseño y evaluación de implantes y sus sistemas de fijación sobre estructuras óseas es el análisis de la distribución de tensiones en la estructura en aras a prevenir sus efectos negativos una vez sometida ésta a cargas, tanto estáticas como dinámicas [Ahmed y col., 1990]. La forma de abordar el estudio del campo de tensiones en el hueso es mediante su simulación tridimensional con elementos finitos, puesto que este método es capaz de representar la geometría irregular y la inhomogeneidad de estos tejidos [Huiskes y Chao, 1983; Tensi y col., 1989; Schereppers y col., 1990; Keyak y col., 1990, 1992; Clift, 1992]. El sistema que proponemos constituye pues un primer paso en el intento de disponer de una herramienta que de forma automática proporcione un mapa de la distribución de tensiones en un sistema implante/hueso en un paciente específico.

En este capítulo vamos a describir el proceso de obtención del contorno externo de la tibia en cada uno de los cortes (*slices*) CT. En los dos siguientes abordaremos la identificación de lesiones y la representación del modelo sólido 3D.

Dada la complejidad de las estructuras anatómicas, sobre todo teniendo en cuenta que éstas van a presentar lesiones o algún tipo de patología, la aplicación de cualquier método de bajo nivel no va a ser suficiente. Para obtener resultados fiables va a ser necesaria tanto la combinación de distintos métodos de segmentación, como la introducción de conocimiento específico en el proceso de segmentación y metodologías de verificación de resultados que permitan ajustar parámetros y/o cambiar estrategias de procesado. El sistema para la detección de la superficie externa de la tibia va a constar pues de varios módulos: uno basado en crecimiento de regiones, que se implementa mediante un sistema basado en reglas, otro basado en la detección de bordes, que se implementa mediante un esquema de contornos deformables, y un tercero de integración, en el que se intenta sacar partido de las características complementarias de los dos sistemas anteriores.

Los sistemas de análisis basados en regiones presentan la ventaja de mostrar una menor sensibilidad al ruido que los métodos que usan información de derivadas; además, su rendimiento no se ve demasiado afectado por el hecho de que la información de alta

frecuencia se pierda o sea poco fiable. Sus inconvenientes residen en la dificultad de introducir información de forma en el proceso de crecimiento de las regiones y en que los cambios en las características de los pixels que sugieren la presencia de un contorno pueden quedar diluidos por el carácter general de las propiedades de las regiones, lo que provoca con frecuencia la unión no deseada de regiones. Otro problema típico es la gran dependencia que tiene la segmentación resultante con el orden de evaluación de las partes de la imagen (pixels o subregiones). Adicionalmente, están los problemas de infra y sobresegmentación, difíciles de manejar en general. Los sistemas basados en reglas permiten incluir conocimiento de forma explícita en el proceso de segmentación, lo que ayuda a mitigar estos problemas, pero no les dan una solución definitiva.

Los sistemas basados en contornos presentan problemas de alta sensibilidad a ruido y de interferencia debida a bordes múltiples próximos. Esto provoca la posible extracción de falsos contornos y la no detección de trozos de contorno reales. De entre los métodos orientados a contornos, aquellos que han logrado un mayor rendimiento en el análisis de las imágenes biomédicas han sido los contornos activos (*snakes*). Sus ventajas residen en la facilidad para integrar las estrategias de control descendente guiada por modelos (*top-down*) y ascendente guiada por datos (*bottom-up*), pero siguen sufriendo los problemas de sensibilidad al ruido y de gran dependencia con las condiciones iniciales.

Aunque los métodos basados tanto en regiones como en bordes tienen sus ventajas e inconvenientes, éstos no son necesariamente idénticos; es decir, los problemas no les afectan de la misma forma. Aunque el ruido afecta a cualquier algoritmo de procesado de imágenes, los métodos basados en regiones presentan una menor sensibilidad que los métodos de gradiente. Lo mismo ocurre con la pérdida de información de alta frecuencia. Por contra, los esquemas basados en contornos activos manejan mejor las variaciones de forma. Además, puesto que los métodos basados en la detección de contornos dependen más de los cambios en los niveles de gris que de los valores de gris mismos, son menos sensibles a las distribuciones de la escala de grises. Los métodos basados en gradientes proporcionan, en general, una mejor localización en los límites de los segmentos. Todas estas razones ponen de manifiesto que la integración de ambos métodos debe proporcionar un mejor resultado que cada uno de ellos por separado, al poderse combinar los puntos fuertes de los métodos individuales, como se apuntó en los trabajos de Pavlidis y Liow [1990], Chu y Aggarwal [1993] y Chakraborty y col. [1996]. Otra ventaja adicional que nosotros consideramos en la integración es la posibilidad de flexibilizar los proced-

imientos de regiones y bordes, sacando la mejor ventaja posible del principio de mínimo compromiso, consistente en evitar la realización de operaciones drásticas siempre que sea posible.

Para el proceso de integración hemos desarrollado un esquema basado en el modelo de Campos Aleatorios de Markov. El módulo de integración toma la salida de un módulo de segmentación basado en reglas, que operan sobre regiones, y la salida de un módulo basado en contornos deformables y las conjuga mediante un esquema probabilístico en el que se tienen en cuenta probabilidades *a priori* sobre la pertenencia de una región a hueso, basadas en sus propiedades de densidad y varianza, y probabilidades condicionales basadas en las relaciones con regiones vecinas.

El sistema global que proponemos usa pues conocimiento del dominio de aplicación, bien de forma explícita en el proceso de segmentación basada en conocimiento, bien de forma implícita en la aplicación de modelos deformables y modelos de Campos Aleatorios de Markov. Esto proporciona un método de análisis flexible y eficiente.

En los apartados siguientes, tras analizar las características de las imágenes con las que vamos a trabajar, describiremos las distintas etapas del proceso de extracción de la geometría del hueso. Así, en primer lugar describiremos la etapa de preproceso, que nos va a proporcionar un conjunto de imágenes adecuado para abordar la reconstrucción 3D. Posteriormente describiremos los tres métodos de segmentación comentados: análisis de regiones basado en conocimiento, contornos activos e integración de ambos resultados mediante el formalismo de los Campos Aleatorios de Markov. El capítulo termina comentando los resultados obtenidos al procesar conjuntos de imágenes correspondientes a distintos pacientes.

3.2 Tipo y método de adquisición de imágenes

La modalidad de imágenes médicas con la que nosotros trabajaremos es la de tomografía computerizada (CT) de rayos X. Esta modalidad se basa en el hecho de que los diferentes tejidos del cuerpo presentan diferentes coeficientes de atenuación lineal a los rayos X. El logaritmo de la razón entre el número de fotones que entran al cuerpo y salen es una aproximación de la integral de línea del coeficiente de atenuación lineal para esa energía a lo largo de la línea entre fuente de radiación y detector. Registrando

los coeficientes de atenuación obtenidos para diferentes orientaciones se puede hacer una reconstrucción bidimensional de la sección radiada. En un sistema de tomografía computerizada, el almacenamiento de la información de los coeficientes de atenuación, así como su posterior procesado para la formación de la imagen, se realiza mediante un sistema de computación.

Las imágenes que tendremos que analizar serán pues imágenes de tomografía computerizada, que obtendremos mediante el esquema mostrado en la figura 3.1. Inicialmente se somete al paciente a una exploración CT. De este modo se obtienen imágenes de distintas secciones transversales de la rodilla a diferentes niveles; la figura 3.2 muestra la posición de los distintos cortes obtenidos en un paciente concreto. El sistema de exploración proporciona imágenes bidimensionales correspondientes a cada uno de los cortes realizados, imágenes que posteriormente se imprimen sobre placas radiográficas. Estas placas se digitalizan mediante un escáner y las imágenes digitales generadas se almacenan en el computador. Como las placas radiográficas incluyen las imágenes de varios cortes (figura 3.3), es preciso dividir la placa inicial en tantas partes como cortes contenga. Hecho esto, queda lista la secuencia de imágenes para su posterior procesado.

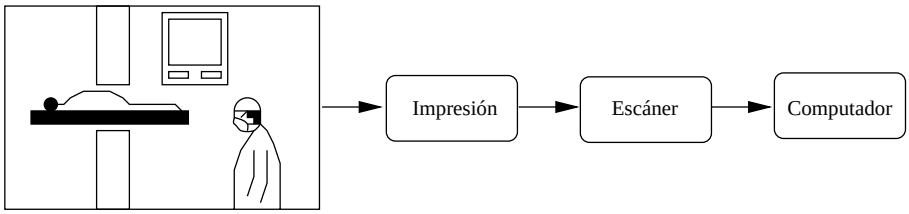


Figura 3.1: Sistema de adquisición de imágenes de CT.

A la baja relación señal ruido inherente a las metodologías de formación de las imágenes médicas hay que añadir tanto el ruido incorporado en el proceso de digitalización mediante el escáner, como los problemas de alineado al separar las imágenes de cada corte contenidas en la placa radiográfica, o los posibles movimientos del paciente durante la realización de la exploración, lo que obligará a una corrección en rotación y traslación de las imágenes. Todos estos elementos hacen que las imágenes finalmente disponibles sean difíciles de tratar.

Una exploración CT puede hacerse de diferentes formas, variando parámetros tales

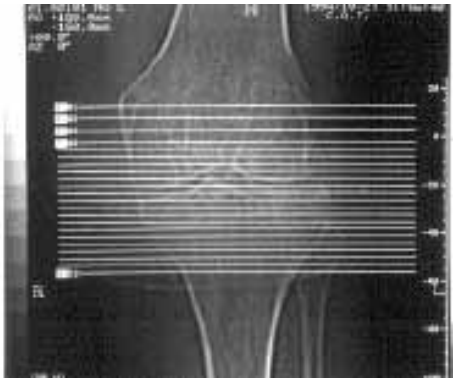


Figura 3.2: Secciones para las que se obtuvieron imágenes en una exploración CT particular.

como grosor del corte, tamaño del pixel o dosis de energía [Sutherland y col., 1986]. Para crear una reconstrucción anatómica tridimensional realista es necesario, en primer lugar, producir cortes CT con el menor grosor posible, dependiendo del tamaño de la estructura y el nivel de detalle deseado. En la mayoría de las aplicaciones ortopédicas será necesario un grosor de 1.5, 2.0 ó 3 mm, al menos, en la región de interés. Los inconvenientes para el trabajo con grosores menores residen tanto en la necesidad de someter al paciente a dosis altas de radiación, como en los mayores costes que implicaría. En general, la definición a lo largo del eje logitudinal es menor que la que se obtiene en el plano del corte. Las imágenes digitales finales presentarán una resolución espacial en el plano de 256×256 pixels y un tamaño de pixel del orden de 0.3 mm^2 sobre la imagen original.

Por otra parte, las imágenes se captan con 256 niveles de gris. A esta representación se llega después de dos transformaciones. La primera es la realizada a la hora de transformar los 4096 niveles de gris originales en los 256 representables por el sistema de impresión. En esta transformación el radiólogo determina la ventana sobre la cual se va a hacer la transformación lineal. La segunda transformación corresponde a la captación, mediante un sistema de digitalización, de las imágenes impresas en la placa radiográfica. Todavía queda un tercer factor por considerar, como es la dosis de radiación. De todas formas, aún es posible relacionar el número CT con la densidad aparente del hueso [Keyak y col., 1993; Kang y col., 1993]. La densidad aparente se calcula a partir del número CT después de un proceso de calibración lineal, a partir de dos puntos de referencia en cualquiera de los cortes CT. El primer punto es el número CT promedio del agua, que representa no-hueso

(0 gcm^{-3}). Para ello se emplea un tubo lleno de agua de un fantoma de calibración Cann-Genant. El segundo punto de calibración es el número CT promedio del hueso cortical, para el que se asume una densidad aparente de 1.8 gcm^{-3} [Keyak y col., 1990; Keyak y col., 1993]. Una vez evaluada una zona representativa de hueso cortical, ya se puede predecir la densidad aparente de cualquier punto del hueso mediante interpolación lineal de los números CT.

En las siguiente subsecciones se describe de forma más detallada el preproceso al que se someten todas las imágenes de forma sistemática con el fin de reducir el ruido, conseguir un buen apilamiento de los cortes y obtener un volumen de datos isótropo.

3.2.1 Apilamiento

Como hemos comentado, las imágenes con las que ha de trabajar el sistema se obtienen mediante la digitalización de placas radiográficas que contienen imágenes de varios cortes anatómicos. La figura 3.3 muestra una de estas placas digitalizada; se trata de una imagen de 1024×1480 pixels obtenida a partir de una placa original mediante un escáner Howteck Scanmaster 3+ con una resolución del orden de 0.3 $mm^2/pixel$.

Con el fin de obtener un apilamiento correcto de la secuencia de cortes en el eje Z hay que recortar adecuadamente la imagen correspondiente a cada uno de los cortes. Para ello es necesario establecer y localizar puntos de referencia fijos en cada una de las imágenes. Como tales puntos usaremos los caracteres L, R, P y A impresos en la placa y que marcan, respectivamente, la parte izquierda, derecha, inferior y superior de cada corte transversal. Una vez localizados estos caracteres dentro de cada una de esas imágenes, se procederá a determinar el ángulo de giro de la placa radiográfica midiendo el ángulo de desviación respecto de la horizontal de cualquier segmento que una el mismo carácter en distintos cortes situados en una misma banda, definida ésta como un intervalo en el eje Y. Dado que las imágenes de una exploración concreta se recogen en varias placas radiográficas, la determinación del ángulo de giro en cada una de ellas y su posterior corrección permitirá establecer un sistema de coordenadas común para toda la secuencia de cortes. Por consiguiente, calculado el ángulo de giro en cada placa, se procederá a su corrección y, finalmente, se recortará cada uno de los cortes sin más que indicar su tamaño y la posición de una esquina relativa a cualquier carácter identificado en ese corte. Todo este proceso se describe con mayor detalle a continuación.

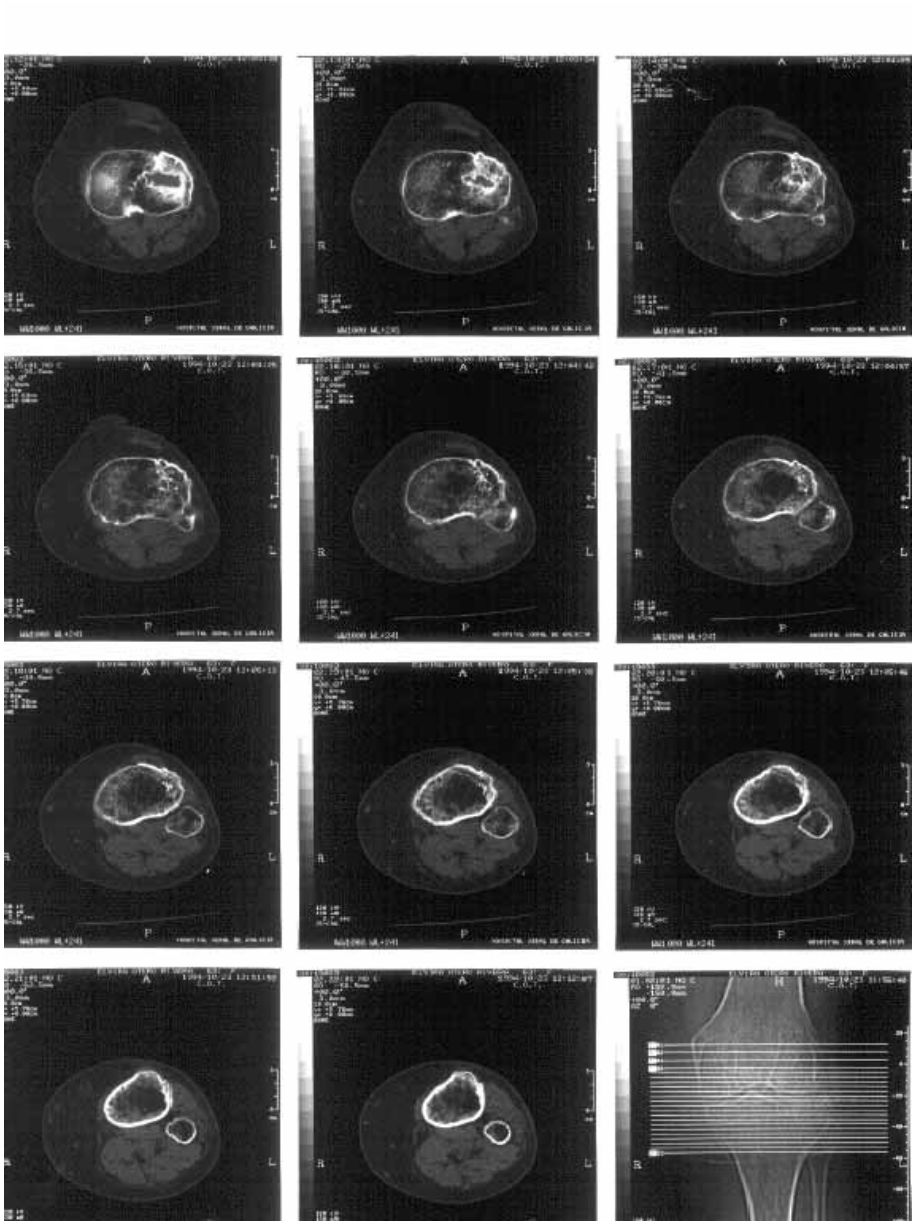


Figura 3.3: Imagen de una placa radiográfica digitalizada mediante un escáner de luz transmitida (1024 × 1480 pixels, 256 niveles de gris).

Para la identificación de cada carácter usaremos una medida de similaridad que tiene en cuenta tanto el parecido de cada carácter con los distintos patrones, como la posición de

los demás caracteres que reciban inicialmente una clasificación semejante. El parecido se establece mediante funciones de correlación normalizadas, ya que son menos dependientes de las propiedades locales de las imágenes de referencia y de entrada que su versión sin normalizar. El proceso de identificación de caracteres es como sigue: como la correlación es una función computacionalmente cara, inicialmente se correlaciona la imagen con sólo uno de los patrones, aquél con forma más genérica, que en nuestro caso consideramos que es la R. Tras la localización de los máximos de la función, y sobre una ventana de imagen centrada en ellos, se establece la correlación con los demás patrones. La medida de correlación empleada es de la forma [Ballard y Brown, 1982]:

$$C(c, c_i) = \frac{E(cc_i) - E(c)E(c_i)}{\sigma(c)\sigma(c_i)} \tag{3.1}$$

siendo:

$$\sigma(q) = [E(q^2) - (E(q))^2]^{1/2} \tag{3.2}$$

En las expresiones anteriores E es el operador valor esperado, c representa el carácter a clasificar y c_i los patrones característicos de las distintas clases: $\{c_1 = R, c_2 = L, c_3 = A, c_4 = P\}$.

Una vez obtenidas las medidas de correlación, cada carácter se asigna inicialmente a la clase con la que presenta mayor parecido (mayor valor de la función de correlación). Sin embargo, esta clasificación no es aún lo suficientemente precisa; para mejorarla realizamos un posterior proceso de clasificación iterativo. Para ello definimos una medida de similaridad de cada carácter con cada una de las clases de la forma:

$$M_{sim}(c, c_i) = a_1 C(c, c_i) + a_2 D(c, c_i) \tag{3.3}$$

donde a_1 y a_2 son constantes de ponderación, $C(c, c_i)$ es la medida de correlación normalizada y $D(c, c_i)$ es una medida de distancia al carácter más próximo con la misma etiqueta. Esta distancia se define de la siguiente forma:

$$D(c, c_i) = \max\{Dim_X, Dim_Y\} - \min\{|x_c - x_{c'}|, |y_c - y_{c'}|, \text{ tal que } c' \text{ tiene etiqueta } c_i\} \tag{3.4}$$

donde Dim_X y Dim_Y representan las dimensiones de la placa radiográfica. En este proceso iterativo, cada carácter se asigna a la clase con la que presente mayor similaridad. Ésta se determina calculando el máximo de las medidas $M_{sim}(c, c_i)$, con $i \in \{1, 2, 3, 4\}$ representando a cada uno de los patrones.

$$M(c) \quad = \quad \max_{c_i}\{M_{sim}(c, c_i) / \, c_1 = R, \, c_2 = L, \, c_3 = A, \, c_4 = P\} \tag{3.5}$$

En este proceso las medidas de $C(c, c_i)$ se calculan una sola vez, mientras que $M(c)$ se calcula de forma iterativa hasta que no haya variaciones en las asignaciones de patrones.

Una vez identificados los caracteres y agrupados por bandas según su coordenada en el eje Y de la imagen, se procede con la determinación del ángulo de rotación de cualquiera de las bandas. Escogemos para ello las correspondientes al carácter P, y considerando ahora los puntos de máxima correlación con estos caracteres, se puede determinar a partir de sus localizaciones el ángulo de giro de la imagen haciendo simplemente el promedio de las pendientes de las rectas que unen cada máximo de correlación con el más próximo. Una vez girada la imagen en ese valor medio, se determinan las nuevas posiciones de los máximos y se recorta la imagen correspondiente a cada corte dando su tamaño y el desplazamiento (*offset*) correspondiente a partir de la localización de las $P's$, o de cualquier otro de los patrones identificados. Como norma general, el tamaño de las imágenes recortadas es de 256×256 pixels, ya que este tamaño es suficiente para abarcar toda la información anatómica de cada corte.

3.2.2 Eliminación de información no anatómica

Una vez recortadas las imágenes, el siguiente paso consiste en la eliminación de toda aquella información no necesaria presente en la imagen (letras, ruido sobre el fondo, etc.) y mejorar la relación señal/ruido. Para ello se aplica un algoritmo para la detección, a partir de la propia imagen, del nivel de gris umbral entre el fondo de la imagen (*background*) y las partes anatómicas [Bae y col., 1992]. Una vez establecido éste, la zona de rodilla se separa del fondo de la imagen mediante un proceso de umbralización.

El cálculo del nivel umbral se basa en dos consideraciones: por una parte, la rodilla ocupa la parte central de la imagen, y por otra, los niveles de gris de la rodilla difieren sustancialmente de los del fondo. Examinaremos pues la distribución de los niveles de

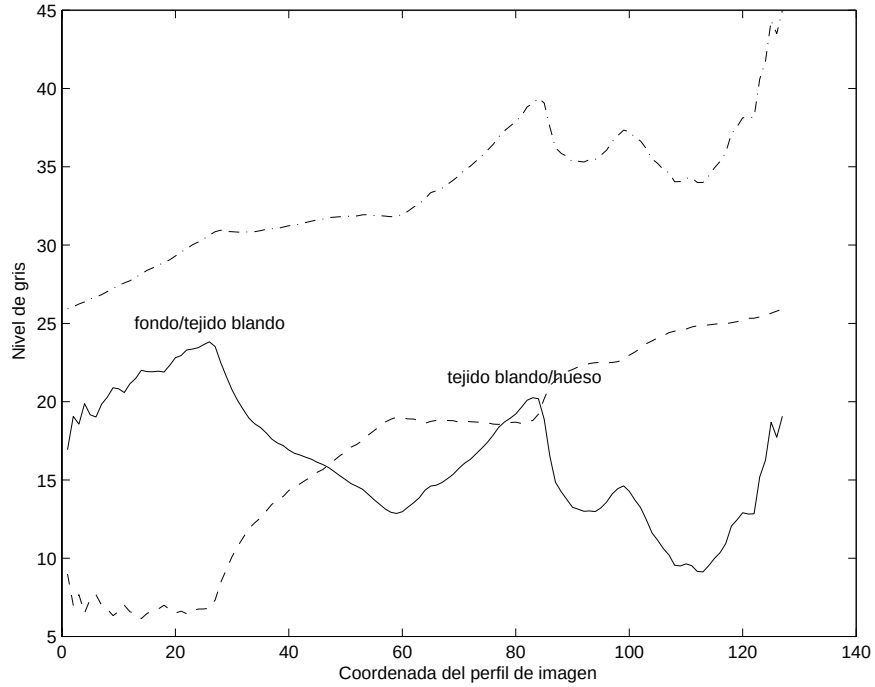


Figura 3.4: Determinación de separación entre fondo y región anatómica. Trazos discontinuos: - - : variación de valor medio acumulado del nivel de gris a lo largo de una línea horizontal trazada en la zona central de la imagen de la figura 3.5.a y recorrida en el sentido de borde a centro. -.-: igual información, pero calculada de centro a borde. El trazo continuo representa la diferencia entre las medias acumuladas.

gris de los pixels en una línea desde el borde al centro de una imagen. El umbral de valores de gris que da la máxima polaridad entre las regiones anatómicas y el fondo se puede determinar de la siguiente forma: se calcula la media acumulada del valor de gris los pixels desde el borde hasta el centro de la imagen. Puesto que la densidad de los pixels del fondo es mucho menor que la de los pixels de las regiones anatómicas, la densidad media es baja y relativamente constante hasta que se encuentran los pixels de rodilla. A medida que se incluyen los pixels más próximos al centro, los valores de la densidad media acumulada aumentan, produciéndose otro cambio brusco al entrar en estructura ósea. La línea en trazo discontinuo de la figura 3.4 muestra esta variación para una línea horizontal de la región central de la imagen mostrada en la figura 3.5.a. De la misma forma se calcula la densidad media acumulada en la dirección opuesta; es decir, del centro al

borde. La densidad media se mantiene a niveles constantes altos en la zona ósea y decrece al entrar en la zona de músculo, sobre la que se mantiene prácticamente constante hasta que se llega al fondo, en donde decrece rápidamente a medida que se incluyen pixels más próximos al borde. La línea representada mediante trazo-punto en la figura 3.4 ilustra este comportamiento. La diferencia entre ambas intensidades medias presenta dos máximos, uno correspondiente al contorno entre fondo y rodilla, y otro entre músculo y hueso, tal y como se puede observar en la línea de trazo continuo de la figura 3.4. Una vez detectados estos máximos, se calcula la densidad media en ambas zonas (fondo-tejido blando, tejido blando-hueso) del perfil de imagen considerado. Es ahora cuando se calcula el umbral TH como el promedio de las intensidades medias anteriores. Con este método se pretende reducir la sensibilidad frente al ruido.

La separación entre zonas de rodilla y fondo se podría obtener ahora mediante una simple operación de umbralización con el valor de umbral TH calculado. Sin embargo, en los casos reales, sobre el fondo van a aparecer números y letras, y son frecuentes los casos de nivel alto de ruido, por lo que implementaremos una variante del método anterior. Esta consiste en clasificar como fondo los pixels con niveles de gris menores a $TH - 0.1 * TH$ y como pixels de rodilla aquellos que tienen un nivel de gris mayor que $TH + 0.1 * TH$, dejando un margen de incertidumbre. Posteriormente realizamos un cierre morfológico de las regiones anteriores y unimos las regiones no clasificadas al vecino con el que comparten el mayor perímetro. A continuación buscamos las áreas más grandes correspondientes a rodilla y a fondo, y finalmente le incorporamos sus huecos. De esta forma desaparecen las letras del fondo y los huecos dentro de la rodilla. Obtenemos así una imagen binaria; en ella los pixels de valor "1"corresponden a zona de rodilla y los pixels de valor "0"corresponden a fondo. Una operación AND de esta imagen con la original da como resultado la eliminación de estructuras no anatómicas. La figura 3.5 muestra los resultados de este proceso; la figura 3.5.a muestra una imagen original con caracteres impresos en la esquina superior izquierda y un fondo no homogéneo; en la figura 3.5.b se muestra el contorno de la región anatómica resultado de aplicar el proceso descrito anteriormente, y que se superpone a la imagen original en la figura 3.5.c. Finalmente, la figura 3.5.d muestra la imagen conteniendo únicamente componentes anatómicas.

Para la eliminación de ruido aplicamos un filtro de mediana. Con ello se pretende regularizar la distribución de los niveles de gris sin la introducción de niveles no existentes y sin suavizar las transiciones fuertes entre regiones claramente diferenciadas.

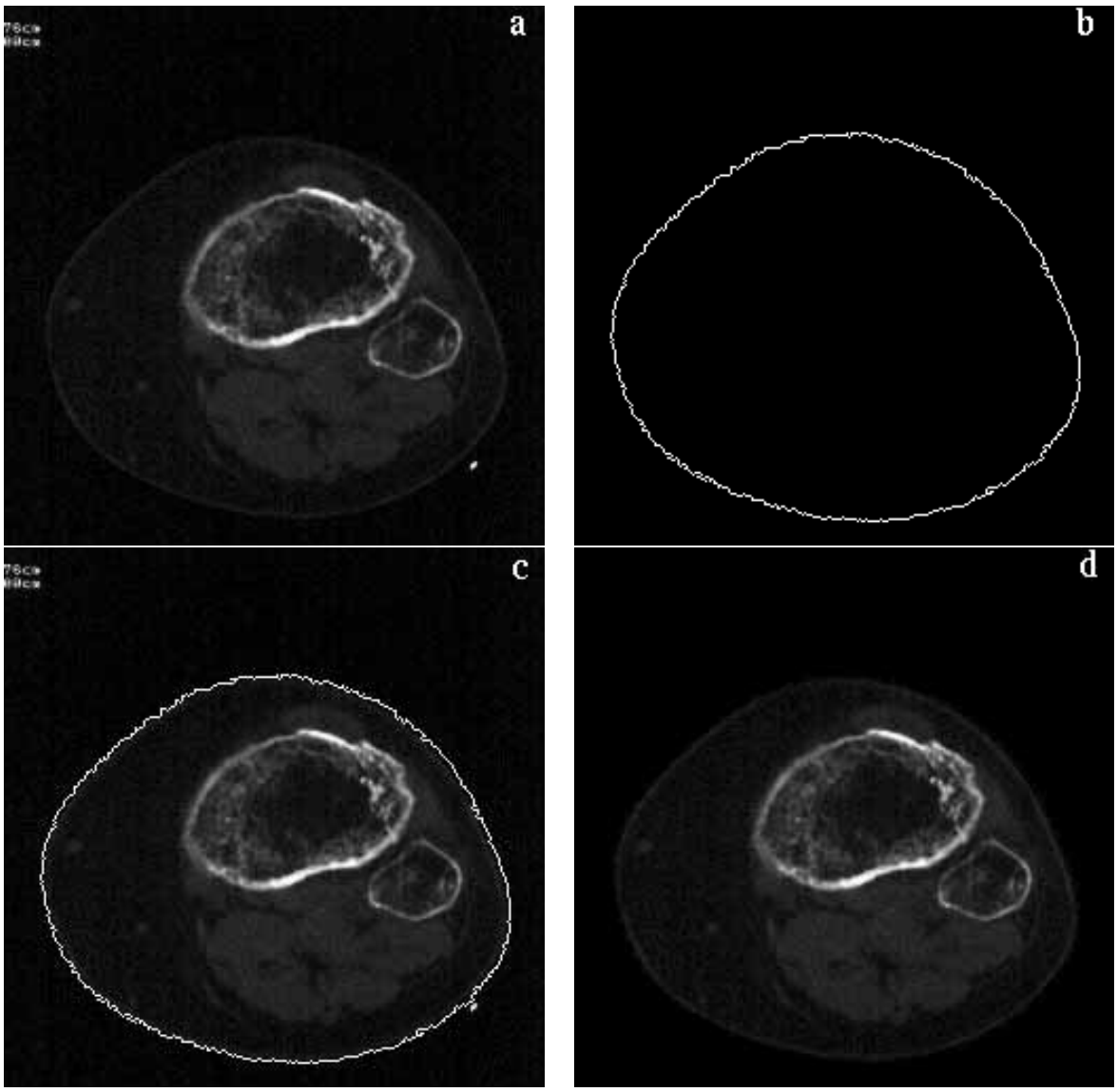


Figura 3.5: Ejemplo de preprocesado en el que se han eliminado las estructuras no anatómicas: (a): imagen original, (b): contorno de zona anatómica, (c): superposición del contorno sobre la imagen original, y (d): imagen con componentes no anatómicas eliminadas.

3.2.3 Corrección de desplazamientos

A parte del problema de recortar de forma adecuada los cortes a partir de la placa radiográfica, hay que resolver el problema de giros relativos entre los diferentes cortes. Esta posible causa de mal alineado de los cortes se debe a los movimientos del paciente durante el proceso de formación de las imágenes. Cuando éstos se producen, se observa un punto de discontinuidad entre dos cortes, ya sea por un fuerte desplazamiento en el centro de masas de la región anatómica de la imagen, ya sea por un giro, o lo más habitual, que sea por las dos causas a la vez. Esto constituye una fuente de error, ya que deja de ser cierta la continuidad 3D, y ahora este tipo de información es de baja fiabilidad. Dado que el proceso de segmentación se realizará corte a corte, propagándose los resultados de cada uno al siguiente, el error producido cuando se observa el movimiento del paciente puede arrastrarse durante el procesado de un número importante, o incluso el resto de cortes, conduciendo a una segmentación de baja calidad. Por otra parte, el apilado de cortes no seguiría el *eje* correcto y se obtendría una mala reconstrucción 3D.

Nuestro sistema puede detectar y corregir este tipo de problemas. Para ello será preciso realizar una correspondencia (*matching*) entre características significativas de cortes consecutivos. Dentro de los métodos que persiguen establecer correspondencias entre imágenes podemos distinguir entre aquellos que realizan la anotación de puntos prominentes (*landmarks*) y las técnicas volumétricas [Pun y col., 1994]. Las características geométricas clave usadas como puntos prominentes deben ser invariantes bajo traslación y rotación de los objetos y bajo deformaciones elásticas. Es decir, idealmente deben ser características intrínsecas de los objetos. Entre las características que se podrían usar podemos considerar las siguientes:

pixels : son numerosos y por tanto difíciles de *emparejar* entre cortes. Sin embargo su intensidad es invariante con la posición;

segmentos de líneas : se caracterizan por más atributos que los pixels (p.e. longitud, orientación), por lo que son más fáciles de *emparejar*. Sin embargo, no todos los contornos son invariantes;

regiones : son menos numerosas e incluso más fáciles de emparejar, pero en general más difíciles de extraer y generalmente no invariantes; y

puntos críticos : tales como vértices, puntos de inflexión, o cúspides. Son difíciles de extraer, pero intrínsecos a los objetos y, en consecuencia, características muy robustas para el establecimiento de la correspondencia.

Por otra parte, el proceso de establecimiento de correspondencia puede ser rígido o elástico. El primero es el más simple, puesto que la forma de los objetos se considera fija; los únicos grados de libertad son los que hacen referencia a su desplazamiento. El segundo es mucho más realista y complejo, puesto que tiene en cuenta las posibles variaciones de forma [Malandain y col., 1995; Tagare y col., 1995]. Es en el campo de la robótica, para aplicaciones industriales, donde el problema de establecimiento de correspondencias está mejor resuelto [Grimson, 1990; Ayache, 1991; Besl y McKay, 1992; Klaus y Horn, 1997]. Sin embargo, en las imágenes médicas 3D los datos de entrada no tienen la regularidad geométrica ofrecida por los objetos manufacturados, por lo que el problema resulta mucho más complejo. En la bibliografía podemos encontrar distintos planteamientos, tanto para correspondencia rígida como para elástica. Así, en Malandain y col. [1995] se puede encontrar una amplia revisión de estas técnicas para aplicaciones médicas.

En nuestro caso el problema es el de lograr un buen apilamiento de los cortes en aras a obtener una buena reconstrucción 3D de las estructuras anatómicas de interés. Por tanto, las características a considerar deben ser de muy fácil extracción y con variaciones muy ligeras entre cortes consecutivos. Una característica que cumple estas condiciones es la máscara de la parte anatómica de la imagen, cuya forma de obtención ya explicamos en la subsección anterior. Además, como los cortes están muy próximos (de 1 a 3 mm), las variaciones en la forma de esta máscara entre cortes adyacentes es pequeña. De este modo usamos una característica de fácil extracción y correspondencia rígida, más simple que la elástica. Para caracterizar esta región utilizaremos el centroide para corregir traslaciones, y los segundos momentos centrales para tener en cuenta posibles rotaciones. El centroide y la orientación de la máscara de la parte anatómica de la imagen, medidos corte a corte, no permanecen constantes a lo largo del eje Z, ni siquiera presentarán variación lineal, pero podremos aproximarlas, sin errores importantes, por funciones lineales, extrapolando linealmente al corte actual sus trayectorias en unos pocos cortes anteriores.

Por consiguiente, para la detección y corrección de los desplazamientos determinaremos el centroide y los segundos momentos centrales de la parte anatómica de la imagen en todos los cortes y después buscaremos variaciones bruscas entre cortes adyacentes. El

centroide $M(x_{av}, y_{av})$ y los segundos momentos (μ_{20}, μ_{02}) de la máscara anatómica en cada corte se obtienen mediante las siguientes ecuaciones:

$$x_{av} = m_{10}/m_{00} \tag{3.6}$$

$$y_{av} = m_{01}/m_{00} \tag{3.7}$$

$$\mu_{pq} = \sum (x - x_{av})^p (y - y_{av})^q \tag{3.8}$$

donde:

$$m_{rs} = \sum x^r y^s \tag{3.9}$$

con $r, s \in \{0, 1\}$, $p, q \in \{0, 2\}$ y los sumatorios extendidos a todos los pixels de la máscara. m_{10} y m_{01} representan la acumulación de las coordenadas x e y, respectivamente, de los pixels de la región y m_{00} su área.

Una forma de calcular la orientación de una región a partir de sus ejes principales es a través de la siguiente ecuación:

$$\tan 2\phi = 2m_{00}/(\mu_{20} - \mu_{02}) \tag{3.10}$$

donde ϕ es el ángulo de inclinación de los ejes principales respecto de los canónicos.

El par de características M y ϕ se calcula para todos los cortes de la secuencia. A continuación se buscan las variaciones bruscas respecto del comportamiento lineal. Para ello se considerarán los dos cortes anteriores al actual y se extrapolarán linealmente los valores de las características anteriores para el corte actual, obteniéndose de esta forma los valores M^{ext} y ϕ^{ext}. En el punto en el que se detecte una variación brusca entre valores extrapolados y valores reales se considerará que durante la toma de la imagen de ese corte el paciente se movió. Se considerará que existe un cambio brusco cuando la diferencia entre alguno de los parámetros respecto de los obtenidos para el corte previo es mayor del doble de la diferencia obtenida para el valor extrapolado. Así, consideraremos que los cortes previos corresponden a la posición correcta y los siguientes estarán todos desplazados en la misma medida. Por tanto, cada vez que se detecte un desplazamiento de este tipo, habrá que aplicar la corrección a los cortes siguientes (según el orden de procesado). Lo más adecuado será detectar y cuantificar todos los desplazamientos y

hacer una única corrección al final, según el valor acumulado. La nueva posición de los cortes obtenidos con posterioridad al desplazamiento habrá que corregirla en la siguiente medida:

$$T = M_i - M_i^{ext} = M_i + 2M_{i-1} - M_{i-2} \quad (3.11)$$

$$\Delta\phi = \phi_i - \phi_i^{ext} = \phi_i + 2\phi_{i-1} - \phi_{i-2} \quad (3.12)$$

donde T representa la traslación y $\Delta\phi$ la rotación. Estos parámetros se ilustran en la figura 3.6. En las ecuaciones anteriores M representa el centroide, ϕ la orientación, y el subíndice i hace referencia al número del corte en el que se detectó el desplazamiento como cambio brusco en la variación lineal del centro de masas o de la orientación. Para cada pixel (x, y) del corte j, siendo $j \geq i$, la nueva posición $(x^{'}, y^{'})$ será la correspondiente a los enteros más próximos a

$$(r, s) = (T_x + x\cos\Delta\phi - y\sin\Delta\phi, T_y + x\sin\Delta\phi + y\cos\Delta\phi) \quad (3.13)$$

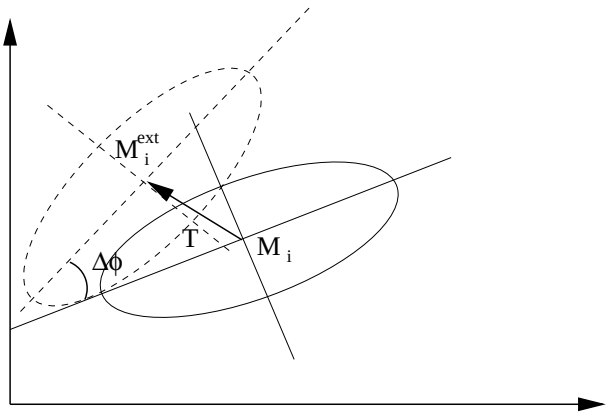


Figura 3.6: Parámetros de rotación y traslación: los trazos continuos representan la posición real y los trazos discontinuos representan la posición extrapolada.

La figura 3.7 muestra un ejemplo de la corrección del movimiento en un corte. La figura 3.7.a corresponde a un corte de la secuencia y la figura 3.7.b a un corte adyacente, desplazado y rotado. Este corte presenta respecto del anterior un desplazamiento del

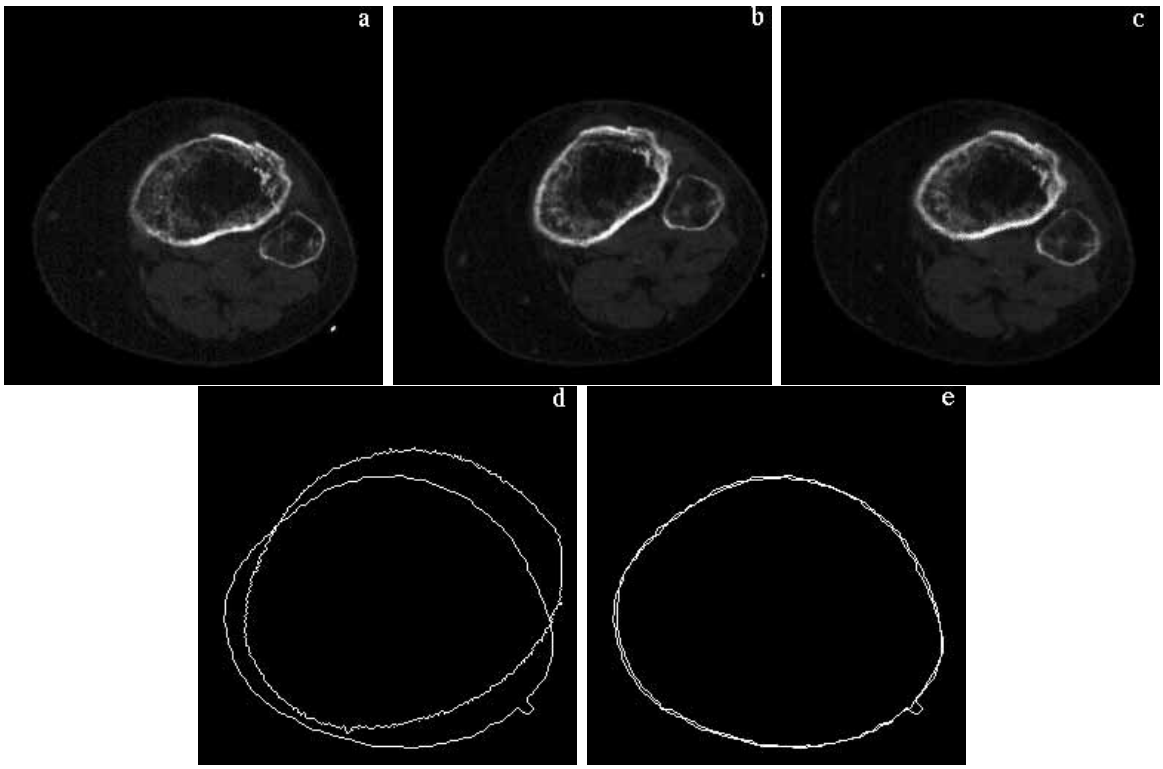


Figura 3.7: Ejemplo de corrección de posición. El corte (a) corresponde al corte anterior en la secuencia, el (b) al corte desplazado y el (c) al resultado de la corrección. La figura (d) muestra la superposición de los contornos anatómicos de (a) y (b), y la figura (e) muestra la superposición para (a) y (c).

centroide de 10 pixels en el eje X y de 15 pixels en el eje Y, y una rotación de 20^o. La traslación y el giro relativos al corte anterior se observan mejor en la figura 3.7.d, donde se muestran los contornos de las dos máscaras de las regiones anatómicas. En la figura 3.7.c se muestra el resultado de la corrección del desplazamiento relativo, que se constata en la figura 3.7.e. El rendimiento del algoritmo es completamente independiente de las magnitudes del desplazamiento y funciona correctamente en cualquier caso.

3.2.4 Interpolación: generación de imágenes intermedias

Para la visualización y análisis en 3D los cortes bidimensionales se registran en una malla 3D. Puesto que la separación entre dos cortes consecutivos es mayor que la separación entre los centros de dos pixels adyacentes dentro de un corte, serán necesarias operaciones de interpolación con el fin de producir una adecuada representación de datos, tanto desde el punto de vista geométrico como desde el punto de vista densitométrico. En la bibliografía podemos encontrar diversas técnicas de interpolación basadas en niveles de gris o basadas en la forma [Ballard y Brown, 1982; Parrott y col., 1993]. Los métodos basados en los niveles de gris usan los valores de intensidad de los pixels. Estos métodos estiman el valor asociado con un pixel (voxel) p mediante la interpolación lineal de los valores asociados con los pixels de la imagen en una vecindad de p en cada uno de los cortes entre los que se interpola. Un método de éstos típico selecciona un conjunto de puntos característicos de cada imagen, determina las correspondencias y utiliza éstas para obtener la función de transformación [Goldwasser y col., 1986; Wood, 1986; Levoy, 1988; Lin y col., 1989]. Por su parte, los métodos basados en la forma usan el contraste en los contornos entre diferentes materiales para realizar la interpolación. Estas técnicas empiezan por identificar una estructura de interés en cada corte mediante operadores de segmentación. De esta forma se crea un volumen de datos binarios en donde cada voxel tiene un valor 0 ó 1 según esté dentro o fuera de la estructura. Usando la imagen binaria se calcula la distancia de cada voxel al contorno, con el convenio de que las distancias de puntos internos son positivas y la de los externos son negativas. Estas distancias son después interpoladas punto a punto [Raya y Udupa, 1990; Udupa y Herman, 1991; Herman y col., 1992].

Las técnicas de interpolación por niveles de gris provocan un efecto de sombreado de un corte sobre otro cuando las diferencias entre imágenes son grandes; a causa de ello no es posible proporcionar contornos bien definidos en los cortes interpolados. A su vez, las técnicas basadas en la forma, aunque proporcionan un mejor contorno que las anteriores, tienen el inconveniente de la fuerte dependencia con la segmentación de las imágenes originales, de tal forma que si los contornos obtenidos no son los adecuados, se pueden obtener estructuras irregulares o deformadas en la representación 3D [Dhawan y Arata, 1991]. Además, nuestro sistema se basa en la proximidad y continuidad en las formas de las estructuras a detectar en cada corte con el fin de disponer de un modelo aproximado de contorno que guíe la segmentación. Por este motivo, la necesidad de la interpolación

es previa a la segmentación. Es por ello por lo que decidimos emplear técnicas basadas en los niveles de gris.

Partiendo ya de que las imágenes se encuentran perfectamente apiladas, un método típico consiste en la selección de puntos característicos para cada imagen, determinar la correspondencia entre ellos, y utilizar estas correspondencias para determinar una función de transformación que pueda hacer la proyección (*mapping*) de una imagen sobre otra.

El método de interpolación que aquí utilizaremos se basa en el de Goshtasby y col. [1992], es totalmente automático, y emplea aquellos pixels con magnitud de gradiente mayor que un umbral como puntos característicos. La correspondencia entre estos puntos en las imágenes se establece en base a propiedades espaciales de las imágenes tomográficas. Éstas tienen varias propiedades en común que relatamos a continuación. En primer lugar, las imágenes no tienen diferencias de escalado, y la correspondencia de un punto de una de las imágenes de referencia cae en un área pequeña centrada en las mismas coordenadas en la otra imagen. Además, hay continuidad en las correspondencias; es decir, puntos vecinos en una de las imágenes de referencia se proyectan en puntos vecinos de la otra imagen de referencia. Puede ocurrir que un punto de una imagen se proyecte en varios de la otra imagen, debido a la expansión del tejido, y que varios puntos sean la proyección de uno sólo, debido al efecto de contracción, pero en cualquier caso se mantiene la continuidad en las correspondencias. La excepción la constituyen los casos de la desaparición de un tejido entre cortes. Por último, la diferencia geométrica entre cortes consecutivos es local. Por lo tanto, no se pueden aplicar funciones de transformación globales para la puesta en correspondencia; ésta debe tener lugar a nivel local.

La selección de los puntos característicos se hace en base a la magnitud del gradiente. Para el posterior establecimiento de la correspondencia entre pixels característicos de dos cortes diferentes se evalúa el siguiente vector de coste:

$$\begin{aligned} C(x,y,x',y') &= u_1 \times [I(x,y) - I'(x',y')]\vec{i} + \\ &\quad u_2 \times [D(x,y) - D'(x',y')]\vec{j} + \\ &\quad u_3 \times [\theta(x,y) - \theta'(x',y')]\vec{k} + \\ &\quad u_4 \times \sqrt{(x-x')^2 + (y-y')^2}\,\vec{l} \end{aligned} \tag{3.14}$$

donde $I(x,y)$, $D(x,y)$, $\theta(x,y)$ son, respectivamente, la intensidad, la magnitud del gradiente y la dirección del gradiente del punto (x,y). Con (x,y) se hace referencia a puntos

característicos de una de las imágenes (I) y con ($x^{'}, y^{'}$) se hace referencia a los de la otra imagen ($I^{'}$).

Para determinar la correspondencia entre puntos característicos primero se evalúa la proyección en el sentido $I \rightarrow I^{'}$, y después en el contrario, $I^{'} \rightarrow I$. En cada caso se evalúa el módulo del vector de coste para obtener el punto de proyección. Esta evaluación se realiza sobre todos los puntos característicos de la otra imagen situados dentro de una ventana (10×10) centrada en el pixel actual. Evaluada esta proyección en los dos sentidos, nos quedaremos sólo con las correspondencias consistentes. Esta prueba de consistencia se hace para eliminar errores provocados por la existencia de regiones homogéneas relativamente grandes, en las cuales la correspondencia (*matching*) no resulta muy fiable.

La figura 3.8 muestra la puesta en correspondencia entre pixels característicos A y B de dos imágenes entre las cuales se va a generar un conjunto de imágenes interpoladas; el segmento a trazos que une estos dos puntos marca la posición que corresponde a ese pixel en los cortes interpolados, cuya intensidad se obtendrá por interpolación lineal de la de los pixels A y B. La proyección de los puntos no característicos y de aquellos para los que no se obtiene una correspondencia consistente entre las dos imágenes se hace por interpolación bilineal de las correspondencias consistentes obtenidas para los cuatro puntos característicos más próximos no colineales. Establecida esta correspondencia, las imágenes intermedias se crean determinando la nueva posición de los pixels y su valor de intensidad como promedio ponderado (en relación con la distancia en la 3ª dimensión) entre los puntos correspondientes de las imágenes de referencia. Para eliminar posibles errores individuales en las asignaciones se aplicará con posterioridad un filtro de mediana (3×3) sobre los vectores de desplazamiento que definen las correspondencias; con ello se asegura la continuidad en la correspondencia (vecindad con vecindad). Puede encontrarse una valoración del rendimiento del sistema en el trabajo de los autores iniciales [Goshtasby y col., 1992].

En la figura 3.9 se muestra un ejemplo de la generación, mediante este método, de 4 imágenes interpoladas entre dos reales. En este caso las imágenes originales corresponden a cortes obtenidos con 3 mm de separación a lo largo del eje Z, mientras que la resolución en el plano XY era de 0.58 mm de lado cada pixel. En este ejemplo se puede observar como las imágenes intermedias muestran la transición de forma entre las imágenes originales.

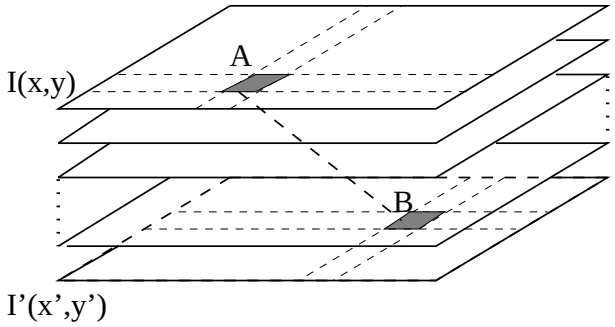


Figura 3.8: Puesta en correspondencia entre los pixels A y B de dos imágenes para la generación de imágenes intermedias interpoladas. El segmento que los une marca la posición del punto interpolado en los cortes intermedios generados.

Por otra parte, la interpolación no sólo será precisa cuando la distancia entre cortes sea muy grande; puede serlo también cuando la diferencia de formas se haga muy acusada entre dos cortes. Esto suele ocurrir en puntos de bifurcación o cuando el hueso presenta grandes deformaciones o lesiones.

3.3 Características de las imágenes CT de rodilla

En la introducción a este capítulo realizamos una serie de consideraciones de carácter general sobre el comportamiento de los métodos de segmentación basados en regiones o en bordes. Comentemos ahora las características de las imágenes con las que hemos de tratar para poner de relieve las dificultades que entraña la obtención de una segmentación correcta. Para ello nos ayudaremos con la figura 3.10, en la que se presenta un esquema identificando tibia y peroné y sendos cortes transversales a distintas alturas.

Nuestro propósito consiste en obtener de forma automática y precisa el contorno exterior del hueso cortical en todos los cortes. Como ya podríamos suponer, el ancho y la alta densidad del hueso cortical, así como la clara separación entre tibia y peroné, hacen que la segmentación de la tibia en la parte distal no ofrezca excesivos problemas. Sin embargo, a medida que nos aproximamos a su extremo proximal, la anchura del hueso cortical decrece y la distancia al peroné se reduce. Cuando además existan lesiones, nos encontraremos

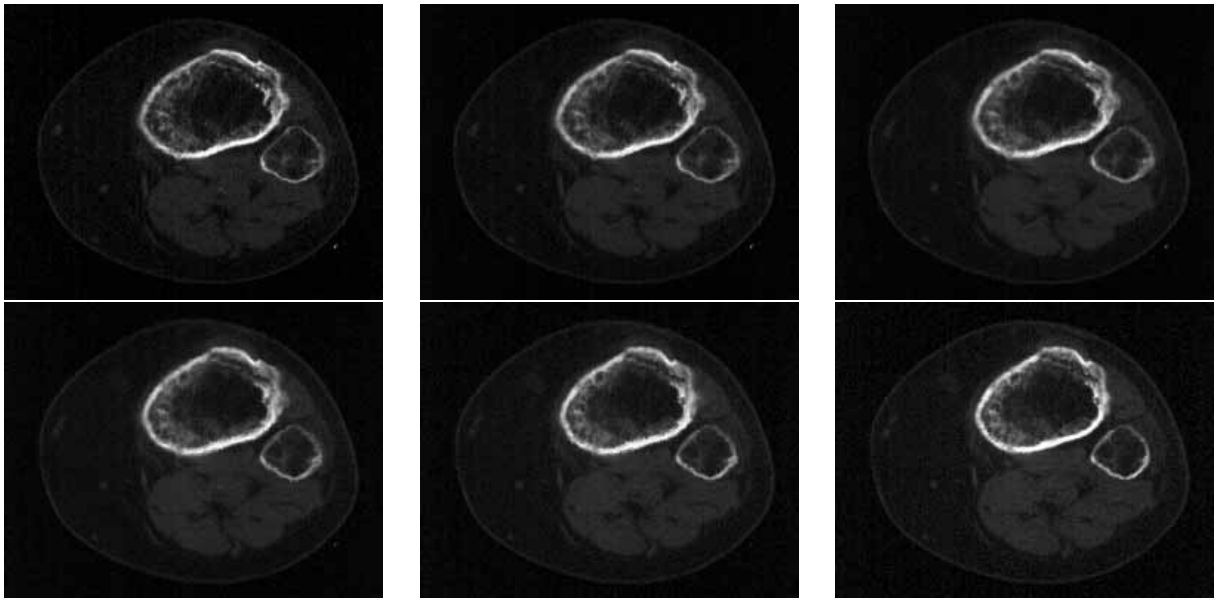


Figura 3.9: Ejemplo de generación de imágenes interpoladas. La primera y la última son las imágenes originales.

con problemas de pérdida ósea, estrechez del hueso cortical, malformaciones, fracturas, etc. Esto se traduce en contornos exteriores del hueso probremente definidos, mínima separación entre contornos internos y externos del hueso cortical y/o proximidad entre contornos de diferentes estructuras (tibia y peroné, tibia y fémur). La figura 3.10.a ilustra un caso con varios de estos problemas. Estas consideraciones implican varios problemas en la aplicación de los métodos de segmentación antes comentados. Así por ejemplo, el comportamiento de los contornos activos quedará fuertemente influenciado por la existencia de ruido y texturas en el fondo y dentro de las estructuras (estructura trabecular del hueso), por el hecho de que las estructuras cuyos contornos queremos determinar puedan estar muy próximas a otras, o por el hecho de que la aproximación inicial esté localizada dentro de una estructura que presenta dobles contornos (hueso cortical). En estas situaciones el contorno puede quedar atrapado por puntos de contorno espúreos, bordes internos o contornos de estructuras próximas. Esto hace que el resultado final sea muy sensible a las condiciones iniciales.

En general, los métodos basados en contornos deformables, aunque han demostrado una gran eficiencia en tareas de segmentación de estructuras médicas, no están libres

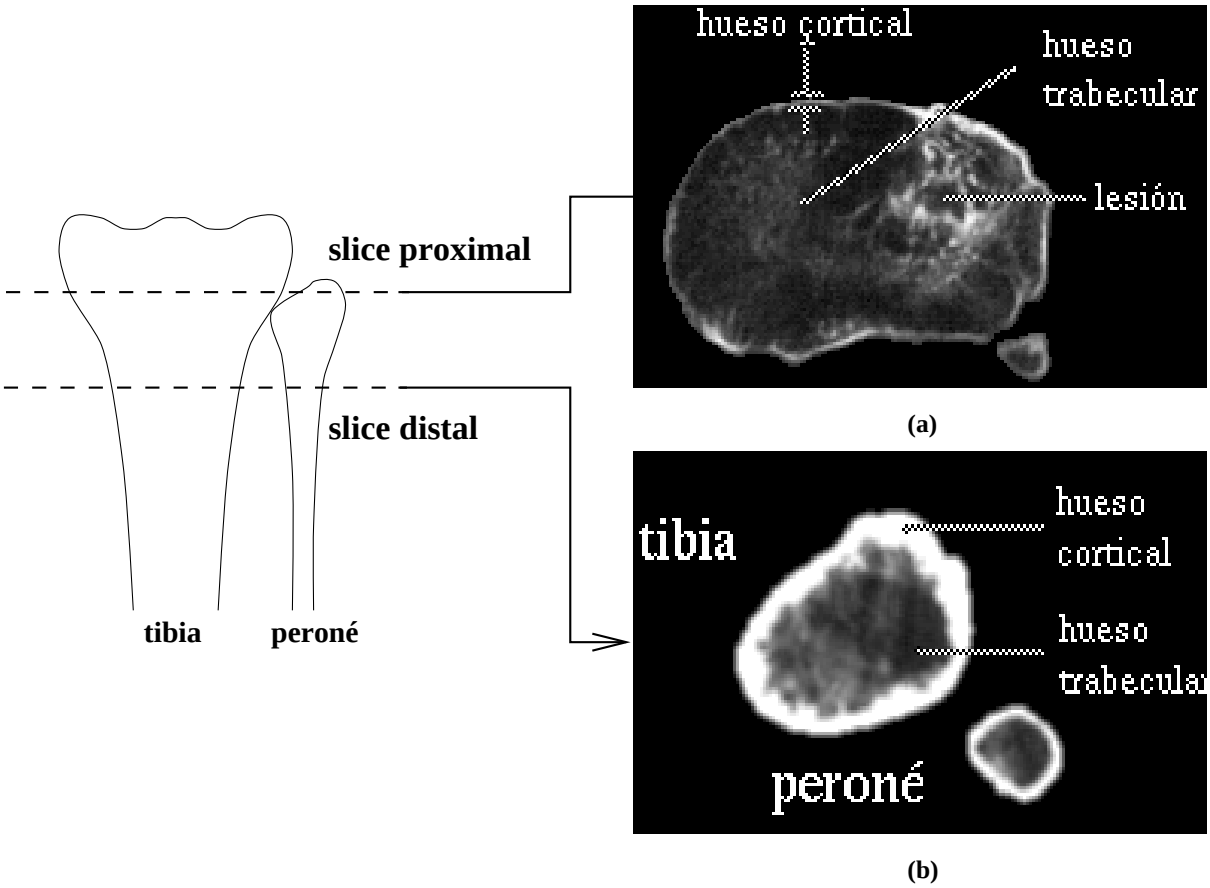


Figura 3.10: Esquema de los huesos de la parte baja de la rodilla: (a) sección transversal en zona proximal; (b) sección transversal en zona distal.

de limitaciones. Así, estos métodos siguen sufriendo problemas de dependencia con los contornos iniciales. La mayoría de los algoritmos basados en los modelos de contornos activos pueden manejar objetos con geometría y topología simples, pero resultan muchas veces inadecuados para objetos con cavidades profundas y para estructuras compuestas de diferentes partes [McInerney y Terzopoulos, 1995]. Estos inconvenientes aparecen con frecuencia al tratar estructuras óseas con lesiones.

Por otra parte, aún cuando los sistemas de análisis basados en regiones presenten la ventaja de mostrar menor sensibilidad al ruido, también presentan inconvenientes. Éstos residen básicamente en la dificultad para introducir información de forma en el proceso

de crecimiento de las regiones y en que los cambios en las características de los pixels que sugieren la presencia de un contorno pueden quedar diluidos por el carácter general de las propiedades de las regiones, lo que provoca con frecuencia la unión no deseada de regiones. Las características ya comentadas de las imágenes CT de rodilla provocarán que estos métodos presenten problemas de incapacidad de separación de estructuras óseas próximas, de dislocación de los contornos de los huesos y no detección de hueso cortical muy estrecho, pudiéndose provocar en ciertos casos uniones con tejidos exteriores al hueso. Esta infrasegmentación se pueden paliar con procesos de división que extraen información a partir de un mapa de bordes. Sin embargo, en zonas de lesión o de poca distancia entre estructuras la información de bordes puede perderse o resultar incompleta debido al efecto de interferencia entre transiciones reales próximas, por lo que no es posible efectuar una división precisa.

Una vez que todas las imágenes han sido sometidas a la etapa de preprocesado consistente en el apilamiento de los cortes, eliminación de la información no anatómica, mejora de la relación señal ruido y obtención de un volumen de datos isótropo mediante la generación de cortes intermedios, se está en condiciones de abordar el proceso de delimitación del contorno exterior de la tibia en cada uno de los cortes. Como ya hemos comentado, el análisis se realizará corte a corte, propagándose los resultados a los cortes adyacentes. Como apuntan algunos autores, de la integración de la información proporcionada por métodos de regiones y métodos de contornos podrá resultar una optimización de la segmentación [Pavlidis y Liow, 1990; Chu y Aggarwal, 1993; Pankanti y Jain, 1995]. Así pues, se abordará la detección de los contornos de las estructuras de interés mediante la intervención de tres módulos: uno de unión de regiones, otro basado en contornos deformables y un tercero de integración de los resultados de ambos métodos. En las siguientes secciones describiremos cada uno de ellos.

3.4 Segmentación mediante unión y división de regiones

Este módulo implementa un sistema de crecimiento de regiones basado en conocimiento. La idea básica es la siguiente: para cada corte se parte de una sobresegmentación de la imagen creada por un clasificador basado en redes neuronales [Cabello y col., 1993]. A

continuación se extraen propiedades morfométricas y densitométricas de las regiones resultantes y se crea un grafo de adyacencias de regiones (RAG) en el que los pesos de cada arco que une dos regiones representan el tamaño del contorno compartido. En la siguiente etapa se procede con el crecimiento de regiones a partir de núcleos formados por regiones de gran densidad localizadas dentro del área de interés de la imagen. El crecimiento está guiado por un conjunto de reglas que hacen referencia a criterios como compactación, cierre, proximidad espacial y proximidad a un modelo aproximado representado por la segmentación del corte previo ya procesado. Tras el proceso de crecimiento puede ocurrir que no exista una adecuada correspondencia entre modelos y objetos en la escena debido, por ejemplo, a la unión entre tibia y peroné en zonas de separación muy estrecha. En esos casos se recurre a información de bordes. Esta información en muchos casos permitirá la división de la región, lo que llevará a la separación de las estructuras. Como el hueso es una estructura constituida por tejidos de características muy diversas, las propiedades más importantes en la operación de crecimiento son las que hacen referencia a forma más que a características de homogeneidad [Pardo y col, 1995a, 1995b].

El procesado se divide en dos etapas según el nivel de conocimiento involucrado: bajo nivel y alto nivel. En la etapa de bajo nivel se particiona la imagen en un conjunto grande de regiones, que se describen mediante propiedades densitométricas, morfométricas y relacionales, y se extrae información de bordes. En la etapa de alto nivel se realiza la interpretación de las regiones segmentadas mediante procesos de crecimiento y división, de acuerdo con sus propiedades específicas y modelos aproximados, usando una base de conocimiento. En la figura 3.11 se muestra el esquema del sistema global en el que se representan los módulos de bajo y alto nivel y sus tareas asociadas, así como la información leída y/o escrita por cada uno de estos módulos en una base de datos y una base de conocimiento sobre el dominio.

En la segmentación de bajo nivel se adopta un planteamiento conducido por datos (*data driven*), puesto que el ruido y la natural diversidad de las estructuras anatómicas, sobre todo debido a la presencia de deformaciones o lesiones, hacen que las variaciones en la imagen de entrada sean la consideración más importante. El conocimiento en esta fase es independiente del dominio; el objetivo del procesado es la creación de regiones (entidades de mayor nivel de abstracción que los pixels individuales) en base a su homogeneidad y el establecimiento de relaciones entre ellas. Como resultado se obtiene una descripción más elaborada de la escena que incluye un conjunto de relaciones entre elementos funda-

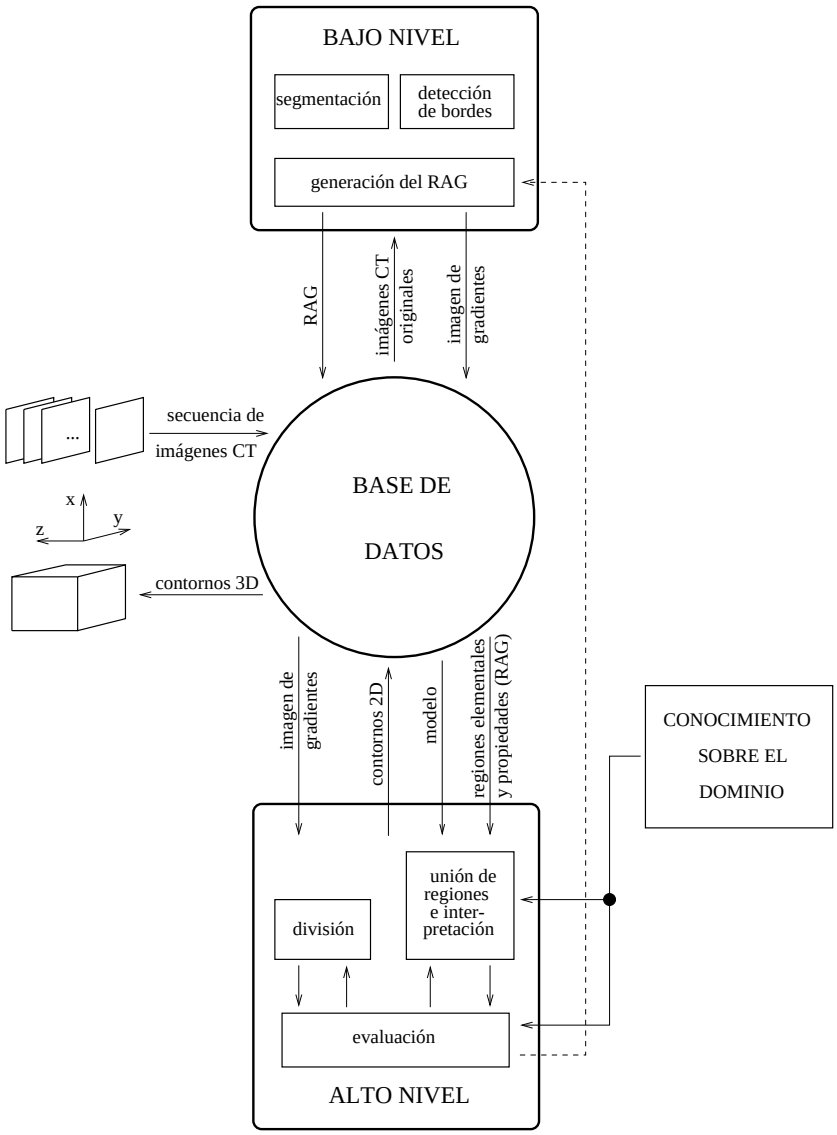


Figura 3.11: Esquema global del sistema de crecimiento y división de regiones.

mentales de la imagen. En la segunda fase, bloque de alto nivel, se aplica conocimiento dependiente del dominio relativo a forma y relación entre estructuras anatómicas para interpretar la imagen segmentada. El sistema reconoce aquí tres clases de objeto: tibia, peroné y fondo. El conocimiento se formaliza en forma de reglas de producción. Las motivaciones para el empleo de un sistema basado en reglas para el crecimiento de regiones

son las siguientes [Raya, 1990]:

1. La representación del conocimiento a través de reglas de producción ofrece una gran flexibilidad en la codificación de conocimiento proveniente de diferentes fuentes.

2. Los sistemas basados en reglas proporcionan una estrategia de control eficiente consistente en la ordenación de reglas y la selección del camino dentro de la imagen.

3. El sistema se puede extender fácilmente con la inclusión de nuevas heurísticas o la mejora de las existentes.

El sistema basado en reglas consta pues de una base de conocimiento sobre el dominio y una base de datos que incluye las imágenes originales, las resultantes de los procesos realizados y las propiedades calculadas. El sistema incluye también una componente de control para determinar las acciones a ejecutar. El conocimiento a priori sobre la anatomía y el conocimiento específico adquirido en el proceso de reconstrucción se implementan como un conjunto de reglas y propiedades cuya aplicación permitirá ir reconstruyendo el hueso corte a corte a modo de transición de estados, donde el estado actual se representa mediante el corte previo ya procesado, la entrada actual por el corte actual y la salida la constituirá el etiquetado final de las regiones del corte actual.

El sistema realiza el procesado corte a corte, ya que la desigual resolución de la exploración CT provoca discontinuidades en la escena de salida que restringen el uso directo de verdaderos operadores 3D [Raya, 1990]. Por ello nuestro sistema identificará el espacio 3D ocupado por el hueso haciendo uso sólo de una vecindad en 2D. Esto implica que se está realizando el crecimiento de regiones 2D dentro de cada sección de la rodilla; el modelo 3D se obtendrá mediante el apilamiento de contornos en el eje Z. Esto requiere la búsqueda de localizaciones independientes dentro de cada sección transversal por las que iniciar el proceso de crecimiento. En cualquier caso, se hace uso de información 3D resumida en la segmentación obtenida para el corte adyacente ya procesado.

Un problema importante de la segmentación es el de la no unicidad de la solución. Idealmente, una segmentación completa se debería corresponder con los objetos a interpretar. Sin embargo, los procesos de bajo nivel sólo pueden producir particiones sobre una base no semántica, y aquéllas podrían no corresponder necesariamente con los objetos de interés presentes en la escena. Tales particiones no son siempre únicas, incluso para

el operador humano. La solución a este problema podría estar en el uso de una realimentación guiada por modelos (*top-down*) que permita controlar el estado del proceso. Para ello es preciso establecer medidas que indiquen la distancia a esa referencia conocida u otras medidas de rendimiento, que evaluaremos durante el proceso de análisis. Estas medidas actuarían a modo de sintonizador del modelo, guiando las siguientes acciones del proceso. Para ello, el sistema que proponemos implementa una estrategia de control mixta guiada por datos (bottom-up) y guiada por modelo (top-down), representado en nuestro caso por el contorno extraído en el corte previo. Esta estrategia proporciona interpretaciones correctas de la escena. No obstante, cuando además de reconocer un objeto se pretende definir sus contornos con la mayor precisión posible, este planteamiento puede no ser suficiente, ya que ahora no existe un modelo preciso. Es por ello por lo que para hacer la segmentación más robusta proponemos la integración de la segmentación de regiones proporcionada por este método con la obtenida mediante un proceso de detección de contornos, tal como comentaremos en la parte final del capítulo.

En los apartados siguientes vamos a describir los distintos bloques de procesado que incluye el sistema de análisis de regiones, así como las distintas tareas que desarrollan: segmentación de bajo nivel, extracción de propiedades y estructura de datos, extracción de información de bordes y bloque de alto nivel.

3.4.1 Segmentación de bajo nivel mediante un mapa de propiedades autoorganizado

Un ejemplo del tipo de imágenes con el que vamos a tratar se muestra en la figura 3.12.a. Esta corresponde a un corte transversal de la tibia proximal. En ella se pueden observar las características de baja densidad y estrechez del hueso cortical, así como la existencia de texturas asociadas a la estructura trabecular. Nuestro objetivo es obtener una primera partición de la imagen que sea un compromiso entre un número no excesivamente elevado de regiones y no infrasegmentación. Es decir, se tratará de obtener una primera clasificación de los pixels basada en características de intensidad que proporcione una base adecuada para el procesado de alto nivel. Una forma típica de abordar la clasificación inicial de pixels sería mediante procesos de umbralización, tras el correspondiente análisis del histograma. Sin embargo, a la vista del histograma que presenta esta imagen, y que mostramos en la figura 3.12.b, resulta evidente la dificultad en establecer umbrales

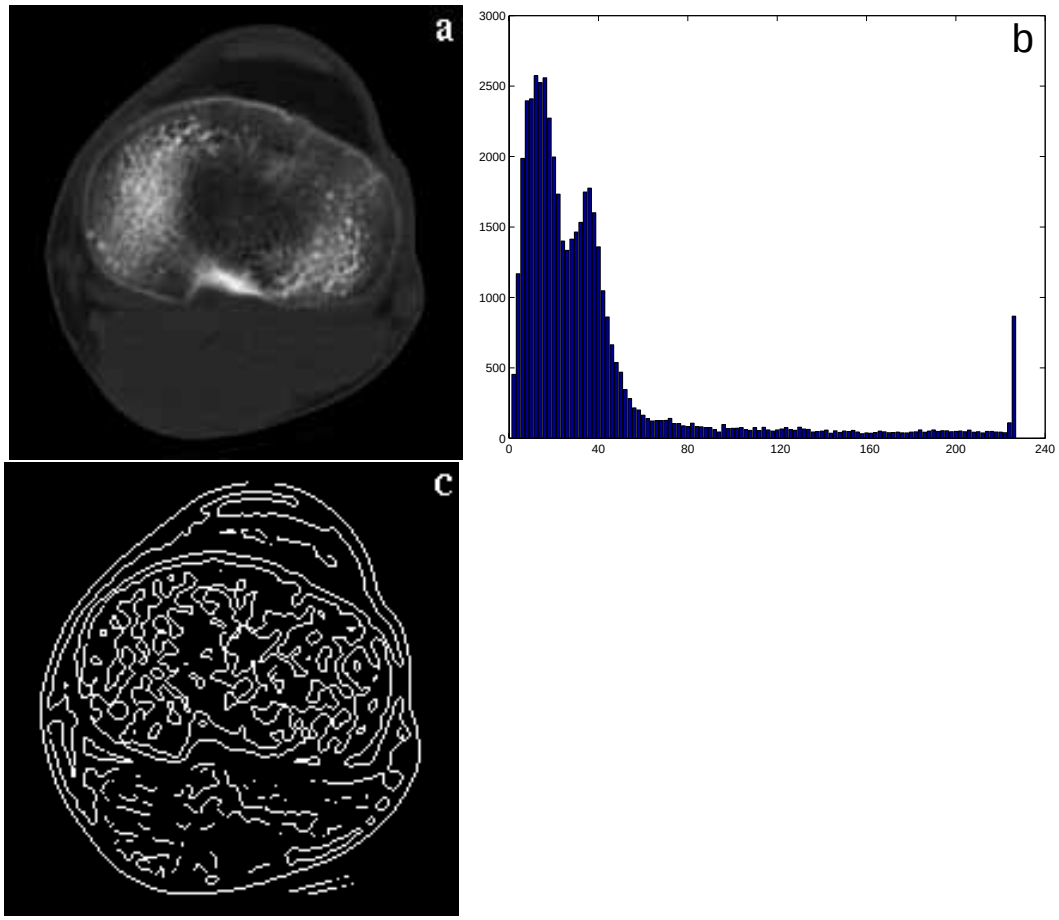


Figura 3.12: (a) Corte CT de tibia proximal; (b) histograma correspondiente; (c) detección de cruces por cero sobre la salida del operador de Marr-Hildredth con σ igual a 2.

de separación entre los diferentes segmentos de la imagen, ya que el hueso está constituido por tejidos cuya intensidad corresponde a la parte media alta de la escala de grises; el hueso no es una estructura homogénea. Por otra parte, los tejidos correspondientes a la trabécula y al músculo comparten un amplio rango de valores de gris. De este modo la umbralización no permite realizar una segmentación adecuada de este tipo de imágenes.

Los inconvenientes producidos por el ruido, junto con las abundantes transiciones de niveles de gris, así como el efecto de interferencia entre éstas, lo que provoca la generación de bordes espúreos y la desaparición de contornos verdaderos, hacen que la presegmentación mediante detectores de bordes no sea adecuada. Estos algoritmos presentan

además, el problema de tener que cerrar contornos, tal como se puede inferir de la imagen de la figura 3.12.c, obtenida mediante un detector de cruces por cero de la salida del operador de Marr-Hildredth al actuar sobre la figura 3.12.a.

Para la segmentación de bajo nivel emplearemos pues un esquema de clasificación de pixels alternativo, basado en un mapa de propiedades auto-organizado [Kohonen, 1982]. Este tipo de red neuronal se entrena mediante un algoritmo de aprendizaje autoorganizado; este tipo de entrenamiento es no supervisado y de naturaleza competitiva. El aprendizaje se lleva a cabo tomando patrones de datos escogidos aleatoriamente entre todas las imágenes, de modo que se realiza un único entrenamiento para toda la secuencia o secuencias a segmentar. En este sistema se introduce como parámetro el tamaño de la ventana que conforma la vecindad de un pixel, lo que determina la dimensionalidad del vector de propiedades que lo caracterizará; sus componentes serán los niveles de gris del propio pixel y los de sus vecinos. Otro parámetro a introducir es el número de clases totales en que se pretende clasificar los pixels de la imagen, lo que establecerá el número de neuronas en la red. Estos dos parámetros se eligen por inspección visual de los resultados, buscando un compromiso entre sobresegmentación de la imagen que evite la unión de estructuras que al final queramos separar, y un número de regiones no excesivamente alto. Tras el análisis de varias secuencias de exploraciones CT, se obtuvieron como valores adecuados el de 3×3 para tamaño de ventana y el de 12 para número de clases. Estos parámetros son los que usa el sistema que proponemos para la segmentación de imágenes de rodilla. Tras el entrenamiento de la red, la posterior proyección en las entradas de los vectores de propiedades que caracterizan a los pixels proporciona una primera discriminación entre las diferentes estructuras existentes en las imágenes, sensible tanto a la distribución global como a los valores locales de los niveles de gris.

Veamos la estructura de la red; ésta se muestra en la figura 3.13. Cada una de sus neuronas N_j, con $j = 1, 2, \ldots, J$, recibe como entrada los niveles de gris de los pixels de una ventana de una imagen CT de un corte de la tibia. Representamos esta ventana mediante un vector de propiedades $P = (p_1, \ldots, p_n)^T$, donde cada elemento p_i representa el nivel de gris del i-ésimo pixel cubierto por la ventana centrada en el pixel bajo análisis. Cada neurona N_j recibe los valores p_i como entradas moduladas mediante un vector de pesos $W_j = (w_{1j}, \ldots, w_{nj})^T$, donde w_{ij} es el peso asociado con la i-ésima entrada a la neurona j. Cuando se aplica una entrada P a la red, se obtiene un vector de salidas

$O = (o_1, \ldots, o_J)^T$, cuyas componentes toman los siguientes valores:

$$o_j = (P - W_j)^T * (P - W_j) \tag{3.15}$$

La neurona con la salida más baja será la que mejor se corresponda con el patrón usado como entrada a la red.

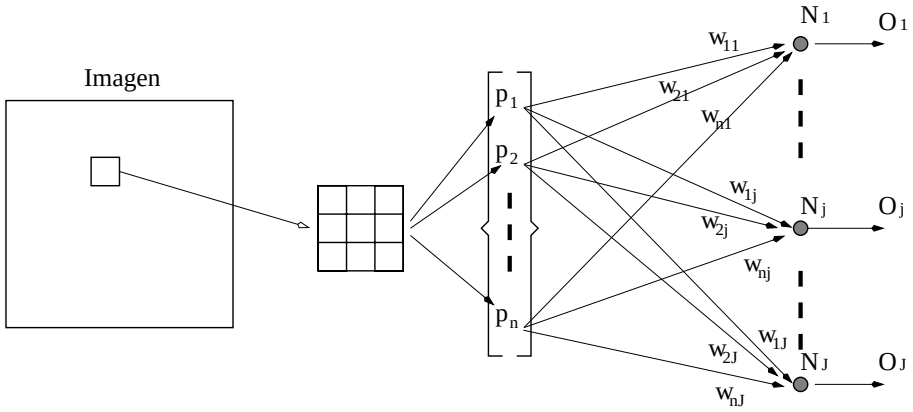


Figura 3.13: Estructura de la red neuronal

El proceso de entrenamiento produce un ajuste de los vectores de pesos asociados con cada neurona de la red. El objetivo de este proceso es hacer las neuronas sensibles a las diferentes estructuras fundamentales existentes entre los patrones elegidos para el entrenamiento. Éste se efectúa del modo que se indica a continuación:

---

1. *Se inicializan todos los pesos a valores aleatorios en el rango de las entradas. Se inicializa un factor de vecindad entre neuronas a $NF = J/2$, donde J es el número de neuronas en la red. Se inicializa $\Delta W_j, 1 \leq j \leq J$, a cero, y el factor de aprendizaje, L, a un valor entre 0 y 1.*

2. *Se considera un conjunto de entrenamiento constituido por los patrones $P_1, \ldots, P_K$. Para cada patrón de entrada a la red P_k se selecciona la neurona N_j con salida mínima. Los pesos W_s de las neuronas N_s, $\forall (j - NF) \leq s \leq (j + NF) | s \in \{1, \ldots, J\}$ se modifican de acuerdo a las siguientes expresiones:*

$$W_s = W_s + L \times \Delta W_s \tag{3.16}$$

$$\Delta W_s = P_k - W_s \tag{3.17}$$

El proceso se repite hasta que para cada P_k, se tenga que $\Delta W_s < \epsilon$.

3. *Si $NF = 0$, entonces se produjo convergencia y termina el entrenamiento. Si no, $NF = NF - 1$ y se vuelve al Paso 2.*

El entrenamiento de la red se hizo sobre un conjunto de patrones escogidos aleatoriamente de varios conjuntos de secuencias de imágenes CT; de esta forma no es preciso entrenar la red para cada secuencia de imágenes, sino que el conjunto de pesos obtenido en este entrenamiento se utilizará en la clasificación de todas las imágenes. Después del proceso de entrenamiento, a cada neurona j de la red se le asocia como etiqueta el nivel de gris $j * 255/12$, dado que como hemos comentado, fijamos $J = 12$.

Una vez entrenada la red, ésta se puede aplicar sobre cualquier imagen para la clasificación de sus pixels. El proceso de clasificación consiste en introducir en la red el vector de propiedades de cada pixel y asignarle a la salida la etiqueta de la neurona ganadora. Las regiones elementales surgen al trasladar esta información de clases al espacio imagen: pixels adyacentes con la misma etiqueta formarán una región elemental.

Un ejemplo de la aplicación de este clasificador se puede ver en la figura 3.14. Este caso corresponde a una imagen de alta calidad, en la que se han eliminado las estructuras no anatómicas y se ha regularizado la imagen mediante un filtro de mediana. Además, esta imagen, procedente de una base de datos del Mallinckrodt Institue of Radiology, corresponde a un corte alejado de la zona proximal en el que tibia y peroné están muy separados y el hueso cortical es ancho y denso. Las regiones en la figura 3.14.b se representan mediante máscaras. Podemos observar en ella una buena delimitación de la zona de hueso cortical.

A medida que nos aproximamos al extremo proximal, la clasificación resultará más compleja, como se puede observar en la figura 3.15. Ésta muestra una nueva imagen correspondiente a la misma base de datos. Ahora la tibia y el peroné están muy próximos, y el ancho y la densidad del hueso cortical disminuyen. Aparece además una lesión que complica aún más el proceso de segmentación. La segmentación actual queda, por tanto, incompleta. Por otra parte, la clasificación que proporciona la red sobre la imagen de la figura 3.12.a es la que se muestra en la figura 3.16. No obstante, como se puede observar, las segmentaciones responden básicamente a las estructuras anatómicas subyacentes y constituyen un buen punto de arranque para el bloque de alto nivel; las distintas etique-

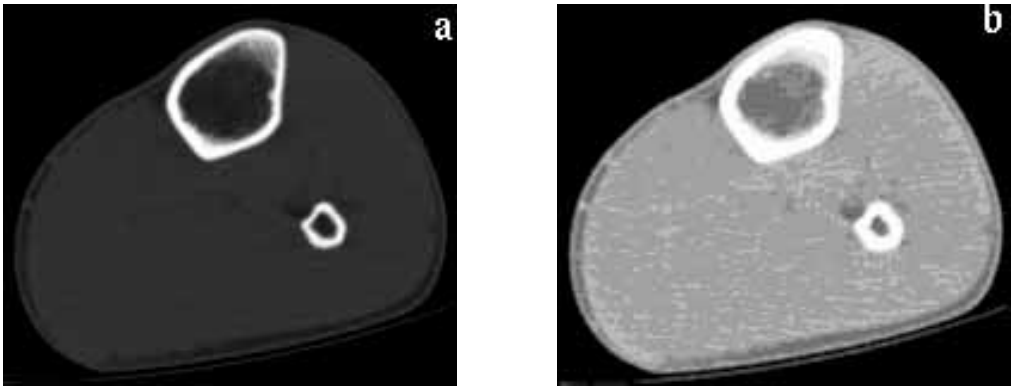


Figura 3.14: Clasificación de pixels mediante un mapa de Kohonen que proporciona una delimitación correcta del hueso: (a) imagen original obtenida a partir de una base de datos del Mallinckrodt Institute of Radiology; (b) imagen etiquetada mediante el clasificador neuronal ($J = 12$ y vector de propiedades de dimensión 3×3 pixels).

tas/densidades existentes dentro del hueso se corresponden con las diferentes estructuras trabeculares. Como el hueso es una estructura heterogénea, en esa nueva etapa será necesario considerar otras propiedades, a parte de la característica de homogeneidad. Así, a partir de esta primera clasificación de los pixels se obtendrá un conjunto de propiedades para cada región conectada, incluyendo entre ellas relaciones de vecindad, y ésta será la información que procesará el bloque de crecimiento de regiones.

En esta primera clasificación se obtiene un número de regiones comprendido entre 3000 y 4000, de las cuales el 75% se corresponde con regiones con un tamaño menor o igual a 25 pixels. En principio se podría pensar en obtener un gran ahorro en memoria y tiempo de procesado eliminando estas regiones; sin embargo las vamos a conservar, ya que en muchos casos van a indicar la presencia de una transición fuerte en los niveles de gris de regiones adyacentes, como, por ejemplo, entre hueso cortical y músculo. Además, nuestro sistema no va a necesitar del procesado de la mayoría de las regiones, como se mostrará posteriormente, gracias a un módulo de foco de atención.

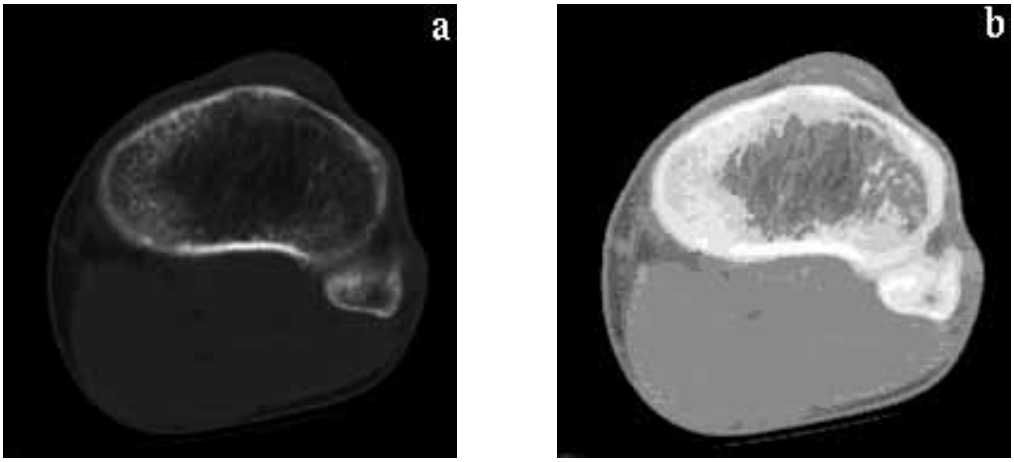


Figura 3.15: Clasificación de pixels mediante un mapa de Kohonen que produce segmentación incompleta de hueso. (a) Imagen original, (b) imagen segmentada.

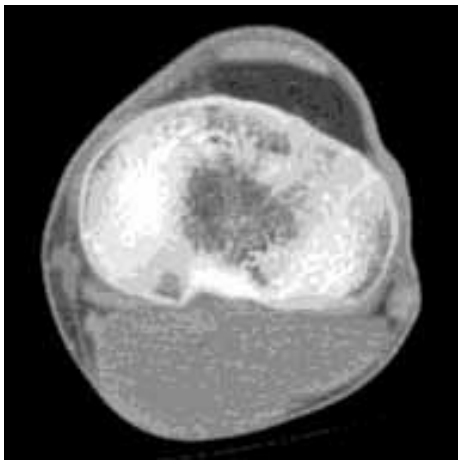


Figura 3.16: Segmentación mediante un mapa de Kohonen de la imagen de la figura 3.12.a.

3.4.2 Representación de características

Una vez realizada una primera reducción de la complejidad de la imagen a través de la clasificación de los pixels antes descrita, se procederá con el análisis de regiones basado en conocimiento. Para ello cada una de las regiones formada por pixels adyacentes con la misma etiqueta quedará descrita mediante un conjunto de propiedades densitométricas,

morfométricas y relacionales. Sin embargo, antes de entrar en este sistema de análisis para el crecimiento de regiones es preciso estructurar la información de una forma adecuada.

Dado que los algoritmos de división y unión usarán información a cerca de las adyacencias entre regiones, así como sobre sus características, una estructura adecuada para representar esa información es la conocida como *grafo de adyacencias de regiones* (RAG) [Jain y col., 1995]. En este grafo los nodos representan regiones, mientras que los arcos entre nodos representan las longitudes de los contornos comunes entre dichas regiones. En las estructuras de datos de los nodos se representan las diferentes propiedades de las regiones que sean útiles al proceso de segmentación. El grafo se construye habitualmente tras la segmentación inicial, y los datos en él almacenados se podrán combinar después para obtener una mejor segmentación. Así, los algoritmos de división y unión manipularán esta estructura de datos de forma que represente en cada momento la nueva estructura de regiones creada. La figura 3.17 ilustra la idea del RAG.

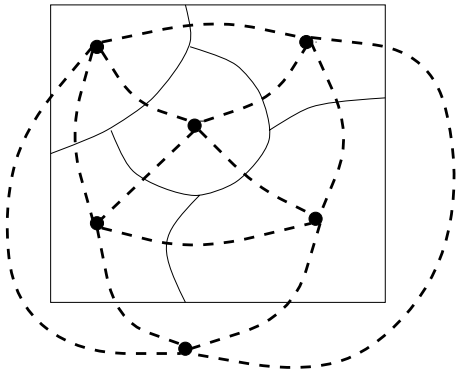


Figura 3.17: Partición de una imagen y RAG asociado. Las curvas a trazos corresponden al grafo de adyacencias, y las demás corresponden a la partición de la imagen. El nodo situado fuera del rectángulo que delimita la imagen representa el fondo (*background*) de la imagen. Cuando, como en este caso, no existe una verdadera región de fondo, el nodo correspondiente se sitúa fuera de la imagen y representa el mundo exterior.

Tal como hemos comentado, a la hora de etiquetar los pixels de una imagen para obtener una segmentación de la misma que nos permita identificar los objetos de interés sería conveniente disponer no sólo de una descripción de la imagen de entrada, sino de un **modelo** para cada una de los objetos a identificar. Ballard y Brown [1982] introducen dos tipos de modelos: *declarativos* y *procedurales*. Los modelos declarativos consisten en un

conjunto de ligaduras sobre las propiedades de, o las relaciones entre, las partes en el nivel de descripción dado. Las partes cuyas descripciones satisfacen estas ligaduras constituyen la clase definida por el modelo. Un modelo procedural, en general, es cualquier proceso que genera o reconoce objetos en las imágenes; las clases que él define están formadas por los objetos que genera o acepta.

Para el proceso de segmentación basado en reglas adoptaremos una representación de las regiones mediante un grafo de adyacencias. Ahora la entidad a procesar ya no es el pixel, como en los procesos de clasificación de bajo nivel, sino las regiones, que se verán sometidas a procesos de unión o división. Las reglas evaluarán las propiedades de las regiones, que constituirán la información almacenada en el nodo correspondiente del RAG, y los procesos de unión se ejecutarán sobre algunos de los vecinos con los que comparte contorno y a los que está unido en el RAG mediante arcos. El modelo que representa la estructura a identificar se almacenará considerando el mismo esquema de representación declarativo usado para las propiedades de las regiones; en este caso, en los nodos se incluirán las propiedades del modelo.

Una vez realizada la segmentación de bajo nivel y adoptado el formalismo para representación de los datos, la siguiente tarea a realizar es el agrupamiento en regiones de las componentes conectadas resultantes de la presegmentación. De esta forma se pasa de trabajar a nivel de pixels a hacerlo a nivel de regiones. El algoritmo para la identificación de regiones, usando conectividad 4, es el siguiente:

---

1. *Explorar la imagen I de izquierda a derecha y de arriba abajo.*
2. *Si $I(x,y) \neq 0$ entonces:*
   (a) *Si $ETIQ(I(x-1,y)) \otimes ETIQ(I(x,y-1))$, entonces copiar etiqueta.*
   (b) *Si $ETIQ(I(x-1,y)) = ETIQ(I(x,y-1))$, entonces copiar etiqueta.*
   (c) *Si $ETIQ(I(x-1,y)) \neq ETIQ(I(x,y-1))$, entonces copiar $ETIQ(I(x,y-1))$ y hacer $ETIQ(I(x-1,y)) \equiv ETIQ(I(x,y-1))$.*
   (d) *En otro caso, asignar una nueva etiqueta.*
3. *Si hay pixels sin explorar, ir a 2.*
4. *Propagar las equivalencias de las etiquetas.*

Una vez identificadas todas las regiones, el paso siguiente consiste en el cálculo de propiedades y la creación del grafo de adyacencias. Dentro de todas las posibles propiedades que se pueden extraer de una región, nosotros hemos escogido las siguientes: área (A), nivel medio de intensidad (ρ), varianza (σ^2), centro de masas ($\bar{X}, \bar{Y}$), momentos principales de inercia ($M^{r,s}$), perímetro (Per) y matriz de dispersión (C). Las expresiones para su cálculo se indican a continuación:

$$A_i = \sum_{j \in R_i} 1 \tag{3.18}$$

$$\rho_i = \frac{1}{A_i} \sum_{j \in R_i} I_j \tag{3.19}$$

$$\sigma_i^2 = \frac{1}{A_i} \sum_{j \in R_i} (I_j - \rho_i)^2 \tag{3.20}$$

$$\bar{X}_i = \frac{1}{A_i} \sum_{j \in R_i} X_j \tag{3.21}$$

$$\bar{Y}_i = \frac{1}{A_i} \sum_{j \in R_i} Y_j \tag{3.22}$$

$$C_i = \frac{1}{A_i} \sum_{j \in R_i} ((X_j, Y_j) - (\bar{X}_i, \bar{Y}_i))((X_j, Y_j) - (\bar{X}_i, \bar{Y}_i))^t \tag{3.23}$$

donde el índice i referencia la región e I_j representa el nivel de gris del pixel con coordenadas (X_j, Y_j). Los sumatorios se extienden a todos los pixels de la región R_i. El perímetro de la i-ésima región se obtiene mediante la ecuación:

$$Per_i = \frac{1}{2} \sum_{P \in R_i} [b_1(P) + b_2(P)\sqrt{2}] \tag{3.24}$$

siendo

$$b_1(P) = \sum_{k=0}^{3} [(P \wedge P_{2k}) - (P_{2k} \wedge (P_{2k-1} \vee P_{2k-2}) \wedge (P_{2k+1} \vee P_{2k+2}))] \tag{3.25}$$

$$b_2(P) = \sum_{k=0}^{3} [(P \wedge P_{2k+1}) - (P_{2k} \wedge P_{2k+2})] \tag{3.26}$$

donde P hace referencia a un pixel de la región, y $P_0, \ldots, P_7$ hacen referencia a sus 8 vecinos, numerados en sentido antihorario comenzando por el situado a mano derecha. P_j

es 1 si pertenece a la región R_i y 0 en otro caso. $\wedge$ y $\vee$ son los operadores lógicos AND y OR, respectivamente. Los subíndices de las dos últimas ecuaciones son todos módulo 8. El sistema hará uso no sólo de este conjunto de propiedades, sino también de otras que, aunque no estén explícitamente representadas, si lo están de forma implícita, ya que pueden calcularse a partir de las representadas en el marco. Entre estas propiedades están las relaciones de **es-un-hueco-de**, **tiene-huecos**, **es-adyacente-a**, **posición-relativa**, y otras propiedades numéricas tales como el índice de compactación y la elongación, que se pueden calcular según indican la siguientes ecuaciones:

$$\text{COMPACTACIÓN}_i = \frac{4 \times \Pi \times A_i}{(Per_i)^2} \tag{3.27}$$

$$\text{ELONGACIÓN}_i = \log \frac{M_i^{20}}{M_i^{02}} \tag{3.28}$$

donde M^{20} y M^{02} son los segundos momentos centrales con respecto al eje X y al eje Y, respectivamente. Para que la medida de elongación sea invariante es preciso calcular los momentos centrales con respecto a los ejes principales de inercia. Para ello se calculan los autovalores de la matriz de dispersión; en otro caso, la medida dependería de la orientación de la región dentro de la escena. La matriz de dispersión representa el área elíptica que aproxima la forma de la región, como se ilustra en la figura 3.18. Los autovectores de esta matriz representan los valores invariantes de los segundos momentos centrales, que serán los que realmente almacenaremos.

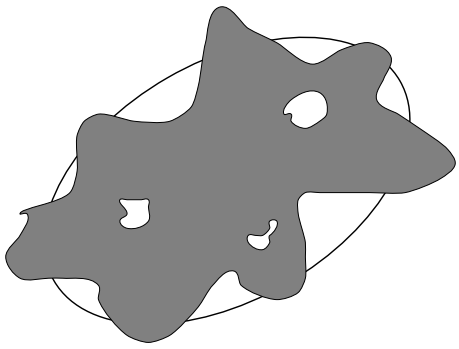


Figura 3.18: Elipse correspondiente a la matriz de dispersión.

Los autovectores de la matriz de dispersión dan una medida de orientación de la región. Un autovector v de la matriz C será aquel que para algún escalar λ, llamado autovalor,

cumpla,

$$vC = \lambda v \tag{3.29}$$

Los autovectores se calculan a partir del polinomio característico:

$$\det(C - \lambda I) = 0 \tag{3.30}$$

Los autovalores, que serán independientes de la orientación de la región, se corresponden con los segundos momentos centrales definidos de la siguiente forma:

$$M_i^{r,s} = \sum_{j \in R_i} (X_j^p - \bar{X}_i^p)^r (Y_j^p - \bar{Y}_i^p)^s \tag{3.31}$$

$$\tag{3.32}$$

donde (X_i^p, Y_i^p) representan las coordenadas respecto de los ejes principales de inercia (autovectores), con $r, s \in \{0, 2\}$. La razón de considerar simultáneamente la matriz de dispersión y los segundos momentos centrales es que la primera es fácil de actualizar tras los procesos de unión, y los segundos contendrán en cada momento los valores con respecto a los ejes principales de inercia.

Una vez que se realiza la segmentación de bajo nivel y la extracción de propiedades, las características de cada región se organizan como un marco (*frame*):

(REGIÓN
(IDENTIFICADOR (VALOR 1234))
(CLASE_REGIÓN (VALOR 6))
(ÁREA_DE_SOPORTE (VALOR 10000001))
(ETIQUETA(TIBIA(VALOR(0.9)) PERONÉ(VALOR(0)) FONDO(
VALOR(0))))
(ÁREA (VALOR 123))
(PERÍMETRO (VALOR 245))
(DENSIDAD_MEDIA (VALOR 84.32))
(VARIANZA (VALOR 10.12))
(CM_X (VALOR 112.25))
(CM_Y (VALOR 165.15))

(M_DISPERSIÓN (VALOR 106.01 1800.40 1800.40 80.50))
(MOMENTO_02 (VALOR 102.31))
(MOMENTO_20 (VALOR 99.37))
(ADYACENCIA(VALOR(1902 (GRADO(VALOR(50))) 1822 (GRADO(VALOR(10))) 1002 (GRADO(VALOR(200))))))

En el marco se asigna a cada región un número identificador, la clase a la que pertenece cada región inicialmente tras la presegmentación, el área de soporte a la que pertenece (información necesaria para el proceso de crecimiento; en su momento la definiremos), la etiqueta que se le asigna tras proceso de crecimiento (tibia, peroné o fondo), las regiones adyacentes (identificador y perímetro compartido) y el resto de propiedades que antes hemos comentado. Estas características se escogen no sólo por la información que encierran, sino también por la facilidad de recálculo cada vez que se hace una unión. De esta forma, el proceso de unión/división de regiones se realiza sobre el RAG, sin necesidad de volver sobre la imagen etiquetada. Así, por ejemplo, el área resultante de la unión de dos regiones es simplemente la suma de las áreas individuales. El perímetro viene dado por la siguiente expresión:

$$P_U = P_{R1} + P_{R2} - 2 \times \mathrm{PC}_{R1,R2} \tag{3.33}$$

donde P es el perímetro, PC el perímetro compartido, el subíndice U indica la región resultante de la unión y los subíndices $R1$ y $R2$ denotan las regiones individuales que se están uniendo. El centroide de la región resultante de la unión de otras dos se puede calcular a partir de los centroides de las regiones individuales, de acuerdo con las siguientes ecuaciones:

$$\begin{aligned} \overline{X}_U &= \frac{(\overline{X}_{R1}A_{R1} + \overline{X}_{R2}A_{R2})}{(A_{R1} + A_{R2})} \\ \overline{Y}_U &= \frac{(\overline{Y}_{R1}A_{R1} + \overline{Y}_{R2}A_{R2})}{(A_{R1} + A_{R2})} \end{aligned} \tag{3.34}$$

donde $(\overline{X}, \overline{Y})$ es el centroide y A es el área . Para la densidad media se tiene una expresión idéntica, sin más que intercambiar $\overline{X}$ por ρ. La varianza se recalcula de la siguiente forma:

$$\sigma_U = \frac{(\sigma_{R1}A_{R1} + \sigma_{R2}A_{R2} + A_{R1}(\rho_U - \rho_{R1})^2 + A_{R2}(\rho_U - \rho_{R2})^2)}{(A_{R1} + A_{R2})} \tag{3.35}$$

Y la matriz de dispersión se recalcula como:

$$C_U = \frac{A_{R1}(C_{R1} + (\overline{X}_{R1} - \overline{X}_U)^2) + A_{R2}(C_{R2} + (\overline{X}_{R2} - \overline{X}_U)^2)}{(A_{R1} + A_{R2})} \tag{3.36}$$

Tras las operaciones de unión, las regiones finales obtenidas se comparan con el modelo correspondiente (tibia o peroné), el cual se representa mediante un marco del tipo que se indica a continuación:

(MODELO
(ETIQUETA (TIBIA—PERONÉ))
(NUM_MÁSCARA (VALOR 1))
(ÁREA (VALOR 1003))
(CM_X (VALOR 140.25))
(CM_Y (VALOR 185.15))
(MOMENTO_XX (VALOR 1005.31))
(MOMENTO_YY (VALOR 10.37)))

La característica de número de máscara (**NUM_MÁSCARA**) sirve para asignar un número al modelo; su función está relacionada con la característica de área de soporte que aparece en el marco de las regiones. El resto de propiedades hacen referencia a la forma de la máscara de hueso del corte previo, que como hemos comentado, es el modelo que guía el análisis en el corte actual. En la figura 3.19 se muestra un ejemplo de un grafo de adyacencias de regiones (RAG). Una vez creado el RAG, ya se está en condiciones de abordar los procesos de unión de regiones del bloque de alto nivel. Las operaciones afectarán al RAG, que habrá de ser actualizado tras cada suceso. Al final, cada región tendrá asignada una etiqueta del conjunto {tibia, peroné, fondo}.

La existencia de un modelo aproximado de las formas de las estructuras de interés permite introducir medidas dentro del proceso de crecimiento que indiquen el grado de aproximación entre el estado actual de la segmentación y la estructura buscada, al tiempo que permite identificar las estructuras mismas cuando aparezcan otras del mismo tipo en la segmentación de la imagen. Describiremos más adelante las medidas de distancia de forma empleadas por nuestro sistema.

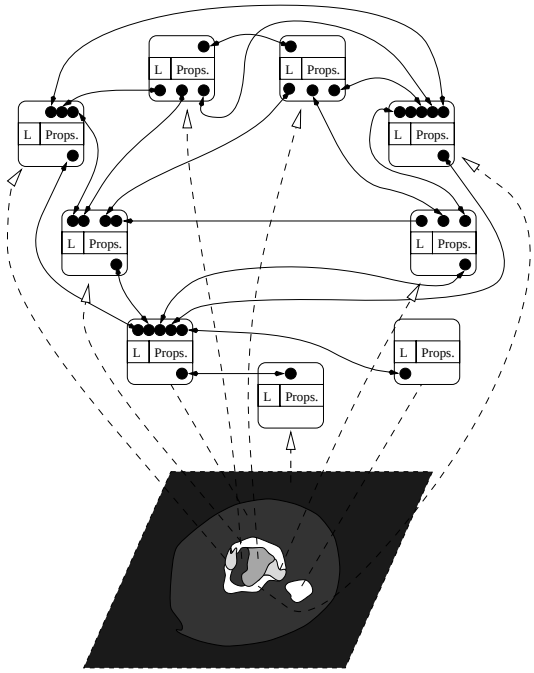


Figura 3.19: Grafo de adyacencias de regiones. Cada nodo guarda las propiedades individuales de cada región (**Props.**) y su etiqueta (**L**), mientras que los arcos determinan las relaciones de vecindad. Propiedades tales como **es un hueco de**, **posición relativa**, etc., se obtienen a partir de las propiedades de cada uno y de las relaciones de adyacencia.

3.4.3 Obtención de información de bordes

Ya hemos comentado anteriormente que la presegmentación mediante detección de bordes no era el método más adecuado para este tipo de imágenes, pues planteaba una serie de problemas difíciles de resolver, como eran la recuperación de bordes perdidos o la eliminación de bordes falsos. Sin embargo, si puede resultar útil la inclusión de información de bordes en el proceso de segmentación para resolver problemas de infrasegmentación en alguna parte de la escena. Hay que decir en este punto que este módulo de procesado no se aplica de forma sistemática a todas las imágenes, sino que es un proceso lanzado por el sistema de control en el bloque de alto nivel sobre una determinada zona de la imagen cuando reconoce la necesidad de realizar operaciones de división de regiones.

Como los procesos de unión son más simples que los de división, nosotros partiremos

siempre de una sobresegmentación de la imagen, que simplificaremos mediante procesos de agrupación. Cuando el algoritmo de preprocesado no produzca en algunas zonas suficiente segmentación o cuando la unión (*merging*) de resultados erróneos, recurriremos a la información de bordes para dividir regiones y/o corregir la posición de los contornos.

La aplicación de los procesos de crecimiento de regiones puede conducir a tres tipos de errores [Pavlidis y Liow, 1990]:

1. Un contorno no es un borde y no hay bordes próximos.
2. Un contorno se corresponde con un borde, pero no coincide con él.
3. Existen bordes para los que no se detectan contornos próximos.

Con el fin de corregir errores del tipo (1) hemos implementado un sistema de unión (*merging*) de regiones basado en conocimiento. Con el objeto de corregir errores de tipo (2) se someterá a las imágenes de máscaras a operaciones morfológicas y a suavizado de contornos, con lo que se pretende obtener resultados más realistas. La probabilidad de errores del tipo (3) se puede reducir significativamente, si no eliminar, mediante una relajación en los parámetros del proceso de segmentación. Esto, sin embargo, tiene como resultado una sobresegmentación de la imagen que hace crecer el número de errores del primer tipo.

La principal razón por la que los procesos de crecimiento producen falsos contornos u olvidan otros es que las definiciones de uniformidad son demasiado estrictas para poder verificar todos los requerimientos que se puedan imponer al sistema. El principal objetivo de la integración de información de bordes con la del crecimiento de regiones es añadir un criterio de heterogeneidad adicional en aquellas partes donde el crecimiento de regiones no encuentra suficiente evidencia de ello. La presencia de bordes dentro de una región, junto con el conocimiento de la forma aproximada de éstos, son los elementos básicos que conforman los criterios para la toma de decisiones en el proceso de división (*splitting*).

Los principales problemas asociados con la detección de contornos en imágenes médicas radican en el bajo contraste y la pobre relación señal ruido, que limitan la correcta estimación de los cambios de intensidad y crean numerosos artefactos. En las imágenes CT de rodilla hemos de considerar además la fuerte textura introducida por la estructura trabecular, así como la presencia de patologías (osteoporosis, osteofitos,...) o lesiones

que modifiquen la forma normal de estas estructuras o no dejen suficiente evidencia de la separación entre ellas. Por otra parte, la proximidad entre tibia y peroné y entre tibia y fémur puede impedir la detección de la transición entre esas estructuras.

El primer paso que hay que dar para la obtención de los contornos que delimitan las estructuras de interés es la detección de puntos de borde. De entre los detectores de bordes clásicos, aquellos que gozan de mayor popularidad son el operador de Marr-Hildredth [Marr y Hildredth,1980] y el detector de bordes de Canny [Canny, 1986]. Sin embargo, el problema es demasiado complicado para ser resuelto por simples algoritmos [Taratorin y Sideman, 1993]. La tendencia actual es la incorporación de información *a priori* en los algoritmos de detección. La detección de bordes no puede resolverse contando únicamente con información de intensidades; es preciso incorporar información *a priori* sobre propiedades geométricas y, en nuestro caso, anatómicas.

La reconstrucción consistente de un contorno basándose en los puntos candidatos obtenidos a partir de un operador de bordes sólo es posible si se pueden resolver los dos problemas siguientes:

- Extrapolación de puntos de contorno en regiones donde no se detecta su presencia.
- Eliminación de falsos contornos del conjunto de puntos candidatos.

Daremos solución a estos problemas mediante el análisis del comportamiento de los puntos de borde en el espacio de escalas y usando información *a priori* obtenida del procesado del corte precedente en la secuencia. Esta información *a priori* la incorporaremos como un conjunto de ligaduras. Comenzaremos la descripción del sistema para la obtención de información de bordes considerando el operador de bordes y el análisis en el espacio de escalas; finalmente describiremos la forma de incorporar la información *a priori*.

Detección de bordes: operador de Marr

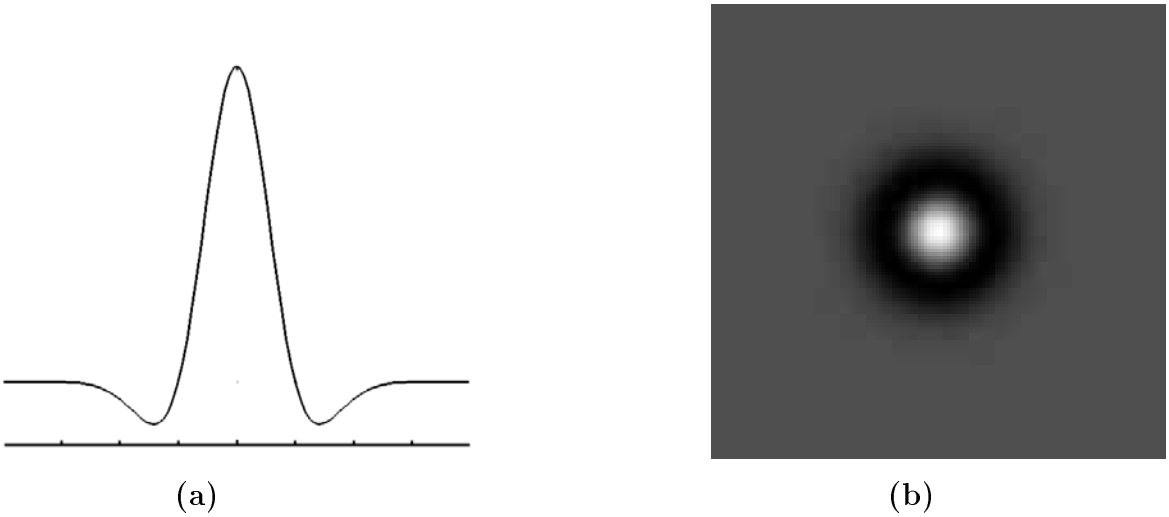
La operación del detector de bordes la realizaremos en un espacio 2D, y no en 3D como en ocasiones se ha efectuado [Bomans y col., 1990]. Aunque el conjunto de datos pertenece a un volumen, la densidad de cortes en la exploración CT pudo no ser lo suficiente densa como para obtener una continuidad adecuada para la aplicación de operadores 3D, aunque las formas de los contornos se asemejen. La desigual resolución de la exploración

CT provoca discontinuidades en la escena de entrada que restringe el uso de verdaderas operaciones 3D [Raya, 1990]. Aunque, como ya hemos indicado, se obtendrá un volumen de datos isótropo mediante la generación de cortes intermedios, los efectos de interacción entre diferentes superficies y entre diferentes partes de la misma superficie se agravan en relación con lo que ocurre en la interferencia entre curvas del mismo plano. Por otra parte, cuando se procesa corte a corte es posible disponer de modelos aproximados a partir del corte adyacente, cosa que no ocurriría en el procesado 3D puro.

Para la detección de las transiciones dentro de cada corte emplearemos el operador de Marr-Hildredth, definido por:

$$C(x,y) = \nabla^2(I(x,y) * G(x,y,\sigma)) \tag{3.37}$$

donde $I(x,y)$ es la imagen a procesar, $G(x,y,\sigma)$ la función Gaussiana con desviación típica σ, ∇^2 el operador Laplaciano y $C(x,y)$ es la imagen de contornos. La figura 3.20 muestra el operador Laplaciana de la Gaussiana (LoG) en 1D y 2D.



(a) (b)

Figura 3.20: Operadores de Marr-Hildredth: (a) 1D, (b) 2D.

El operador de Marr-Hildredth suaviza primero los datos para eliminar estructuras con componentes de alta frecuencia. De esta forma se hace una limitación en el rango de frecuencias sobre el que puede ocurrir un cambio de intensidad, considerándose entonces

sólo cambios de intensidad a resoluciones específicas. Marr y Hildredth mostraron que un operador Gaussiano constituye un compromiso óptimo entre suavidad y precisión en la localización de contornos, puesto que la distribución Gaussiana es suave y localizada en los dominios espacial y frecuencial. Otra ventaja de la función Gaussiana es que la cantidad de suavización, y por tanto el rango de frecuencias considerado, se puede ajustar fácilmente escogiendo un valor apropiado para la desviación estándar. Con un valor de σ bajo la operación de suavizado involucra valores situados en un entorno reducido, mientras que con una desviación mayor la suavización involucra a una vecindad de mayor tamaño, con lo que se pierde precisión en la localización de los puntos de borde. Otra ventaja del operador es que es rotacionalmente invariante. Si hay un cambio de intensidad en una orientación particular dentro de una imagen, habrá un pico en la primera derivada direccional y un cero en la segunda derivada direccional tomada perpendicular a la orientación del cambio de intensidad. Sin embargo, la detección con un operador direccional, como es el de Canny, tiene la desventaja de que debe aplicarse el operador con diferentes direcciones con el fin de seleccionar la dirección de mayor gradiente. Así pues, hemos elegido un operador invariante rotacionalmente para reducir a uno el número de convoluciones.

La detección de los cruces por cero en la respuesta del operador de Marr-Hildredth nos indicará la localización del contorno; mediante el estudio de las posiciones de las transiciones determinaremos la orientación y magnitud de éstas. Para ello, con el fin de preservar el tamaño y la forma de las regiones a ambos lados de las transiciones y dar una localización más precisa a los puntos de cambio de intensidad aplicamos el algoritmo basado en predicados de Huertas y Medioni [1986]. En este algoritmo se inspecciona cada vecindad 3×3 en la imagen convolucionada para ver si cuadra con alguno de los once predicados permitidos para cruce por cero, los cuales definen 24 posiciones del borde.

Análisis en el espacio de escalas

El grado de suavización que introduce este operador de bordes antes de la extracción de los cambios de intensidad se ajusta escogiendo el tamaño adecuado para el operador de suavización. En la mayoría de las aplicaciones, las médicas particularmente, no es suficiente con una única escala de suavización. Para detectar transiciones que ocurren a diferentes escalas es preciso aplicar operadores de diferentes tamaños y posteriormente integrar la información mediante un análisis multiescala. En general, los detalles más

finos en cambios de intensidad se obtienen con parámetros de escala pequeños, mientras que los operadores con mayores parámetros de escala proporcionan información menos fina, tal como ilustra la figura 3.21. De esta forma, el análisis multiescala se emplea para eliminar ruido de escala fina.

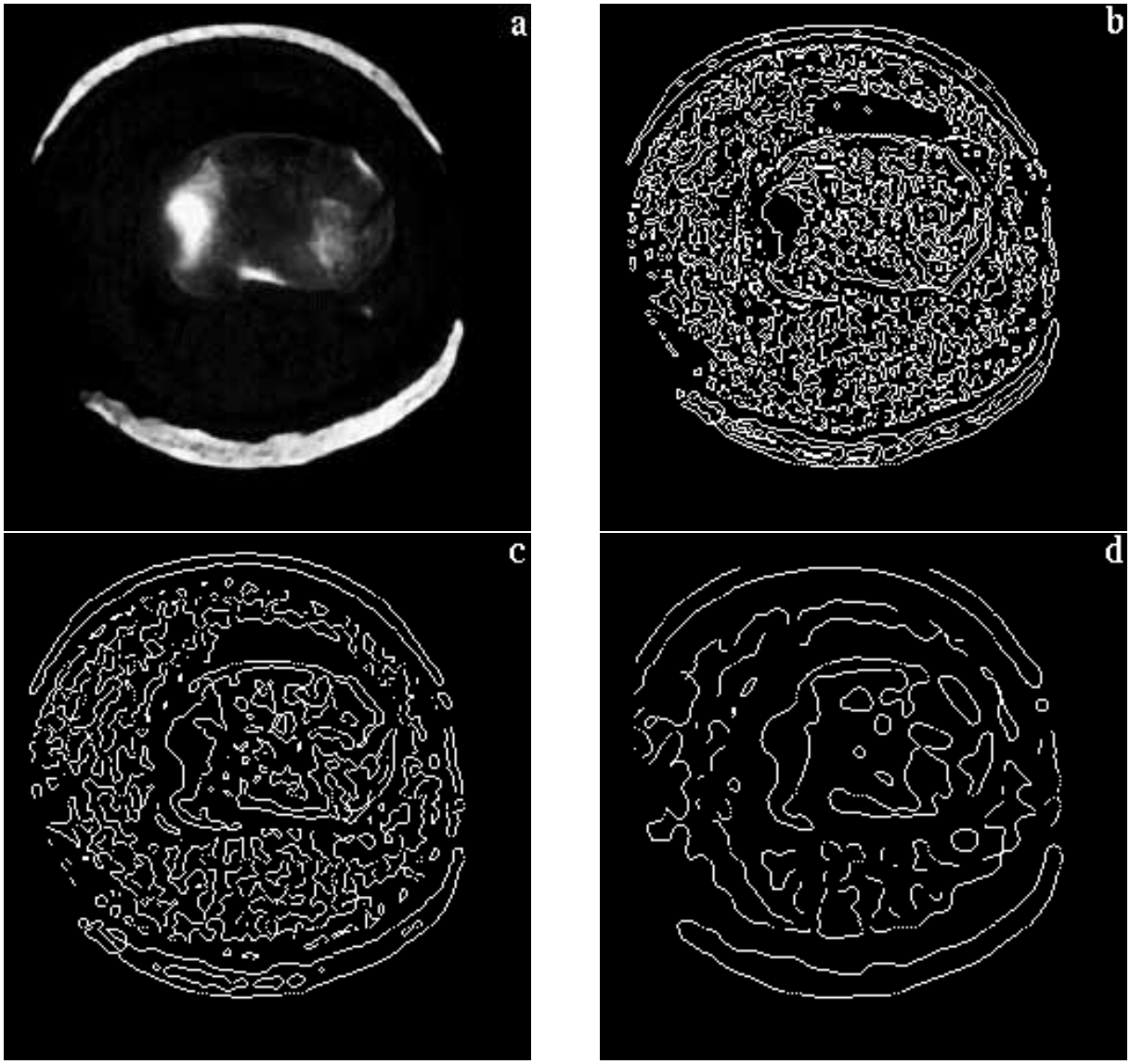


Figura 3.21: (a) Imagen original, (b), (c) y (d): imágenes de contornos con $\sigma = 1.0$, 2.0 y 4.0 respectivamente.

Los dos problemas iniciales con los que nos enfrentamos son la selección de las escalas de los operadores y la combinación de información de bordes obtenida a diferentes escalas. La selección de los parámetros de escala es absolutamente dependiente del tipo de imágenes concretas con las que se está trabajando. Sin el conocimiento a cerca de las características deseadas es difícil seleccionar los parámetros de escala e integrar la información obtenida con las diferentes escalas. Nosotros hemos seleccionado las escalas mediante inspección visual, evitando aquellas que introducen demasiados detalles y contornos ruidosos, y aquellas otras que introducen una suavización excesiva que impide identificar transiciones significativas.

Por otra parte, la combinación de la información ha de basarse en el conocimiento del comportamiento de los bordes a diferentes escalas para el operador concreto. Los tres comportamientos típicos en el espacio de escalas Gaussiano son [Lu y col., 1992]:

- *Dislocación de bordes.* La posición obtenida para un borde en la imagen filtrada puede ser diferente de su posición en la imagen real.
- *Pérdida de bordes.* Existen bordes de la imagen real que no tienen correspondencia en la imagen filtrada.
- *Bordes espúreos.* Existen bordes que aparecen en la imagen filtrada y que no existen en la imagen real.

Los problemas relativos a dislocación y a falsedad de bordes los intentaremos resolver mediante la aplicación de algoritmos basados en el comportamiento de bordes no lineales en el espacio de escalas descrito por Lu y col. [1989, 1992]. En el análisis en el espacio de escalas se procede desde los tamaños de escala mayores hacia los más finos; ello se justifica en base a los siguientes comportamientos:

- Es fácil identificar diferentes curvas en una imagen de cruces por cero con el operador de mayor escala.
- Los bordes falsos desaparecen cuando el parámetro de escala cambia de mayor a más fino.
- Sólo pueden desaparecer bordes falsos cuando el parámetro de escala cambia en este sentido.

- Los bordes perdidos en las escalas mayores siempre pueden recuperarse en las escalas más pequeñas.

El problema de dislocación de bordes en el análisis multiescala es inevitable. Los cruces por cero correspondientes a las mismas curvas de borde tienen diferente localización dependiendo de la escala. El conocimiento usado para corregir la posición de los bordes se representa con las reglas de la tabla 3.1. Del conjunto de reglas indicadas en ella se deduce que las posiciones más correctas de los bordes se dan en sigmas pequeños, y que la determinación de los desplazamientos no puede hacerse para la curva completa, sino que hay que hacerlo punto a punto. Si I_{σ_j} representa el mapa de contornos para el parámetro de escala σ_j, la posición más precisa de estos contornos se obtendrá en la imagen $I_{\sigma_{j-1}}$. Centrándose en cada posición de punto de borde en I_{σ_j}, se define un área de búsqueda de radio $d = \|\sigma_j - \sigma_{j-1}\|$ en $I_{\sigma_{j-1}}$ y se busca la distancia más próxima al centro de los cruces por cero dentro de ese área. La posición encontrada definirá la nueva localización del punto de borde.

Tabla 3.1: Reglas para corrección de localización [Lu y col. 1992].

Regla 1:
SI una curva borde es no lineal
ENTONCES los puntos del borde sobre esta curva tienen diferentes valores de dislocación y de dirección de desplazamiento.

Regla 2:
SI un punto de borde se localiza en σ_1
ENTONCES su dislocación es menor o igual que σ_1.

Regla 3:
SI la dislocación de un punto de borde detectado en σ_1 es x_1 y en σ_2 es x_2
ENTONCES en la mayoría de los casos tendremos que $\|x_2 - x_1\| \leq \|\sigma_2 - \sigma_1\|$.

El comportamiento de los bordes en el espacio de escalas muestra que los operadores LoG con escalas mayores eliminan tanto ruido como algunos bordes deseados. Hay dos objetivos dentro de la recuperación de bordes perdidos. El primero de ellos es recuperar los cruces por cero que conecten los segmentos de curva que representan una curva borde. El segundo es recuperar los bordes perdidos en los parámetros de escala mayores. El primero

de los objetivos se puede alcanzar recuperando todas las curvas de borde conectadas a los bordes proyectados desde la escala mayor. Para recuperar en lo posible estos bordes perdidos se implementan procedimientos algorítmicos basados en las reglas de la tabla 3.2. Sin embargo, el segundo de los objetivos es más complicado de alcanzar; para ello se integrará conocimiento proveniente de técnicas de crecimiento de regiones y de formas de contornos aproximadas para la situación concreta.

Dos son los tipos de cruces por cero no deseados que hemos de eliminar: los cruces por cero falsos generados por modelos de borde tipo escalera y los cruces por cero debidos a ruido. La naturaleza de los procesos grosero-a-fino indican que todas las curvas borde falsas desaparecen en parámetros de escala bajos. Esto implica que si una curva de cruces por cero falsa desaparece en la escala σ^0, en el proceso de corrección de posición de bordes no se encontrarán las posiciones correspondientes a curvas falsas en escalas $\sigma \leq \sigma^0$. Sin embargo, debido a la presencia de ruido, aún pueden quedar cruces por cero correspondientes a curvas falsas. Este conocimiento se representa por las reglas de la tabla 3.3.

Dentro de la tarea de eliminación de bordes falsos y ruidosos se pueden marcar tres objetivos. El primero es comparar la longitud de las curvas de cruces por cero en la escala actual con las obtenidas para escalas mayores. Si la longitud de la curva en la escala actual es mucho menor que la obtenida para escalas mayores, es posible que represente una curva falsa y se ha de eliminar. El segundo objetivo es eliminar curvas abiertas. El tercero es eliminar curvas de cruces por cero que se puedan haber obtenido por fusión de curvas correspondientes a un modelo de intensidad pulso; para este último objetivo necesitaremos aplicar conocimiento concreto a cerca del problema actual.

Sistema para la obtención de contornos externos de tibia y peroné

Como primer paso para la obtención de los contornos externos de tibia y peroné en zonas de interés hemos de establecer el número y magnitud de los parámetros de escala con los que va a trabajar el sistema. La selección de sus magnitudes se efectuó en base a las siguientes consideraciones: sea $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_k\}$ la secuencia de parámetros de escala, donde $\sigma_i < \sigma_{i+1}$, I_σ la imagen de cruces por cero obtenida con σ y $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ la secuencia de contornos a detectar. Entonces se deben cumplir las siguientes condiciones para cada curva borde Γ_j:

Tabla 3.2: Reglas para recuperación de bordes perdidos [Lu y col., 1992].

Regla 4:
SI las curvas de borde deseadas forman un modelo de intensidad de pulso múltiple o de escalera
ENTONCES estas pueden desaparecer a escalas grandes.

Regla 5:
SI las curvas de cruces por cero desaparecen
ENTONCES lo harán a pares.

Regla 6:
SI una curva de cruces por cero desaparece en σ_1
ENTONCES siempre se podrá recuperar en una escala más pequeña $\sigma < \sigma_1$.

Regla 7:
SI una curva borde se detecta en σ
ENTONCES los posibles bordes perdidos sólo pueden estar en su vecindad R_σ

Regla 8:
SI una curva deseada es no lineal
ENTONCES esta se puede unir con sus curvas vecinas para dar una curva mayor.

Regla 9:
SI una curva cerrada forma modelos escalera con sus curvas vecinas
ENTONCES sus segmentos de cruces por cero pueden desaparecer con curvas de cruces por cero vecinas falsas, y se puede convertir en una curva abierta en un parámetro de escala mayor.

Regla 10:
SI una curva de borde cerrada forma modelos escalera con sus curvas vecinas
ENTONCES esta se puede dividir en varias curvas en un parámetro de escala mayor.

Regla 11:
SI una curva borde no lineal y sus curvas vecinas forman un modelo escalera y una relación posicional anidada
ENTONCES puede desaparecer completamente en un parámetro de escala mayor.

---

1. Γ_j debe existir en σ_i.

2. I_{σ_i} debería contener pocos cruces por cero irrelevantes.

Tabla 3.3: Reglas para eliminación de contornos falsos [Lu y col., 1992].

Regla 12:
SI las curvas deseadas forman un modelo escalera
ENTONCES se generarán falsos bordes en las escalas mayores.
Regla 13:
SI existe una curva de cruces por cero en una escala grande y desaparece en una escala pequeña
ENTONCES es una curva falsa.
Regla 14:
SI una curva borde falsa está presente en σ_1
ENTONCES ésta desaparecerá en un σ, donde $\sigma < \sigma_1$.
Regla 15:
SI una curva cambia de forma significativamente con σ
ENTONCES está formada por ruido.
Regla 16:
SI una curva continúa dividiéndose cuando σ decrece
ENTONCES ésta podría estar formada por ruido.

---

3. $I_{\sigma_1}, \ldots, I_{\sigma_{i-1}}$ debe proporcionar la adecuada información para recuperar la localización y forma adecuada de Γ_j.

4. i debe ser lo más pequeño posible, siempre que cumpla los criterios previos.

Aunque en la literatura están propuestos diversos métodos para la selección de parámetros de escala para contornos de tipo lineal, éstos dejan de ser eficientes cuando las curvas borde no son lineales. Esto nos ha llevado a realizar su selección tras la inspección visual de su comportamiento en diferentes situaciones. Como los problemas de dislocación y de falsos bordes se deben a la influencia mutua de bordes, no sólo de diferentes curvas sino también de la misma curva, se deben tomar valores de σ_{min} pequeños. Nosotros hemos elegido $\sigma_{min} = 1.0$, ya que para valores más pequeños se observan en las imágenes demasiados detalles y contornos ruidosos. Para el valor de σ_{max} encontramos que $\sigma_{max} = 2.0$ es un valor adecuado, en general. Así pues, nuestro sistema considera únicamente dos parámetros de escala, lo cual redundará en un ahorro del esfuerzo computacional.

Una vez que se dispone de las imágenes de los cruces por cero para los valores de σ considerados, se realizará un procesado de los puntos de borde con el fin de obtener la información de contornos más completa y precisa posible en la zona de interés. Para ello se ejecutarán una serie de algoritmos sobre los dos mapas de bordes previos. Estos algoritmos se fundamentan en los conjuntos de reglas descritos en las tablas 3.1-3.3, las cuales resumen conocimiento de bajo nivel a cerca del comportamiento de los bordes en el espacio de escalas, y en otra información que podemos considerar de alto nivel, ya que se corresponde con el conocimiento aproximado sobre formas y localizaciones de los contornos deseados. Este último conocimiento se representará mediante un conjunto de ligaduras. Consideremos una función de contorno $f(x, y)$; las ligaduras que le podemos imponer pueden ser de los siguientes tipos [Taratorin y Sideman, 1993]:

- Ligaduras de amplitud, las cuales establecen el conocimiento sobre límites en amplitud, por ejemplo:

$$f(x, y) \geq A \tag{3.38}$$

- Ligaduras de soporte, las cuales estipulan que los valores de la imagen sólo pueden ser distintos de cero en una región específica del espacio.

- Ligaduras de forma de onda: las cuales implican el conocimiento de la proximidad de la imagen a cierta imagen prototipo $g(x, y)$; por ejemplo:

$$\|f(x, y) - g(x, y)\| \leq \epsilon \tag{3.39}$$

- Ligaduras funcionales, las cuales imponen límites sobre las transformaciones funcionales de las imágenes; por ejemplo:

$$\|Cf(x, y) - g(x, y)\| \leq \delta \tag{3.40}$$

donde C es algún operador lineal, y ϵ y δ representan valores umbrales *suficientemente* pequeños. Estos tipos de ligaduras son ampliamente usados en la teoría de reconstrucción de la señal [Youla y Webb, 1982; Taratorin y Sideman, 1993]. El encontrar una solución al problema de obtención del contorno exterior del hueso es equivalente a encontrar la función $f(x, y)$ de contorno que verifique las ligaduras anteriores. En las imágenes con las que trabajamos los contornos que interesan son los correspondientes a hueso cortical,

mientras que los correspondientes a estructuras no óseas o los que corresponden a hueso trabecular deben eliminarse.

Para la incorporación de información *a priori* definimos la función de contorno $f(x, y)$ como aquella que toma el valor de la derivada en la dirección normal al contorno en los puntos en los que se detecta cruce por cero en la imagen resultante de aplicar el operador de Marr-Hildredth, y valor cero en el resto. Para cada una de las ligaduras anteriores definimos un operador de proyección, que transforma la función de contorno en la función más próxima en el conjunto de ligaduras. Empezaremos considerando una función $g_i^0(x, y) = P_1[f(x, y)]$ como la estimación inicial de los contornos en el corte i, donde P_1 es la función de proyección de la ligadura de amplitud, definida del siguiente modo:

$$P_1[f(x,y)] = \begin{cases} 1; & f(x,y) > Grad_{th} \\ 0; & f(x,y) \leq Grad_{th} \end{cases} \tag{3.41}$$

y $Grad_{th}$ representa un valor umbral de gradiente. De esta forma se eliminan contornos no reales y aquellos que aunque sean reales, se corresponden a cambios de intensidad no significativos, como los que se observan dentro de regiones con texturas (diferentes regiones de la trabécula). Éstas se pueden eliminar, al menos en parte, imponiendo un umbral mínimo para el gradiente sobre los contornos.

El uso de información *a priori* permitirá obtener una estimación consistente del contorno a partir de la estimación inicial $g_i^0(x, y)$. El contorno real ha de estar contenido en una determinada *región de soporte* S de la imagen. Esta región de soporte se obtendrá como una dilatación de la máscara de las estructuras óseas detectadas en el corte anterior, ya procesado. Si $g_i^0(x, y)$ es la estimación actual de la posición de puntos de borde en el corte i y $m_{i-1}(x, y)$ representa la máscara dilatada definida por los contornos finales obtenidos para el corte $i - 1$, será razonable asumir que los puntos de borde verificarán $g_i^1(x, y)$ AND $m_{i-1}(x, y) = 1$. Esta función de contorno se obtendrá utilizando un nuevo operador de proyección $g_i^1(x, y) = P_2[g_i^0(x, y)]$, que definimos de la siguiente forma:

$$P_2[g_i^0(x,y)] = \begin{cases} g_i^0(x,y); & m_{i-1}(x,y) \geq 0 \\ 0; & \text{otro caso} \end{cases} \tag{3.42}$$

donde $m_{i-1}(x, y) \geq 0$ representa puntos interiores a la máscara.

Otro tipo de ligadura semejante a la anterior es la representada por el conocimiento aproximado del contorno verdadero. Si $g_i^1(x,y)$ es la estimación actual de la posición de puntos de borde en el corte i, y $g_{i-1}^f(x,y)$ es el contorno final extraido para la misma estructura en el corte previo, será razonable asumir que los puntos de borde verificarán $\|g_i^1(x,y) - g_{i-1}^f(x,y)\| \leq \epsilon$, siendo ϵ un parámetro fijado de forma heurística a valor 5. Llamando $g_i^2(x,y) = P_3[g_i^1(x,y)]$, $g_i^3(x,y) = P_4[g_i^2(x,y)]$ y $g_i^4(x,y) = P_5[g_i^3(x,y)]$ a la estimación de contorno resultante de aplicar el criterio anterior, con la siguiente definición de los operadores obtendremos:

$$P_3[g_i^1(x,y)] = \begin{cases} 1; & g_i^1(x,y) = 1 \text{ y } \|g_i^1(x,y) - g_{i-1}^f(x,y)\| \leq \epsilon \\ 0; & g_i^1(x,y) = 0 \\ 0.5; & \text{otro caso} \end{cases} \tag{3.43}$$

$$P_4[g_i^2(x,y)] = \begin{cases} 1; & g_i^2(x,y) = 1 \\ 1; & g_i^2(x,y) = 0.5 \quad y \quad g_i^2(x+\delta x, y+\delta y) = 1 \end{cases} \tag{3.44}$$

$$P_5[g_i^3(x,y)] = \begin{cases} 1; & g_i^3(x,y) = 1 \\ 0; & \text{otro caso} \end{cases} \tag{3.45}$$

donde δx, δy toman valores 0 y 1. La ligadura P_4 se aplica de forma iterativa hasta que no haya cambios en la asignación de valores; su objetivo es hacer que los segmentos de borde próximos al contorno que define la forma aproximada se conserven en su totalidad, aún cuando alguna de sus partes esté alejada de esa forma aproximada. De este modo inicialmente se ponen a uno las partes de los segmentos de borde próximas al contorno aproximado y a 0.5 las partes alejadas. En las sucesivas iteraciones se irán poniendo a 1 todos los puntos de un segmento que tuviese inicialmente alguna parte con valor 1. La ligadura P_5 eliminará finalmente los puntos de contorno con valor distinto de 1.

Describamos el algoritmo al que sometemos las imágenes de bordes obtenidas con $\sigma = 1$ y $\sigma = 2$. Mediante la aplicación de los operadores P_1 y P_2 nos quedamos con los bordes significativos dentro de la región de interés en cada uno de los mapas de bordes (uno por cada escala). Dentro del conjunto de bordes resultantes sólo nos interesan los que definen el contorno exterior, por lo que le aplicamos los operadores P_3 y P_4. El operador P_3, al igual que los anteriores, se aplica una sola vez, mientras que el P_4 se aplica de forma iterativa para incluir de forma completa aquellos segmentos de borde que tienen partes próximas al contorno aproximado y otras alejadas. Finalmente, P_5 eliminará los segmentos de contorno sin partes próximas al contorno aproximado. Llegados a este punto

y obtenidos dos mapas de bordes simplificados restringidos a la zona de interés, es cuando comienza el análisis en el espacio de escalas. Este proceso lo representamos mediante la aplicación de un nuevo operador P_6 sobre el conjunto de mapas de bordes previo. La aplicación de este último operador proporcionará un único mapa de bordes de salida en el que se habrá ajustado la localización de los bordes, eliminado los bordes espúreos y recuperados algunos de los bordes perdidos.

Inicialmente se ejecutará un algoritmo basado en las Reglas 1-3 de la tabla 3.1, en el que para cada punto de borde del filtrado con parámetro $\sigma = 2.0$ se buscará su localización más próxima en un entorno de radio 3 de la misma posición en la imagen de bordes de $\sigma = 1.0$. De esta forma obtenemos la localización más precisa de los contornos verdaderos y se eliminan los bordes ruidosos obtenidos con $\sigma = 1.0$, pero no con $\sigma = 2.0$.

A continuación, para recuperar bordes reales no presentes en el filtrado con el parámetro de escala mayor, y considerando las Reglas 4-11 de la tabla 3.2, lo que se hace es recuperar los bordes presentes en la imagen del parámetro de escala menor conectados a aquellos identificados como correspondientes a los bordes del σ mayor. Es decir, para el contorno encontrado en el parámetro de escala mas elevado buscamos el correspondiente en el parámetro de escala más bajo, incluyendo además sus prolongaciones. Este proceso se va repitiendo de forma iterativa hasta que no haya más segmentos adyacentes a los proyectados desde el sigma mayor, o a los añadidos en el proceso de prolongación. Como paso final se unen los extremos de contornos abiertos que estén muy próximos y que se pudieran dividir por los motivos indicados en las reglas que afectan a este proceso. Puede ocurrir que la forma de la curva tenga, en esta parte, una escala menor a la mínima considerada.

La eliminación de falsos contornos se basa en las Reglas 12-16 de la tabla 3.3, muchas de ellas implícitas en el proceso de corrección de localización de posiciones de contornos. En este momento es en el que se procede con la eliminación de contornos abiertos, que en general se van a corresponder con texturas del hueso trabecular.

Aún resta por hacer la extrapolación de los segmentos de borde para obtener los contornos reales no detectados. Para realizar esta tarea se dispone de dos vías de información. La primera la constituye la máscara del corte previo, que permitirá clasificar los contornos actuales como pertenecientes a alguna de las estructuras óseas presentes o como pertenecientes al fondo. Esta clasificación impedirá la extrapolación de contornos

hacia estructuras diferentes. La extrapolación se basará en criterios de proximidad entre vértices de bordes. La segunda fuente de información la aportará la situación actual del crecimiento de regiones, que permitirá el establecimiento de contornos obtenidos en aquellas partes en donde el operador de bordes no detecte transiciones suficientemente fuertes, aún cuando si exista una separación entre diferentes tejidos. Hay que aclarar en este punto que este proceso de integración de información corresponde al sistema de alto nivel, que guía el sistema completo de unión y división de regiones. Es el sistema de control del alto nivel el que decide incorporar información de bordes en el proceso global, para lo cual ha de efectuarse todo el procesado que estamos comentando. Con este procesado de bordes no se pretende dar solución al problema de la segmentación, sólo se persigue la incorporación de información que pueda ser relevante al proceso global. Corresponde al sistema de alto nivel el conjugar todas las fuentes de información de la mejor forma posible y complementarlas. En ocasiones los métodos basados en regiones no advertirán la presencia de contorno que si advierte el método para detección de contornos y viceversa.

Todo el procesado previo se efectuó utilizando directamente las imágenes de bordes. A partir de este punto ya se pasa a trabajar con segmentos de borde como unidad de procesado, por lo que es conveniente disponer la información de bordes en un grafo en el que la relación de adyacencia viene dada por los puntos de borde compartidos entre diferentes segmentos. Sobre la imagen de bordes se ejecuta inicialmente un procesado de adelgazamiento (*thinning*) con el fin de obtener contornos de anchura unidad y conectividad 8. Una vez hecho ésto se asigna un tipo de vértice a cada punto borde: -1 a puntos borde aislados, 0 a aquellos puntos intermedios de un segmento, 1 a los extremos libres de los segmentos, 3 ó 4 a aquellos extremos de segmentos de borde compartidos por 3 ó 4 segmentos (no son posibles comparticiones entre más segmentos en una imagen binaria). Los vértices aislados se pueden eliminar, ya que en general corresponden a procesos ruidosos. Una vez etiquetados según el tipo de vértice, se aplica el algoritmo de construcción de la lista de bordes siguiente:

```
Inicio :
       MIENTRAS ∃(i,j)| Pasado(i,j)!=1:
1.             SI TipoVert(i,j) ≥ 1 IR A 3
2.             SI TipoVert(i,j) = 0 IR A 4 con (i,j)
               Vértice1=(k,l)
               SI (i,j) ≠ (k,l)
```

```
                    IR A 4 con (i,j)
                    Vértice2=(k,l)
                SI NO
                    Vértice2=(k,l)
                    Reordena(Segmento(Vértice1,...,Vértice2))
                FIN SI
                IR A 5
3.              IR A 4 con (i,j)
                Vértice1=(i,j)
                TipoVert(i,j) = TipoVert(i,j) - 1
                SI TipoVert(i,j) = 0 Pasado(i,j)=1
                Vértice2=(k,l)
                IR A 5
4               (k,l) = (i,j)
                MIENTRAS TipoVert(k,l)!=0
                         GUARDAR (k,l)
                         Pasado(k,l)=1
                         BUSCAR borde adyacente (k,l)=(k+m,l+n),con
                                m, n ∈ {−1, 0, 1}
                FIN MIENTRAS
                TipoVert(k,l) = Mínimo(0,TipoVert(k,l) - 1)
                SI TipoVert(k,l) = 0 ENTONCES Pasado(k,l)=1
5.      PASAR Segmento((i,j),...,(k,l)) a GRAFO
        FIN MIENTRAS
```

Con la lista de contornos construida, su manipulación resulta más simple. La eliminación de uno de ellos se consigue eliminándolo de la lista, actualizando el tipo de vértice de sus extremos y uniendo nodos cuando se pasa de compartir el vértice entre tres segmentos a hacerlo sólo por dos.

Inicialmente, a cada pixel de contorno de la imagen se le asigna un valor que indica su tipo: aislado (-1), extremo libre (1), intermedio (0), y compartido ($\geq$3, según nº de contornos confluyentes). Este número será útil a la hora de etiquetar un contorno para indicarnos si podemos marcar un vértice como totalmente recorrido, o a la hora de eliminar contornos abiertos y unir dos contornos cuando el tipo de vértice compartido

es 2; es decir, se ha eliminado alguno de los contornos confluyentes, se ha disminuido el grado del contorno, y ahora sólo está compartido por dos de los contornos iniciales, lo cual indica en realidad que nos queda un único contorno.

Para el etiquetado de los bordes lo que se hace es, una vez encontrado el primer punto de ese borde, buscar los dos extremos. Si el primer punto encontrado es un punto intermedio, se ha de reordenar la lista de puntos intermedios. Cada punto intermedio recorrido se marca como tal, y para los extremos hay que llevar cuenta de las veces que se ha recorrido para marcarlo como tal una vez que el número de veces recorrido alcanza el tipo de vértice. Si el tipo de vértice es -1 quiere decir que ese borde consta de un único pixel, y por tanto no hay que buscar más.

Una vez etiquetados todos los contornos de la imagen, la información relativa a cada uno de ellos se organiza en un marco con la siguiente estructura:

(BORDE
(ETIQUETA (VALOR 34))
(LONGITUD (VALOR 120))
(PRIMER_VÉRTICE (VALOR 120 80))
(SEGUNDO_VÉRTICE (VALOR 84 132))
(TIPO_PRIMER_VÉRTICE (VALOR 1))
(TIPO_SEGUNDO_VÉRTICE (VALOR 1))
(COORDENADAS_PUNTOS (VALOR 34 120 35 120 ... 84 132))
(GRADIENTE_PUNTOS (VALOR 4 50 ... 100))

Esta información es la que se almacena en cada nodo de la lista de bordes, y una vez organizada la información de bordes en la estructura anterior, se continúa con el procesado de los segmentos de borde.

Como comentamos, la cuestión ahora es la de establecer conexiones entre los diferentes segmentos de borde a fin de obtener los contornos externos a las estructuras de interés: tibia y peroné. Para ello se clasifican los segmentos de borde como continuación de alguna de las estructuras presentes en el corte previo: tibia, peroné y fondo. Esta clasificación se establecerá en base al grado de pertenencia a cada una de las clases, que vendrá determinado por medidas de proximidad. Cuando el grado de pertenencia a una de las clases es mucho mayor que a las otras (4 veces mayor), se etiqueta con esa clase. Si un segmento se etiqueta como fondo, quiere decir que no es continuación de ninguna

estructura previa y por tanto se elimina. Esto ocurre, por ejemplo, cuando se encuentran las primeras regiones correspondientes al fémur, que a pesar de dar lugar a contornos reales, no interesan. Cuando alguno tiene un grado de pertenencia relativamente alto con respecto a más de una clase, queda todavía sin clasificar. Para despejar esta incertidumbre lo que se hace es buscar los segmentos con los *extremos* más próximos a éstos sin clasificar; si las dos clases coinciden se le asigna al segmento la misma clasificación, y en caso contrario se clasifica como fondo para su eliminación. El último paso es el de la conexión de los segmentos través de sus *extremos* con los más próximos dentro de la misma clase siempre y cuando la distancia sea menor que 5 pixels.

El esquema anterior da resultados satisfactorios cuando las discontinuidades (*gaps*) entre los contornos no son demasiado grandes, pero falla, por ejemplo, cuando se trabaja con imágenes de rodillas con problemas de pérdida ósea, en donde las transiciones entre regiones de hueso y fondo son muy suaves, o cuando el hueso cortical es muy estrecho. En estos casos, y otros similares, o no se detectan bordes o éstos son fácilmente clasificables como ruido. Por este motivo el sistema global no puede basarse únicamente en información de bordes, como tampoco lo puede hacer considerando únicamente información de regiones. El bloque de alto nivel evaluará el estado del crecimiento de regiones y la información de bordes proporcionada por estos algoritmos para integrarlos con el fin de obtener la mejor solución posible.

En la figura 3.22 se muestra un ejemplo del uso de la información de bordes para guiar el proceso de división de las regiones de tibia y peroné unidas durante el proceso de crecimiento de regiones basado en conocimiento. En la figura 3.22.a se observa la proximidad entre tibia y peroné, lo que provoca su unión durante la fase de crecimiento de regiones, como se muestra en la figura 3.22.b. Advertida esta unión por el sistema de control del alto nivel, éste pone en funcionamiento el bloque para la obtención de información de bordes de bajo nivel. En este bloque se empieza por determinar los mapas de contornos de las figuras 3.22.c-d, que serán procesados de la forma descrita en este apartado para obtener el mapa de la figura 3.22.e. Los contornos definidos en este mapa de bordes serán utilizados por el bloque de alto nivel, junto con el mapa de regiones de la figura 3.22.b, para separar tibia de peroné. El resultado de la división se muestra en la figura 3.22.f. La información de bordes resulta imprescindible para determinar el lugar por el que se debe establecer la separación entre tibia y peroné una vez que el sistema de control advierte la unión indebida de ambas estructuras. En esta imagen se puede

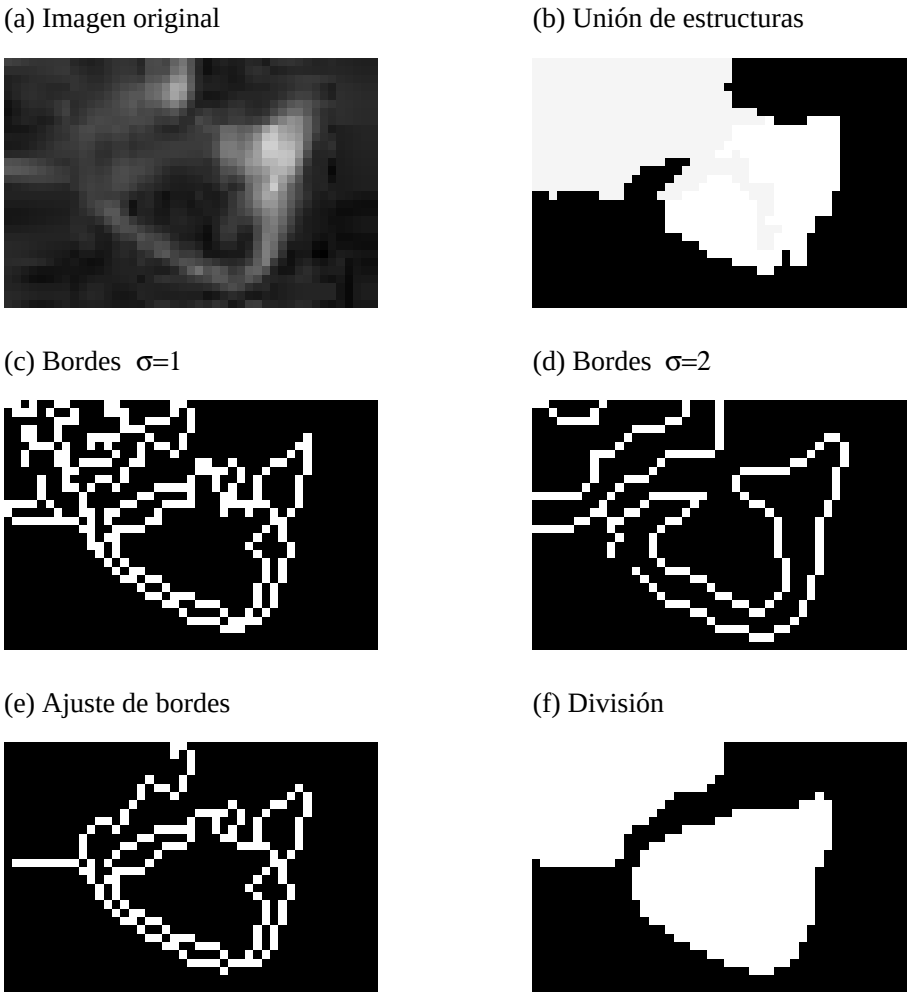


Figura 3.22: Ejemplo del uso de la información de bordes para la división de regiones. (a): imagen original, (b): situación conflictiva en la etapa de crecimiento de regiones que dispara un proceso de división. (c) y (d): mapa de cruces por cero para $\sigma = 1$ y $\sigma = 2$. (e) resultado final del procesado de bordes de bajo nivel. (f) resultado de la operación de división efectuada por el sistema de alto nivel.

observar además como la información proporcionada por el método basado en regiones y la proporcionada por la detección de bordes se complementa para llevar a cabo la división. Este proceso se ejecuta dentro de la etapa de alto nivel de la forma que se indicará después.

3.4.4 Crecimiento de regiones basado en conocimiento

Como resultado de la presegmentación de bajo nivel hemos obtenido un conjunto de regiones elementales. Como hemos podido observar, estas regiones no coinciden totalmente con los elementos significativos existentes en la imagen, pero sí constituyen el punto de arranque de un nuevo proceso de análisis que persigue identificar las estructuras de interés. Este análisis, basado en modelos, integra de forma explícita conocimiento del dominio, que formulamos mediante reglas de producción y que conduce procesos de unión y división de regiones. Así, una vez realizado el procesado de bajo nivel y creado el grafo de adyacencias de regiones, se inicia la etapa de alto nivel en la que se efectuará el proceso de crecimiento de las regiones resultantes de la presegmentación. Este proceso se puede resumir a grandes rasgos de la forma que se indica a continuación. Dada la continuidad 3D de la estructura, la forma y la posición del hueso en la imagen del corte actual ha de ser semejante a la forma y posición en el corte previo. Por tanto, adoptamos como modelo de las estructuras a buscar en el presente corte la zona de hueso extraida en el corte previo. El análisis se inicia definiendo un área de interés dentro de la imagen mediante la dilatación de la máscara del modelo. Así, de los miles de regiones presentes en la segmentación de bajo nivel, nos quedamos sólo con aquellas cuyo centroide cae dentro de ese área, que constituye el área de soporte de las estructuras buscadas. Queda definido de esta forma un foco de atención inicial.

Dentro del área foco de atención se escogerán los núcleos de crecimiento de las estructuras buscadas. Estos núcleos se corresponderán con las regiones correspondientes a las clases más densas dentro de la clasificación proporcionada por la red neuronal. Esta elección se basa en el conocimiento de que las regiones de hueso cortical son siempre las más densas de las regiones correspondientes a la parte anatómica de la imagen, y de que en todas las imágenes a procesar hay estructuras óseas. El problema está en que el hueso cortical no es totalmente homogéneo, ya que presenta diferentes densidades que dan lugar a clases diferentes, y que a veces es demasiado estrecho como para distinguirlo del hueso trabecular. Éste, a su vez, resulta imposible de distinguir de regiones correspondientes al músculo si no tenemos en cuenta información estructural. Estos núcleos de crecimiento se clasifican como núcleos de crecimiento de tibia o peroné dependiendo del área de soporte en la que estén. El conjunto de núcleos de crecimiento de una misma clase constituyen el estado inicial de un agregado óseo (tibia o peroné). En la figura 3.23 se ilustra la determinación de los núcleos de crecimiento para el agregado óseo de tibia. En las fig-

uras 3.23.a-b se muestra la imagen original y el resultado de la presegmentación inicial. En la figuras 3.23.c se representa el modelo para tibia y peroné, resultado del procesado del corte previo, y en la figura 3.23.d se representan con el nivel de gris más alto los núcleos de crecimiento para la tibia, cuyo conjunto define el agregado de tibia en su fase inicial.

Las decisiones sobre la incorporación de nuevas regiones a un núcleo de crecimiento se hacen teniendo en cuenta las propiedades del agregado completo y su aproximación creciente al modelo, y no únicamente las de núcleos aislados. La razón de ello es que el agregado en conjunto transporta información estructural significativa de la que, en general, no disponen los núcleos de crecimiento de forma individual, al menos en las primeras fases de crecimiento. Entorno a estos núcleos se inicia pues el crecimiento del área de hueso mediante procesos de unión con sus vecinos. Esta unión se implementa como un procesado de las regiones de la clase inmediatamente anterior a la última ya procesada (inicialmente se corresponderá con la inmediatamente inferior en densidad media a la más baja considerada como parte del núcleo de crecimiento). Sobre este conjunto de regiones, que constituirán el *foco de atención* en fase de crecimiento, se aplica un conjunto de reglas que determinarán la conveniencia de su unión o no a los núcleos o su consideración como nuevo núcleo de crecimiento asignado al agregado correspondiente. Cada vez que se produce la incorporación de una región a los núcleos, se adapta el RAG a la nueva situación y se recalculan las propiedades de los agregados.

Las regiones de esa clase no incorporadas después de aplicar el conjunto de reglas relativo a la unión de regiones no serán consideradas de nuevo para el mismo agregado óseo y todos los huecos presentes en los núcleos de crecimiento se añaden a éstos. Con ello evitamos hacer el procesado de las regiones presentes en ese hueco, ya que sólo estamos interesados en contornos externos, y a la vez reducimos tiempo de procesado.

El proceso de crecimiento de regiones continúa hasta que se produce una correspondencia (*matching*) con el modelo, terminando en ese instante el proceso en ese corte. El grado de correspondencia se evalúa mediante medidas de distancia de forma entre los núcleos que constituyen el agregado y el modelo. La razón de no comparar todo el agregado con el modelo en esta etapa de evaluación es que el agregado puede presentar una buena semejanza con el modelo a pesar de estar divido en varios núcleos; por tanto, si la comparación se hace con los núcleos, la correspondencia con el modelo sólo se producirá después de haberse unido un número adecuado de núcleos para dar lugar a una región

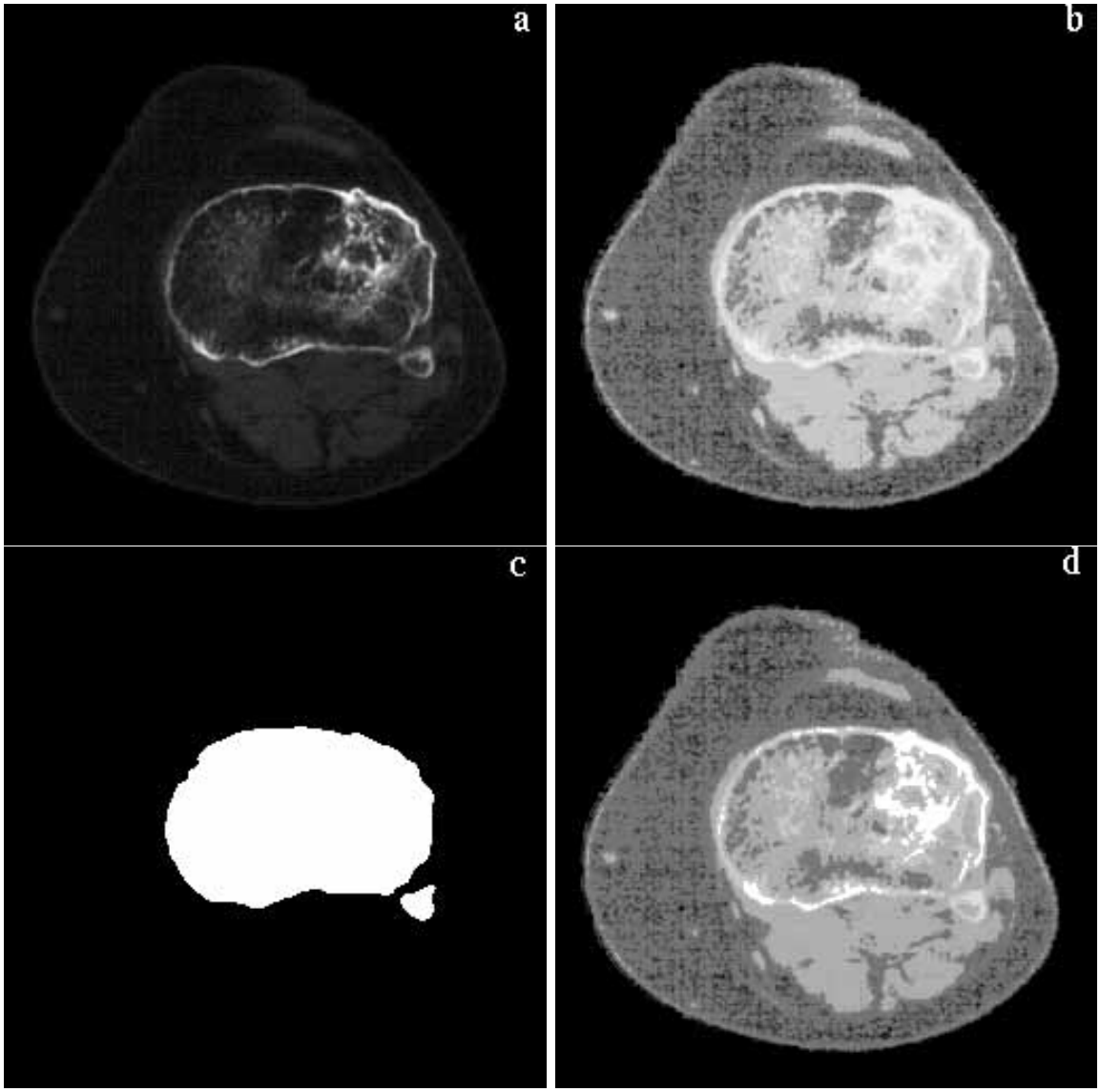


Figura 3.23: (a) Imagen original; (b) presegmentación; (c) modelo; (d) núcleos iniciales para crecimiento de tibia (regiones con nivel 255).

conexa que se asemeja al modelo. Las medidas de distancia de forma las introduciremos en la siguiente subsección. Si en el procesado de una clase de regiones no se produce ninguna incorporación a los agregados de hueso, se selecciona una nueva clase de regiones y se repiten los pasos del proceso de unión hasta agotar todas las clases si no se obtuvo antes

una adecuada correspondencia con el modelo. Si ésta no se alcanzase, se incorporaría conocimiento proveniente de la detección de bordes.

Esta estrategia corresponde al procesado de cortes intermedios, en los que la unión de regiones se hace utilizando un modelo aproximado obtenido de la segmentación concluida para el corte previo. En el corte inicial no existe este modelo; por lo tanto, hay que cambiar la estrategia. Nuestro sistema comienza el procesado siempre por la parte distal de la tibia, ya que es en esa parte donde las componentes corticales son más densas y anchas, por lo que la delimitación de los huesos es mucho más simple que en la parte proximal, donde nos podremos encontrar con problemas de pérdida ósea, lesiones, malformaciones y presencia de otras estructuras óseas (fémur). La única información de que se dispone para el procesado del primer corte es la de la posición central de la tibia en la imagen de entrada, la posición relativa del peroné con la tibia, tamaños mínimos y conocimiento de alto grado de compactación. Esto último se puede asegurar imponiendo como exigencia en la etapa de exploración que el primer corte esté suficientemente alejado de la meseta de la tibia, lo cual va a ser también un requisito para la satisfacción de las adecuadas condiciones de contorno en el análisis mediante Elementos Finitos.

Partiendo del conocimiento a cerca de las estructuras concretas a identificar, la reconstrucción y reconocimiento de tibia y peroné se realiza en fases debidamente ordenadas. Para la identificación de la tibia, éstas son:

1. Identificación de los núcleos de crecimiento como las regiones correspondientes a la clase más densa dentro del área de soporte definida a partir de la máscara de tibia del corte anterior. Estos núcleos constituyen el agregado inicial de tibia.

2. Crecimiento de los anteriores núcleos en base a la unión de regiones que:

   - aumenten la compactación, en cierto grado, del agregado de tibia,
   - aproximen ciertas características (área, orientación, etc.) del agregado a las del modelo.
   - unan dos o más regiones hipótesis de hueso,
   - sean interiores a algún núcleo hipótesis de hueso.

3. Comparación con el modelo. Aún cuando a partir de un determinado momento, el agregado presente una gran proximidad al modelo, lo cual indicaría el final del

proceso, aquél puede estar fraccionado en diversos núcleos. Para evitar parar el proceso es necesario que la comparación se haga entre los núcleos individuales y el modelo. Si ninguno verifica esta prueba de correspondencia, se pasa a considerar las regiones de la clase que representa la densidad inmediatamente inferior a la ya procesada. Se repite así de nuevo el proceso de crecimiento.

4. Si se verifica la correspondencia de un núcleo con el modelo, se reconoce ese objeto como tibia, y se pasa a la identificación del peroné.

Para la identificación de peroné los pasos a seguir son los ya indicados para el reconocimiento de la tibia, pero ejecutados sobre las regiones cubiertas por su correspondiente área de soporte.

Una vez que se han identificado las estructuras de interés, y antes de obtener el contorno externo de hueso cortical de la tibia en ese corte, el siguiente paso consiste en la regularización de formas. Para ello se procede a realizar un cierre morfológico de las estructuras por separado, la extracción y suavización de contornos y la obtención de máscaras.

La figura 3.24 muestra, a nivel de diagrama de bloques, la estructura del sistema que implementa las ideas y estrategias de análisis comentadas. Este sistema consta de un conjunto modular de procesos, más dos memorias al estilo del sistema propuesto por Nazif y Levine [1984]. Éstas son:

- **Base de datos** (*short term memory* (STM)), que contiene las imágenes de entrada, los datos de la segmentación, y la salida. Esta es una memoria de lectura y escritura. Su contenido varía a lo largo del proceso. Esta base de datos coincide con la mostrada en la figura 3.11, que mostraba la estructura global del sistema.
- **Base de conocimiento** (*long term memory* (LTM)), que contiene el conocimiento específico del dominio y estrategias de control. Esta es una memoria de sólo lectura. Su contenido no cambia a lo largo del proceso. La base de conocimiento se corresponde con el bloque de conocimiento sobre el dominio que aparece en la figura 3.11.

Cada proceso tiene acceso a la LTM y a la lectura y modificación de la STM, lo cual permite en sí la comunicación entre procesos.

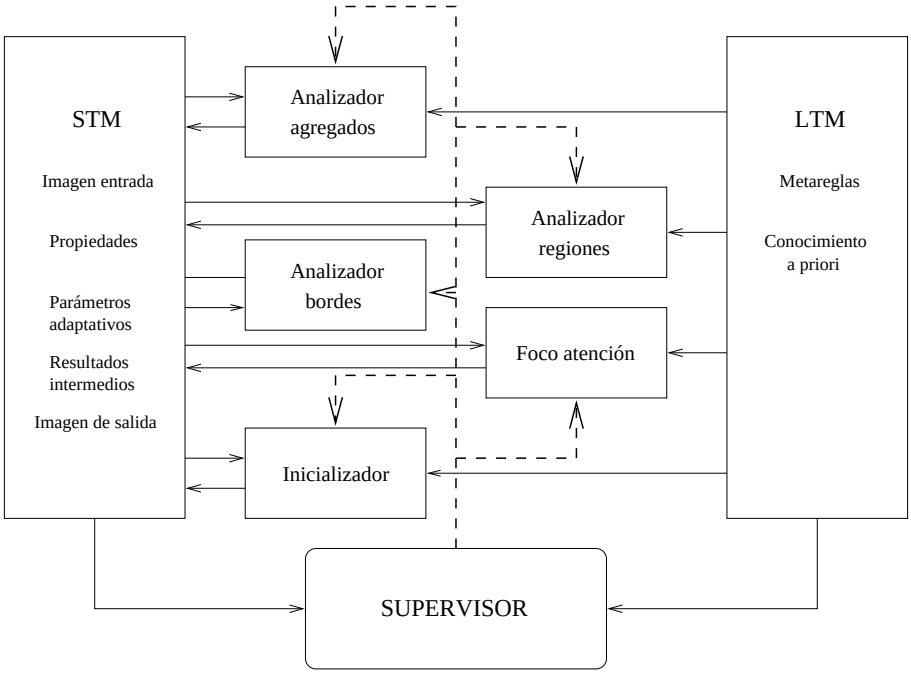


Figura 3.24: Descripción general del sistema de análisis de regiones basado en reglas.

La estructura de control básica es la de un sistema de producción en el que el conocimiento sobre propiedades de bajo nivel independientes del dominio y alto nivel dependientes de la imagen concreta se formula como un conjunto de reglas condición-acción. Sobre la base de datos se analiza la parte de la condición y sobre ella se ejecuta también la acción, que provocará cambios en su contenido. Un proceso del sistema selecciona reglas de LTM basándose en datos de STM y conocimiento de alto nivel de la LTM, y cuando se satisface la condición, se iniciará la acción asociada con esa regla, modificándose con ello la STM.

En el procesado inicial de bajo nivel se realiza una sobresegmentación de la imagen, que clasifica los pixels de acuerdo con la densidad de su vecindad, teniendo en cuenta también la distribución global de las intensidades. En la STM se guarda información tal como la imagen real, la imagen presegmentada, propiedades explícitas de las regiones (nivel medio, dispersión, adyacencias, área, perímetro, centro de masas, segundos momentos centrales) y otras propiedades derivadas de las anteriores (compactación, elongación, posición relativa, forma), en la forma de un grafo de adyacencias (RAG).

La información contenida en la STM se actualiza continuamente durante el procesado,

de tal forma que refleje el estado de la segmentación en cada instante. La segmentación evoluciona a modo de una transición de estados, en la que el estado actual queda definido por el contenido de la STM, la entrada actual es la regla seleccionada en la LTM, y la salida será el nuevo estado de la STM.

El modelo de la LTM está compuesto de dos niveles de reglas de producción:

1. Reglas de conocimiento: codifican la información sobre propiedades de regiones en la forma de pares situación-acción.
2. Reglas de control, entre las que podemos distinguir:
   (a) Reglas de foco de atención: que se encargan de seleccionar los datos a procesar.
   (b) Metareglas. Éstas no modifican datos de la STM, sino que determinan qué reglas se van a evaluar, atendiendo a los datos de la STM.

En el esquema de la figura 3.24 se refleja, a parte de las dos memorias comentadas, toda una serie de módulos encargados de realizar acciones concretas. Los módulos involucrados en la producción de acciones son los siguientes:

- Inicializador: es el encargado de generar el mapa inicial de regiones y calcular y almacenar las propiedades de los datos iniciales de la STM. Se corresponde con el bloque de bajo nivel de la figura 3.11.
- Foco de atención: determina el área de soporte de cada objeto, los núcleos de crecimiento de cada uno de ellos y las regiones a analizar en cada fase del procesado.
- Analizador de regiones: es el módulo encargado de realizar las operaciones de unión de regiones y actualización de propiedades.
- Analizador de agregados: chequea, una vez finalizado el proceso de crecimiento, la posible existencia de uniones entre agregados óseos diferentes. Efectúa la división para casos sencillos.
- Analizador de bordes: entra en funcionamiento siempre que no existe correspondencia entre los objetos y el modelo tras finalizar el proceso de crecimiento de regiones. Este bloque realiza el proceso de división cuando se produce una unión entre tibia y peroné, o de alguna de estas estructuras con otros objetos.

- Supervisor: realiza las tareas de control del sistema.

En los apartados siguientes vamos a describir las medidas de distancia de forma y la estructura de las reglas de conocimiento. Con posterioridad describiremos la actuación de los distintos módulos de procesado a través del análisis del conjunto de reglas que incorporan.

Formulación del modelo y medidas de distancia de forma

La abstracción, representación y modelado de estructuras espaciales se ha introducido en el área de la robótica con considerable éxito. Estas aplicaciones se benefician con frecuencia de la simplicidad de tratar con un entorno estructurado consistente en objetos manufacturados, muchos de los cuales ya existen en el formato de modelos CAD. En el campo de los datos médicos multidimensionales resulta mucho más difícil el definir buenas estrategias para la interpretación basada en conocimiento. Aquí, el modelo no se puede definir de una forma precisa por la variabilidad de estas estructuras, por lo que hay que recurrir a modelizaciones en forma de ligaduras.

En nuestro caso, la posición, la orientación y el tamaño de las estructuras entre cortes adyacentes presentan pocas variaciones, por lo que parece adecuado formular un modelo en el que se representen estas propiedades para el corte previo al actual. Comparando las características de un agregado con el modelo podremos decidir si la inclusión de una nueva región es adecuada o no. Al mismo tiempo, podremos comparar núcleos individuales con el modelo para ver si se alcanzó la correspondencia, y así finalizar el proceso.

Como ya hemos comentado, el modelo de estructura que se manejará, y con el que se comparará el estado de la segmentación, es el que corresponde a la segmentación final obtenida para la misma estructura en el corte previo. Este modelo se representa mediante un marco con el conjunto de características anteriores, tal y como se comentó en el apartado 3.4.2. Por otra parte, es preciso modelar el agregado de regiones correspondiente a esa estructura y que define, tras cada iteración, el resultado de la evolución de proceso de crecimiento de regiones. Para poder realizar el proceso de comparación entre ambos es preciso calcular el mismo conjunto de propiedades con el que se caracteriza al modelo para el agregado de regiones. Estas propiedades deben ser recalculadas tras cada proceso de unión. Características útiles en la comparación con un modelo serán el centroide

(posición), los segundos momentos (orientación) y el área (tamaño), [Dhawan y Arata, 1991].

Dado un grupo de regiones que se identifican como hueso en una primera iteración, la tarea en las siguientes iteraciones es añadir o eliminar regiones vecinas con el fin de mejorar las características en relación al modelo. Mediante las medidas de distancia de forma se proporciona una realimentación desde el modelo para mejorar el proceso de correspondencia ascendente (*matching bottom-up*) inicial. Por tanto, aún cuando algunas de las características de las regiones iniciales no son significativas, sí lo serán cuando se van acumulando al formarse e ir creciendo los agregados de regiones.

La definición y cuantificación de distancias de forma para objetos 2D y 3D es un paso crucial en el desarrollo de sistemas de visión basados en modelos, cualquiera que sea su aplicación [Banerjee y Majumdar, 1996]. Estas medidas darán idea de la correspondencia entre las estructuras que aparecen en cada momento del proceso de segmentación y el modelo de las estructuras de interés. En las aplicaciones con imágenes médicas, las medidas de distancia de forma se emplean sobre todo en el establecimiento de correspondencias entre imágenes del mismo objeto captadas con dos modalidades de imagen diferentes y para la comparación de una imagen con su modelo, como es el caso que nos ocupa [Dhawan y Arata, 1991; Banerjee y Majumdar, 1996]. A continuación vamos a definir la métrica 2D que nosotros usaremos para la determinación de la proximidad de una región binaria a una máscara modelo. Una simplificación importante con la que vamos a contar es que ni la región resultante de la segmentación ni el modelo presentan huecos; éstos se pueden eliminar sin más que incluir en las máscaras las regiones interiores. Para establecer el valor global de la distancia de forma de una estructura (agregado o núcleo) al modelo se realiza la medida de tres distancias: 1) la *distancia del centroide* (D_C), que es la distancia Euclídea entre los centroides de la estructura y el centroide del modelo; 2) La *distancia de área* (D_A), que es una medida de coincidencia de área entre la estructura y el modelo; y 3) la *distancia de momentos* (D_M), que mide la distancia Euclídea entre los momentos de ambos.

$$D_C(m,r) = \sqrt{(\overline{X}_m - \overline{X}_r)^2 + (\overline{Y}_m - \overline{Y}_r)^2} \tag{3.46}$$

$$D_A(m,r) = \min\{1.0, \frac{|A_m - A_r|}{A_m}\} \tag{3.47}$$

$$D_M(m,r) = \sqrt{\sum_{i,j=0}^{1}(M_m^{ij} - M_r^{ij})^2} \tag{3.48}$$

donde el subíndice m hace referencia al modelo y r a la estructura detectada. Estas tres medidas se integran dentro de una medida de *distancia final* (DF) como una suma ponderada:

$$D_F(m,r) = \alpha\frac{D_C(m,r)}{\sqrt{\frac{A_m}{4\pi}}} + \beta D_A(m,r) + \gamma\frac{D_M(m,r)}{\sqrt{\sum_{i,j=0}^{1}(M_m^{ij})^2}} \tag{3.49}$$

Los términos que aparecen dividiendo a D_C y a D_M son factores de normalización. De los dos, el que puede precisar aclaración es el término $\sqrt{\frac{A_m}{4\pi}}$; éste representa una estimación del radio medio del modelo. En la ecuación de *distancia final* se ha de verificar que $\alpha + \beta + \gamma = 1.0$, para tener un valor de distancia de forma en el intervalo $[0, 1]$.

La correspondencia se produce si y sólo si las dos entidades tienen posición, tamaño, orientación y forma semejantes. En nuestro caso, dado que no hay cambio de escala, ni tampoco grandes variaciones en la posición y en la orientación de las estructuras óseas de un corte a otro, no es adecuado considerar la invarianza en tamaño, posición y orientación para la medida de distancia de forma, sino que, al contrario, son factores determinantes. El grado de similaridad de forma μ se define como

$$\mu(m,r) = 1 - D_F(m,r) \tag{3.50}$$

La medida de distancia de forma del agregado y el modelo se efectúa durante los procesos de unión. Sin embargo, cuando se pretende realizar la comparación para decidir si se alcanzó la correspondencia o no, la comparación se efectúa entre las distancias de forma de los núcleos individuales y el modelo. Si para un núcleo no se verifica alguna de las condiciones siguientes,

$$D_A < \frac{1}{3}A_m \tag{3.51}$$

$$D_C < \frac{1}{3}\sqrt{\frac{A_m}{4\pi}} \tag{3.52}$$

$$D_M < \frac{1}{3}\sqrt{\sum_{i,j=0}^{1}(M_m^{ij})^2} \tag{3.53}$$

se considerará que no se corresponde con el modelo, sin necesidad de determinar el grado de similaridad. Si en un determinado momento la medida de similaridad de un objeto con el modelo es mayor que 0.95, se considerará completamente identificado. Si ésto no ocurre, se prosigue con el proceso de unión. Si también finaliza éste sin que para alguno de los objetos la medida de similaridad alcance el umbral mínimo, entonces se integrará esta segmentación con la resultante de la aplicación de modelos deformables.

Por otra parte, se puede dar la situación de que exista un único modelo de objeto en un corte y se detecten dos o más regiones de hueso separadas correspondientes a esa misma estructura; en este caso ninguna de ellas cumplirá por separado la condición de correspondencia con el modelo. Esta situación se da en la parte proximal de la tibia cuando, avanzando en el sentido distal-proximal, aparecen lesiones que puedan provocar una bifurcación de la estructura. La figura 3.25 ilustra uno de estos casos originado por una lesión, en ella se muestran los contornos finales detectados por crecimiento de regiones sobre las imágenes originales. En este caso hay mucha diferencia entre los cortes. En general, cuando esto ocurra se deben generar imágenes interpoladas, de tal modo que los cambios de forma entre cortes adyacentes no sean muy bruscos, haciendo la segmentación más fiable. No obstante, ahora la medida de proximidad deja de tener valor por la gran diferencia existente entre las estructuras detectadas por separado con el modelo. Se define entonces una nueva medida de valoración que integrará aspectos tales como densidad, área y continuidad 3D (solape con la máscara del modelo). El chequeo del resultado de la segmentación tendrá que ser ahora determinado por datos y no por modelo. En este caso se usará siempre información de bordes para dar mayor fiabilidad a la forma de las regiones obtenidas. Esta información de bordes podrá sugerir la conveniencia de realizar divisiones dentro de las regiones resultantes del crecimiento.

De las regiones que al final resulten nos quedaremos con aquellas que presenten un grado de solape (S_A) mayor del 50%, un área (A) de más de 20 pixels y una densidad (ρ) media superior a 25. La medida de grado de solape se define como

$$S_A = \frac{A_r \cap A_m}{A_r} \tag{3.54}$$

si el centro de masas de la región se proyecta sobre la máscara del modelo. En otro caso el grado de solape valdrá 0. Aquí se considera información de densidades (ρ), ya que en esta zona no existe hueso trabecular y la densidad media de zona ósea es alta. Esto permitirá

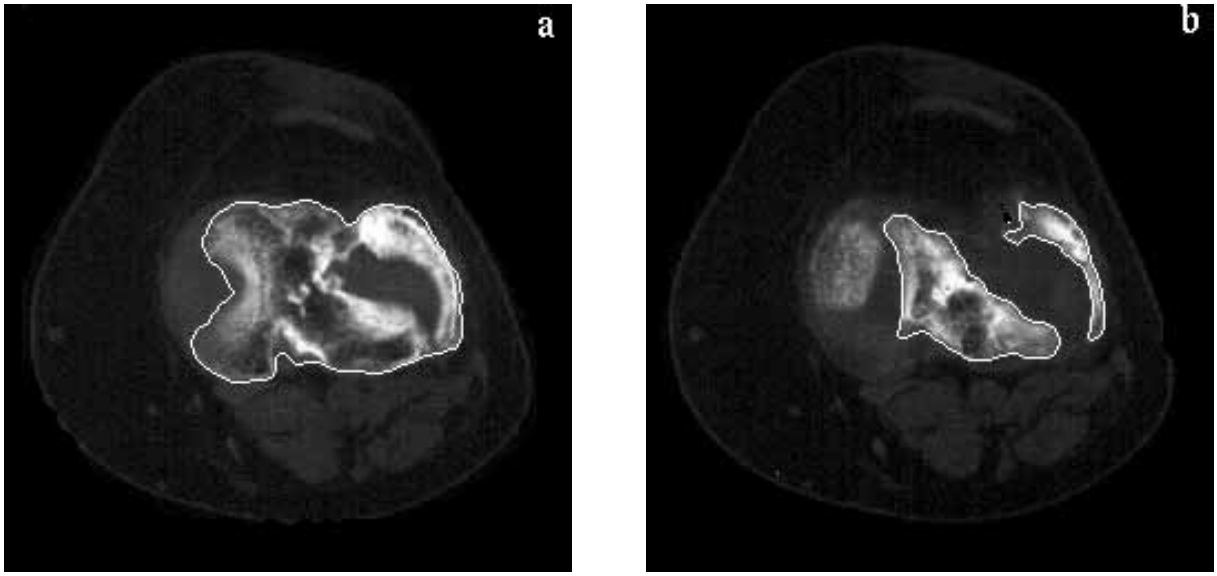


Figura 3.25: Ejemplo de bifurcación de tibia: (a) corte con una única región de tibia en la que se observa una lesión; (b) corte con dos regiones de tibia separadas.

eliminar falsos núcleos. Como en este caso no existe modelo aproximado válido, se debe hacer uso de la mayor cantidad de información posible proporcionada por los datos. Por lo tanto, en el punto de bifurcación se realiza siempre la integración con la segmentación basada en modelos deformables. Las regiones resultantes de la integración definirán las bifurcaciones de tibia.

En el caso del corte inicial no tiene sentido definir medidas de similaridad, dado que no hay modelo y además, al tratarse de un corte suficientemente alejado de la parte proximal de la tibia, las regiones de tibia y peroné son fácilmente identificables. De las dos regiones de mayor área, la tibia será la región más próxima al centro de la imagen, y la otra será el peroné. De esta forma ya habremos reconocido las estructuras y dispondremos de los dos modelos iniciales que se irán actualizando a medida que progrese la segmentación. En cualquier caso, se podría obtener un modelo inicial trazado a mano.

Estructura general de las reglas de conocimiento

Como ya hemos indicado, el sistema diseñado consta de un conjunto modular de reglas y dos memorias: la **STM**, en la que se almacenan las imágenes de entrada, los datos de la segmentación y, al final, las imágenes de salida, y la **LTM**, que contiene el modelo de conocimiento y la estrategia de control.

El conocimiento almacenado en la LTM se compone de 2 niveles de reglas de producción. En el primer nivel, las reglas de conocimiento codifican la información sobre propiedades de regiones en la forma de pares *situación–acción*. Cada situación se codifica como un conjunto de predicados lógicos que componen la parte de condición de las reglas. Cuando se da una determinada situación, es decir, se cumplen todas las condiciones simultáneamente, se dispara una regla. Esto es, se inicia la ejecución de la parte de la acción de la regla. Esta acción puede ser, por ejemplo, la unión de dos regiones o la asignación de etiquetas de hueso.

En el segundo nivel de reglas de la LTM están las reglas de control, las cuales se pueden clasificar en dos categorías : reglas de foco de atención y metareglas. Las reglas de **foco de atención** son responsables de buscar los siguientes datos a considerar: líneas, regiones, áreas, etc. Las **metareglas** se diferencian de otras reglas en que sus acciones no modifican datos en la STM, sino que deciden el orden de la aplicación de los diferentes conjuntos de reglas de conocimiento. Cada módulo de procesado del sistema tiene asociado un conjunto de reglas específico. Las condiciones de las metareglas examinan los datos de la STM y sus acciones especifican el siguiente proceso, y por consiguiente, el conjunto de reglas a activar. Las metareglas son también responsables de evaluar el criterio de fin del procesado. Las reglas de control ejecutan una estrategia de análisis mixta conducida por datos y por modelo; las reglas de foco de atención definen el método de selección de datos a procesar, y las metareglas especifican el orden en el cual se chequean las reglas.

Las condiciones de las reglas son entidades borrosas (*fuzzy*) a las que se les puede asignar diferentes factores de certeza en su cumplimiento. La STM contiene principalmente variables numéricas que representan características de regiones y bordes en la segmentación (área de región, longitud del borde,...). A parte de éstas, se almacenan también datos no numéricos (adyacencia, hueco,...). Utilizamos cuatro tipos de predicados en las condiciones de las reglas:

- Comparaciones sobre variables numéricas: por ejemplo, de áreas $A_{reg} < A_{th}$.
- Comparaciones lógicas sobre variables no numéricas: por ejemplo, reg_a `pertenece al mismo agregado de regiones que` reg_b.
- Evaluaciones no numéricas sobre variables: por ejemplo, `elongación alta`.
- Evaluaciones lógicas sobre variables no numéricas: por ejemplo, reg_i `une` reg_k y reg_l.

Cada regla representa conocimiento estructural, relacional y espacial sobre cada objeto, escribiéndolo como un par múltiple–condición/acción. La parte de la condición tendrá varias premisas que incorporan el conocimiento sobre el objeto, pero no todas serán igual de significativas en el proceso de reconocimiento de los objetos. Por ejemplo, si se identifica la tibia como el objeto de referencia, el reconocimiento del peroné estará más influenciado por las propiedades relacionales que por sus propias características estructurales. Así, a cada premisa p se le asigna un cierto factor de contribución FC_p, según la importancia de la propiedad usada para el reconocimiento del objeto específico. Cuando se chequea una regla, se calcula un factor de verdad (o confianza) FV_p en el cumplimiento de cada una de sus condiciones. Finalmente, a cada regla se le asigna un factor de verificación de la condición VC, que se basa en la suma de los factores de verdad de cada premisa multiplicados por los factores de contribución correspondientes. Si hay N premisas en una regla con los consiguientes factores de contribución FC_i, $i = 1, \ldots, N$, el grado de verificación de la condición VC se calcula como [Dhawan y Juvvadi, 1990]:

$$\mathrm{VC} = \sum_{i=1}^{N} (\mathrm{FC}_i \times \mathrm{FV}_i) \tag{3.55}$$

donde FV_i, $i = 1, \ldots, N$, son los factores de verdad del cumplimiento de las premisas, con valores entre 0.0 y 1.0 y

$$\sum_{i=1}^{N} \mathrm{FC}_i = 1 \tag{3.56}$$

El cálculo anterior es válido para el caso de que las condiciones estén ligadas mediante el operador lógico AND. Cuando la relación lógica entre condiciones se establece mediante

el operador OR, el factor de verificación de la condición asociado a esta relación se obtiene del siguiente modo:

$$VC(P_a \text{ OR } P_b) = \max\{VC_a, VC_b\} \tag{3.57}$$

donde P_a y P_b son premisas individuales o conjunto de premisas relacionadas por el operador lógico AND, por lo que en este caso

$$\text{VC}_c \;=\; \sum_{j \in c} (\text{FC}_j^c \times \text{FV}_j^c) \tag{3.58}$$

donde, de nuevo:

$$\sum_{j \in a} \text{FC}_j^c \;=\; 1 \tag{3.59}$$

donde c representa un conjunto de premisas del tipo de a y b. Los factores de contribución asignados a cada una de las premisas de las reglas se establecen mediante métodos heurísticos.

A partir del factor de verificación de la condición definimos, para cada regla, un grado de certeza (GC) que, en general, podrá tomar valores entre -1.0 y 1.0. La forma de calcularlo dependerá de la regla concreta y lo describiremos posteriormente.

Consideremos una regla escrita con el siguiente formato:

SI A **ENTONCES** X

donde A es la parte de condición y X es la conclusión. Como hemos indicado, antes de ejecutar la regla es preciso evaluar la parte de condición, lo que supone el cálculo de su grado de certeza. Si la parte de condición es cierta, es decir, el grado de certeza toma valor 1.0, entonces se ejecuta la regla considerando la conclusión (acción) acertada. En caso contrario, no se ejecuta la regla. No obstante, el grado de certeza obtenido como resultado de evaluar la parte de condición de la regla puede reforzar o debilitar la conveniencia de tal acción. Para poder contemplar esta posibilidad a cada conclusión se le asociará un valor de certeza C, que almacenaremos en el RAG, en la parte de etiqueta del marco correspondiente a la región sobre la que se evalúa la regla. La asignación de este valor

tiene sentido, ya que es posible llegar a la misma conclusión (acción) después de evaluar premisas (condiciones) diferentes. Por consiguiente, cuando no se alcance el umbral de certeza (1.0) en la evaluación de la condición de una regla, ésta no se ejecuta, pero si que debilitará o reforzará la certeza de su conclusión. Para ello actualizaremos la certeza de la conclusión del siguiente modo [Frost, 1987]. Llamemos $C(X)$ a la certeza de una conclusión en un determinado instante y GC al grado de certeza de una regla que tiene X como conclusión, entonces el nuevo valor de certeza de X, que denotamos por $C(X/A)$ será:

1. Si $C(X)$ y GC son ambos mayores o iguales que 0.0, $C(X/A)$ se calcula de acuerdo con la siguiente ecuación:

$$C(X/A) = C(X) + [\mathrm{GC} * (1.0 - C(X))] \tag{3.60}$$

   Si la certeza $C(X)$ de una sentencia es positiva, entonces lo más que esa regla con valor de certeza positivo GC puede incrementar la certeza de X es $1.0 - C(X)$.

2. Si $C(X)$ y GC son ambos menores que cero, entonces el nuevo valor de certeza de X se calcula de acuerdo con la siguiente ecuación:

$$C(X/A) = C(X) + [\mathrm{GC} * (1.0 + C(X))] \tag{3.61}$$

3. Si $C(X)$ y GC son de signo opuesto, entonces el nuevo valor de certeza de X se calcula de acuerdo con la siguiente ecuación:

$$C(X/A) = \frac{C(X) + \mathrm{GC}}{1.0 - \min\{|\mathrm{GC}|, |C(X)|\}} \tag{3.62}$$

En definitiva, si al final del proceso de análisis de una región no se llegó a ejecutar ninguna acción sobre ella, $C(X)$ reflejará el resultado de evaluar todo el conocimiento aportado por las distintas reglas chequeadas. En consecuencia, una nueva regla chequea su valor y decidirá si la región se etiqueta como hueso (tibia o peroné) o fondo.

Descripción de las reglas

Cada módulo del sistema basado en reglas, cuyo esquema mostramos en la figura 3.24, tiene un conjunto de reglas asociado. Cada regla especifica al módulo de procesado correspondiente la acción a ejecutar cuando se cumplen todas sus condiciones. Según la entidad objetivo de las acciones, podemos clasificar las reglas en diferentes conjuntos:

- Acciones de foco de atención: determinan, en cada momento, la región más adecuada para ser examinada. Esto hace que las acciones se ejecuten sobre áreas localizadas y no de forma general, lo que podría provocar que una acción destruyera progresos previos sobre ciertas regiones de la imagen, al tiempo que se reduce el coste temporal de procesado, al no considerar regiones que no sean de interés.

- Acciones sobre regiones: unión y división de regiones. Este conjunto de acciones se reparte entre el analizador de regiones encargado de la unión y el analizador de bordes, que puede provocar la división de regiones. Dentro de este conjunto de acciones se incluyen también las relacionadas con el mantenimiento del grafo de adyacencias.

- Acciones sobre agregados: determinan la existencia, o no, de uniones entre agregados diferentes y, en caso de producirse, realizan la división en situaciones simples.

- Acciones de supervisión: las acciones de las metareglas, a diferencia de las otras, no modifican los datos de la segmentación en la STM. Especifican qué regla se evalúa en cada momento y, de forma indirecta, activan los diferentes módulos de procesado: inicialización, comienzo de la fase de crecimiento, análisis de regiones o agregados, fin de proceso.

El modelo basado en reglas codifica los criterios que especifican como se va a segmentar una imagen. La información de intensidades, propiedades de regiones, etc., guía el análisis para la creación de una salida acorde con el modelo. Mediante la codificación de estas fuentes de conocimiento en reglas y su aplicación a los datos se obtiene una metodología uniforme y estructurada.

Se podría pensar en la aplicación de varios procedimientos en cascada sobre una imagen, cada uno para una heurística de conocimiento que pudiese mejorar la salida. Sin

embargo, éste no es un método eficiente, puesto que en el curso de mejora de resultados se pueden destruir logros previos, y además, se habría de aplicar un gran número de transformaciones en la imagen, aunque sólo se requiriesen en partes muy concretas. En un sistema basado en reglas se recurre a cada item de información cuando es necesario, de tal forma que una determinada heurística sólo se utilizará para actualizar aquella parte de la imagen sobre la que es aplicable.

Los 5 principios básicos para la agrupación de entidades visuales según las leyes de la Gestalt son: similaridad, proximidad, uniformidad, continuidad y cierre. Cada una de las reglas de la segmentación representa alguno de los principios anteriores, siguiendo además el principio del mínimo compromiso tendente a minimizar el número de acciones drásticas. Nuestro sistema no está orientado a la división (*splitting*); por eso se parte de una sobresegmentación de la imagen original. Las reglas de análisis de regiones conducen la unión de regiones; por tanto, han de especificar los criterios de unión a aplicar a la configuración actual con el fin de producir una partición mejor.

Veamos de forma detallada las reglas asociadas a cada uno de los módulos de procesado comentados.

Reglas de foco de atención El módulo FOCO de ATENCIÓN contiene el grupo de reglas encargado de realizar la selección de datos. Esta selección se realiza en cuatro niveles, como se muestra en la figura 3.26. Inicialmente se crean todas las áreas de soporte de las estructuras de interés. Posteriormente, en el nivel 1 se determinan la siguiente área de soporte a analizar y los núcleos de crecimiento correspondientes. La determinación de todas las áreas de soporte es previo a cualquier otro proceso del sistema. A cada modelo de estructura extraido del corte previo le asignamos un número que se guarda en la característica **NUM_MÁSCARA** de su marco; para reflejar que una región pertenece a un área de soporte dada, lo que hacemos es poner un 1 en la correspondiente posición de la máscara con que se define el valor de la característica **ÁREA_DE_SOPORTE** de su marco. Ésta posee un campo con 8 posiciones; un 1 en la posición $i < 8$ querrá decir que esa región pertenece al área de soporte del i-ésimo objeto a buscar. Más de un 1 (en posiciones que no sean la 8) en la máscara indicará que la región está incluida en más de un área de soporte. Por otra parte, un 1 en esa última posición indicará que la región ya fue analizada en el área de soporte actual. En el nivel 2 de selección se determinan, a requerimiento del sistema de control, la siguiente regiones a procesar dentro del área de

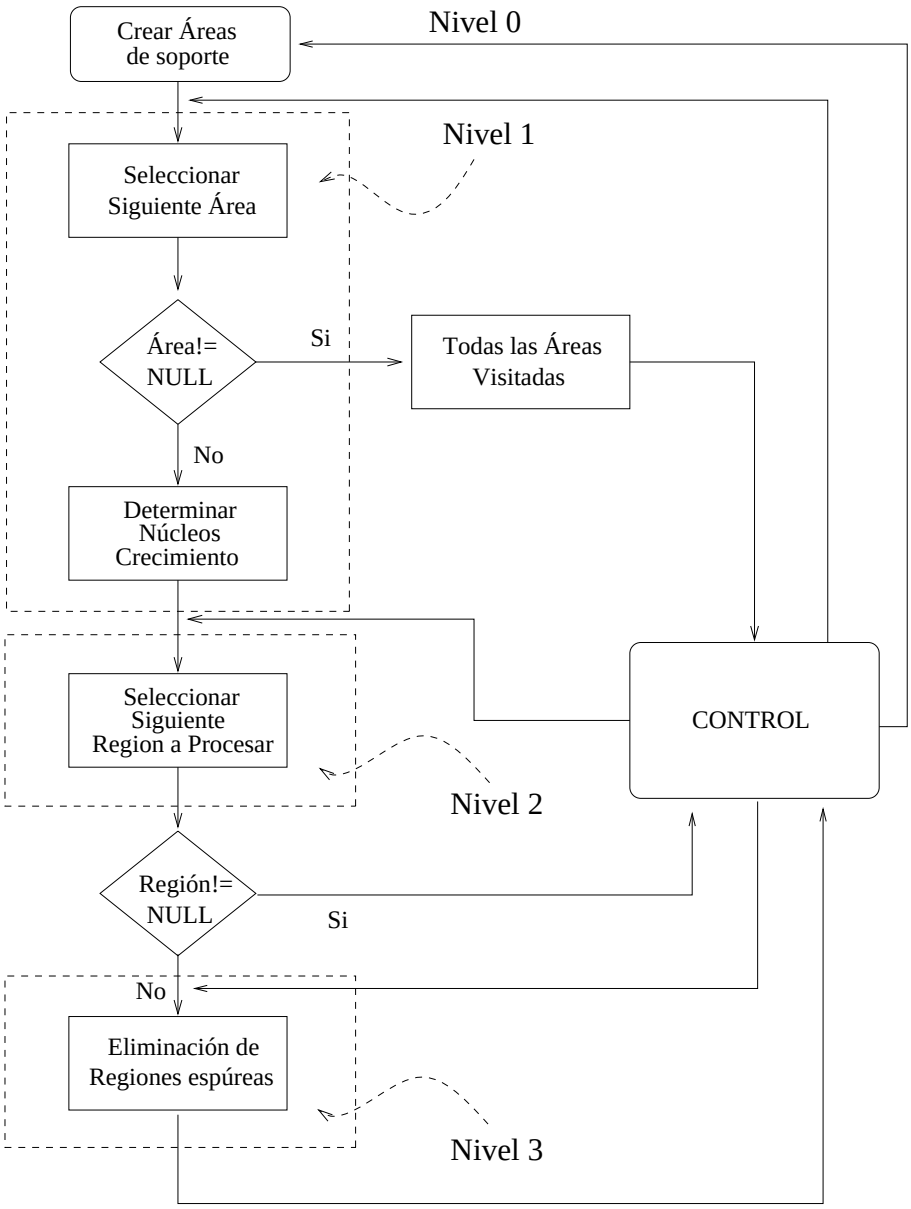


Figura 3.26: Niveles de selección de datos.

soporte que se esté analizando en ese momento. Esta selección se establece en términos de las clases y del tamaño de cada región. En el tercer nivel, de las regiones que no han sido incorporadas a hueso, las que no son continuación de estructuras previas se eliminarán del área de soporte correspondiente. Eliminar de la región de soporte significa poner un cero

en la posición de la máscara de la característica **ÁREA_DE_SOPORTE** de su marco. Además, en el caso de que no pertenezca a ninguna otra región de soporte se unirá al fondo.

El conjunto de reglas de este módulo clasifica inicialmente la escena de entrada en regiones de fondo (*background*) y de primer plano (*foreground).* Por fondo se entiende todas aquellas regiones que no forman parte de ningún área de soporte. Todas estas regiones se unirán en una única región de fondo, que estará conectada a los bordes de la imagen. Será misión del sistema el identificar las regiones que sean realmente hueso dentro del primer plano. Dentro de las áreas de soporte establecidas al comienzo del procesado, las reglas de foco de atención generarán los núcleos de crecimiento y seleccionarán la entidad elemental (región) a procesar en cada momento. Se establece de este modo una jerarquía de regiones dentro de la imagen en cuanto al orden y conveniencia de su procesamiento. El bloque de supervisión será el encargado de mandar la señal para la entrada en funcionamiento del módulo FOCO de ATENCIÓN, indicándole también el nivel de selección a efectuar.

La transición de un área de soporte a otra se produce cuando se da alguna de las siguientes situaciones: no quedan más regiones que procesar en el área de soporte actual, según los criterios de las reglas de FOCO DE ATENCIÓN, o se ha producido la correspondencia de algún núcleo del agregado de regiones con el modelo correspondiente. La iteración en la selección de regiones de un área continúa mientras no se produzca una correspondencia adecuada entre algún núcleo del agregado de regiones con modelo, o no se termine el análisis de regiones del área de soporte actual.

La división de regiones en áreas de interés juega un papel importante en el control de la segmentación, ya que permitirá en unos casos no incorporar regiones de fondo a la máscara de hueso en zonas donde la proximidad entre estructuras diferentes pueda favorecer esta conexión, y en otros casos dará idea de la localización de errores en el proceso de unión. Esto ocurrirá cuando una misma región se incorpore a agregados óseos de áreas de soporte diferentes. Si dichas regiones son pequeñas, se incorporarán al fondo; si son grandes, habrá que someterlas a reanálisis basándose en la información de contornos.

Como ya hemos indicado antes, los modelos de las estructuras se extraen del resultado final del corte previo. Cada uno de estos modelos estará etiquetado como tibia o peroné, y adicionalmente se le asignará un número de máscara. Este número identificará a las áreas de soporte que se creen en el corte actual. Primero se establecen las áreas de soporte de

tibia y peroné asignando valores a la característica de **ÁREA_DE_SOPORTE** de cada región. Después se realiza el análisis en serie de cada una de las áreas de soporte. A continuación se enumeran las reglas empleadas por el módulo de FOCO DE ATENCIÓN.

Determinación de las áreas de soporte:

SI:
(1) CENTROIDE REGIÓN pertenece a ENTORNO MÁSCARA i-ésima corte_previo
ENTONCES:
(1) REGIÓN pertenece a ÁREA de SOPORTE i-ésima

Incorporación a fondo de todas aquellas regiones que no forman parte de ningún área de soporte:

SI:
(1)REGIÓN NO pertenece a ningún ÁREA de SOPORTE
ENTONCES:
(1) CERTEZA de etiqueta de FONDO igual a 1.0

Inicialización de las regiones del área de soporte actual:

SI:
(1) REGIÓN pertenece a ÁREA de SOPORTE actual
ENTONCES:
(1) Marcar REGIÓN como NO ANALIZADA
(2) CERTEZA de etiqueta del HUESO actual a 0.0

Elección de las regiones núcleo de crecimiento iniciales de cada área de soporte:

SI:
(1) DENSIDAD_MEDIA de REGIÓN $\geq$ TH_DENS
(2) REGIÓN pertenece a ÁREA de SOPORTE actual
ENTONCES:
(1) CERTEZA de etiqueta de HUESO actual igual a 1.0
(2) Marcar REGIÓN como ANALIZADA.

Elección de la región a procesar en cada momento:

SI:
(1) REGIÓN pertenece a CLASE inmediatamente ANTERIOR A la ya PROCESADA
(2) ÁREA de REGIÓN MAYOR que las de su CLASE no analizadas
ENTONCES:
(1) DEVOLVER REGIÓN
(2) Marcar REGIÓN como ANALIZADA.

Modificación de la etiqueta de fondo de aquellas regiones que han sido objeto de foco de atención en fase de crecimiento y no incorporadas al área de hueso tras la aplicación de reglas:

SI:
(1) GRADO de CERTEZA de pertenencia de REGIÓN a HUESO BAJO
ENTONCES:
(1) CERTEZA de etiqueta de FONDO igual a 0.5.

Eliminación del área de soporte, una vez finalizado todo el proceso de unión, de aquellos núcleos que no son continuación de ninguna estructura del corte previo:

SI:
(1) PROYECCIÓN de CENTROIDE de NÚCLEO en máscara de HUESO CORRESPONDIENTE en corte previo FALSA
ENTONCES:
(1) ELIMINAR NÚCLEO de ÁREA de SOPORTE

Eliminación del área de soporte, una vez finalizado todo el proceso de unión, de aquellos núcleos con área menor de un umbral:

SI:
(1) ÁREA de NÚCLEO $<$ TH_AGREGADO
ENTONCES:
(1) ELIMINAR NÚCLEO de ÁREA de SOPORTE

Unir a hueso todas aquellas regiones que, una vez finalizado todo el proceso de unión, tienen etiqueta de fondo igual a 0.5 y dentro de un núcleo de hueso

SI:
(1) GRADO de CERTEZA de pertenencia de FONDO $=$ 0.5
(2) REGIÓN INTERIOR a NÚCLEO de HUESO
ENTONCES:
(1) UNIR REGIÓN a NÚCLEO de HUESO
(3) ACTUALIZA GRAFO de ADYACENCIAS

Eliminación del área de soporte, una vez finalizado todo el proceso de unión, de aquellas regiones con etiqueta de hueso menor que 1.0:

SI:
(1) GRADO de CERTEZA de pertenencia de REGIÓN a HUESO actual BAJO
ENTONCES:
(1) ELIMINAR REGIÓN de ÁREA de SOPORTE

Para la determinación de las áreas de soporte lo que se hace es utilizar las máscaras de las estructuras identificadas en el corte previo. Por tanto, con **ENTORNO MÁSCARA**

i-ésima corte_previo hacemos referencia a la estructura modelo (tibia o peroné) con la que estemos trabajando en ese momento. La pertenencia al entorno la determinamos viendo si el centroide de cada región dista como máximo 10 pixels de las máscaras anteriores. El área de soporte de un objeto se puede solapar con la de otro; por consiguiente, podrán existir regiones que pertenezcan a más de un área de soporte. Una vez definida el área de soporte, las regiones que incluye se marcarán como no analizadas. A continuación, con la regla de elección de núcleo de crecimiento se busca la clase de mayor densidad y éste valor se asigna a TH_DENS. En cada iteración del proceso de unión se evalúa una clase de regiones. Finalizada cada una de estas iteraciones, cada región de la clase evaluada tendrá asignado un grado de certeza de pertenencia al agregado óseo como consecuencia de la aplicación de reglas de otros bloques de procesado. Todas aquellas que tengan asignado un valor menor que 1.0 se les asignará una probabilidad de pertenencia a fondo igual a 0.5 para identificarla como región ya procesada y no incorporada a hueso. De esta forma aún no unimos definitivamente estas regiones al fondo, ya que la evaluación de otras regiones podría determinar su consideración final como hueso. Finalmente, analizadas todas las clases, se obtendrán en general agregados de regiones no totalmente conectados. Esto se corresponde generalmente con regiones erróneamente incluidas en el agregado. Una forma de seleccionar el grupo conectado de regiones correcto consiste en eliminar del área de interés aquellos núcleos cuyo centroide no se proyecte en la máscara de la misma estructura del corte previo, o no supere un umbral de área, que tomamos como la mitad del área de la estructura en el corte previo (TH_AGREGADO). Eliminarlos significa modificar la propiedad **ÁREA_DE_SOPORTE** poniendo un 0 en la posición correspondiente a la estructura actual. Si además no existe ninguna otra posición con valor 1, la región se une a fondo. En el FOCO de ATENCIÓN se incluyen también una regla relativa a los huecos dentro de algún núcleo del agregado de regiones, con la que se pretende eliminar huecos de la máscara. Finalmente, todas las regiones que no tengan etiqueta de hueso igual a 1.0 se eliminan del área de soporte.

En todo este conjunto de reglas los factores de contribución (FC) de todas las condiciones son idénticos y los factores de verdad (FV) pueden ser 1 ó 0. Aquí se define el grado de certeza (GC) como 1 si el factor de verificación de la condición (VC) es igual a 1, y cero en otro caso. Las acciones, como siempre, sólo se ejecutarán si GC es 1.0.

En resumen, en este módulo se realizan las operaciones de activación de un área de la imagen, la elección de la clase a evaluar y la región a procesar en cada momento.

Las operaciones relacionadas con la unión de regiones y de actualización de grafo son realizadas por el analizador de regiones. Para ello el módulo de FOCO de ATENCIÓN enviará al sistema de control el requerimiento correspondiente y éste activará los módulos ANALIZADOR de AGREGADOS y ANALIZADOR de REGIONES para la realización de dichas operaciones.

Las acciones anteriores del FOCO de ATENCIÓN corresponden al caso de un corte de la serie que no sea el inicial. En el caso de tratarse del corte inicial, el área de soporte será toda la imagen y el núcleo de crecimiento de ambas estructuras delimita ya los contornos exteriores. Lo demás permanece idéntico.

Reglas de análisis de regiones El módulo ANALIZADOR de REGIONES opera con el conjunto de reglas relativo a la unión de regiones y mantenimiento del RAG. Las reglas del análisis de regiones son reglas orientadas al crecimiento del área de hueso. Para ello se chequean propiedades relacionales y estructurales de las regiones que puedan aconsejar su unión con el área de hueso. Como área de hueso inicial se toman los núcleos de crecimiento generados por las reglas de foco de atención. La unión se producirá siempre y cuando las regiones tiendan a mejorar ciertas propiedades de área de hueso crecida hasta el momento. Tras cada proceso de unión es preciso actualizar el RAG, eliminando los nodos correspondientes y recalculando las propiedades de los nuevos, y obtener los nuevos valores de las características de los agregados óseos. La evaluación de las regiones de cada clase se hará en orden decreciente de tamaño. De esta selección se encarga el módulo de FOCO de ATENCIÓN, como ya hemos dicho. Seleccionada la región a procesar, se ejecutan las reglas indicadas a continuación y en el orden en que se enumeran. La primera regla que se cumpla se ejecuta. Si como resultado de la ejecución de la regla se produce una unión, se pasa a analizar la siguiente región; si no, se le siguen aplicando las siguientes reglas hasta que se llegue a una con grado de certeza 1, o bien, se finalice el conjunto de reglas sin que se verifique ninguna de ellas.

Las reglas del analizador de regiones se aplican sobre las regiones almacenadas en la STM. Cuando ocurre la verificación de la condición de la regla correspondiente, entonces se produce la unión de la región actualmente considerada con otra adyacente, o simplemente, se crea un nuevo núcleo de crecimiento. La función de este conjunto de reglas es pues especificar los criterios de unión a aplicar a la configuración de regiones actual. Al procesar un corte distinguiremos entre dos situaciones: corte inicial y corte no inicial. En el primer

caso no habrá modelo que guíe el análisis y no tendrá sentido definir agregados óseos. Por tanto, el analizador de regiones constará de dos grupos de reglas. Empezaremos por el que corresponde al caso del corte no inicial:

Incorporación al agregado de hueso actual de aquellas regiones que aumenten su compactación:

SI:
(1) CORTE NO INICIAL
(2) La COMPACTACIÓN de la UNIÓN de la REGIÓN y el AGREGADO de HUESO es ALTA respecto de la COMPACTACIÓN del AGREGADO de HUESO
(3) La ELONGACIÓN de la REGIÓN es BAJA
ENTONCES:
(1) UNIÓN a AGREGADO de HUESO
(2) RECÁLCULO de PROPIEDADES de AGREGADO de HUESO
(3) ACTUALIZA GRAFO de ADYACENCIAS

Incorporación al área de hueso de aquellas regiones que actúen de nexo entre dos núcleos si su elongación es baja:

SI:
(1) CORTE NO INICIAL
(2) La ELONGACIÓN de la REGIÓN es BAJA
(3) REGIÓN ADYACENTE a dos NÚCLEOS de HUESO
ENTONCES:
(1) UNIÓN a NÚCLEO de HUESO
(2) RECÁLCULO de PROPIEDADES de AGREGADO de HUESO
(3) ACTUALIZA GRAFO de ADYACENCIAS

Incorporación al agregado de hueso de aquellas regiones que aumentan la proximidad en área, momentos o contorno de la máscara de hueso actual con el modelo:

SI:
(1) CORTE NO INICIAL
(2) AUMENTO de PROXIMIDAD en ÁREA de AGREGADO de HUESO con MODELO
(3) SOLAPE de REGIÓN con MODELO ALTO.
ENTONCES:
(1) UNIÓN a AGREGADO de HUESO
(2) RECÁLCULO de PROPIEDADES de AGREGADO de HUESO
(3) ACTUALIZA GRAFO de ADYACENCIAS

SI:
(1) CORTE NO INICIAL
(2) AUMENTO de PROXIMIDAD en MOMENTOS de AGREGADO de HUESO con MODELO
(3) SOLAPE de REGIÓN con MODELO ALTO.
ENTONCES:
(1) UNIÓN a AGREGADO de HUESO

```
    (2) RECÁLCULO de PROPIEDADES de AGREGADO de HUESO
    (3) ACTUALIZA GRAFO de ADYACENCIAS

SI:
    (1) CORTE NO INICIAL
    (2) AUMENTO de PROXIMIDAD en CONTORNO de AGREGADO de HUESO con MODELO
    (3) SOLAPE de REGIÓN con MODELO ALTO.
ENTONCES:
    (1) UNIÓN a AGREGADO de HUESO
    (2) RECÁLCULO de PROPIEDADES de AGREGADO de HUESO
    (3) ACTUALIZA GRAFO de ADYACENCIAS
```

Incorporación al área de hueso de aquellas regiones que tengan un **GRADO de CERTEZA de pertenencia a AGREGADO de HUESO** alto:

```
SI:
    (1) CORTE NO INICIAL
    (2) CERTEZA de pertenencia a HUESO alto
ENTONCES:
    (1) UNIÓN a AGREGADO de HUESO
    (2) RECÁLCULO de PROPIEDADES de AGREGADO de HUESO
    (3) ACTUALIZA GRAFO de ADYACENCIAS
```

En las reglas anteriores la acción de **UNIÓN a AGREGADO de HUESO** se interpreta de diferentes formas dependiendo de si la región pertenece sólo al área de soporte actual o a alguna otra más. Si sólo pertenece al área de soporte actual, la unión a agregado consiste en unir la región a algún núcleo adyacente, si existe, y eliminar su nodo del RAG, o bien, si es nuevo núcleo de crecimiento, poner a valor 1 la certeza de pertenencia al modelo correspondiente (etiqueta de tibia o peroné en su marco). Si la región pertenece a alguna otra área de soporte, no se eliminará su nodo del RAG. De ser así no se podría evaluar su inclusión en otra estructura. Por consiguiente, si la región se uniese a un núcleo, se actualizarían sus características, pero se mantendría el nodo de la región marcándolo como analizado para el área de soporte actual. Si la región pasase a ser un nuevo núcleo para el agregado actual, se duplicaría en el RAG. Al nodo original se le asignará la etiqueta del hueso que es continuación y se marcará como perteneciente únicamente al agregado actual, y al nuevo se le marcará como analizado para el área de soporte actual dejando patente la pertenencia a todas las áreas de soporte de las que inicialmente formaba parte. De este modo, si una región se incluye como parte de más de un hueso aparecerá en el RAG final un nodo libre que dará cuenta de esta situación.

Para la incorporación de una región al agregado óseo por aumento de su compactación lo que se hace es calcular la medida de compactación que tendría el agregado resultante

de la unión respecto de la medida actual. Para ello es preciso calcular de nuevo área y perímetro del agregado. Las dos medidas de compactación se hacen considerando los posibles huecos del agregado como parte de éste, y también se considerará, a efectos de cálculo, que las regiones interiores a la unión formarán parte del agregado final. Esto se incluye para considerar los casos donde el hueso cortical es muy estrecho y la trabécula poco densa, y en los que la incorporación de la región daría lugar a un agregado cerrado. La medida de elongación se introduce para no incluir regiones de transición entre hueso y músculo como parte del primero. La no consideración de esta medida tendería a dilatar la máscara de hueso y favorecería la inclusión de regiones externas al mismo al aumentar el perímetro compartido entre éstas y el agregado óseo. Este aumento del perímetro compartido aumentaría a su vez la medida *tentativa* de compactación, o podría poner en contacto a esa región con el hueso y favorecer su incorporación al agregado por unión de dos agregados óseos.

El que una región disminuya la compactación del agregado óseo al que es adyacente no implica necesariamente una disminución en el grado de certeza en la pertenencia a la clase de hueso, ya que se puede tratar de una región que aunque presente forma irregular, sea parte del hueso. Sin embargo, el hecho de aumentar la compactación, aunque menos del umbral mínimo exigido, si se puede ver como un aumento en la confianza de pertenencia a la clase hueso, y por ello, en consecuencia, ese grado de certeza se ha de aumentar.

En la primera regla se evalúa el aumento del grado de compactación con la inclusión de la región al agregado óseo. Por otra parte, se exige un máximo de elongación para no incluir zonas de transición externas al hueso. El aumento del grado de compactación se considerará alto con factor de verdad 1, cuando es mayor o igual que el 25%. La elongación se considerará baja, con factor de verdad 1, para valores menores que 0.4. Aquí los factores de verdad (FV) estarán comprendidos entre 0 y 1, según se indica en la figura 3.27. En esta regla la condición de la tiene mayor peso que la de la elongación; por ello, los factores de contribución son distintos: 0.6 para compactación y 0.4 para elongación. La acción se ejecuta para valores de VC mayores que $V_{TH} = 0.75$. En el caso de que el grado de certeza ($GC = \frac{VC}{0.75}$) de esta regla sea menor que 1, no se ejecutará la acción correspondiente (unión), si no que se modificará la certeza (C) de la idoneidad de la acción de la forma que comentamos al describir la estructura de las reglas. A este valor le llamamos **GRADO de CERTEZA de pertenencia a HUESO** y es el que almacenamos en el marco asociado a la región dentro de la propiedad (**ETIQUETA (TIBIA(VALOR(-**

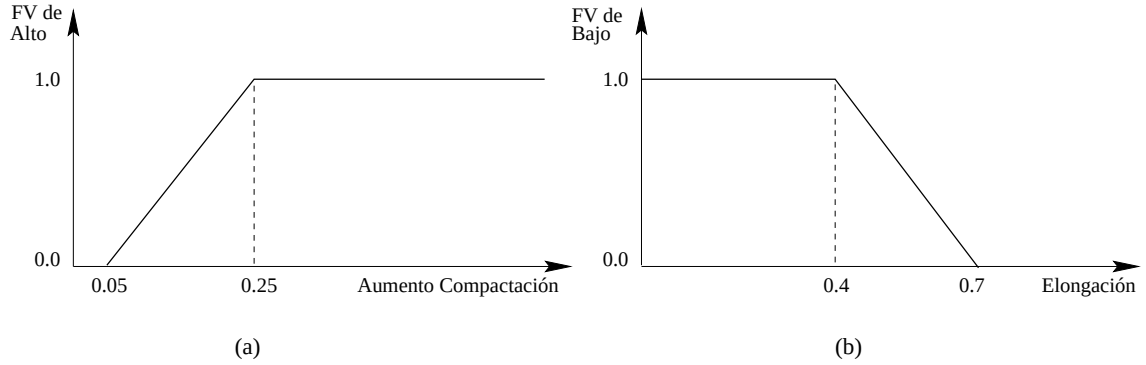


Figura 3.27: Factores de verdad para: (a) aumento de compactación alta y, (b) elongación baja.

)) **PERONÉ**(**VALOR**(-)) **FONDO**(**VALOR**(-)))) . El agregado al que hagamos referencia será tibia o peroné, y esa característica será la que modificaremos.

La regla que considera la adyacencia a dos o más agregados óseos se incluye para favorecer el cierre de estructuras mediante regiones que no cumplan la condición de aumento de compactación. Esta regla es muy significativa en casos de fuerte pérdida ósea y donde el hueso cortical es muy estrecho, de tal forma que la presegmentación puede etiquetarla como una clase más alejada de la clase de hueso cortical de la que le correspondería. Aquí se aplica de nuevo la restricción de la elongación por los motivos antes expuestos. Las condiciones de esta segunda regla para la unión es que la región una dos agregados de hueso ($FV \in \{0, 1\}$) y su elongación sea baja ($FV \in [0, 1]$). La segunda condición es importante para evitar la unión entre estructuras diferentes muy próximas. En este caso los factores de contribución son iguales a 0.5. El GC se define como VC cuando $VC > V_{TH} = 0.75$ y cero en otro caso. De esta forma si no se da la condición de unión, GC será cero y la regla no reforzará ni debilitará la consideración de la región como hueso.

La tres reglas siguientes, relativas al aumento de similaridad con el agregado del corte previo, presentan un elemento nuevo con respecto a las dos anteriores, que es la conveniencia de su aplicación. Nuestro sistema trabaja con modelos aproximados que vienen representados por los objetos reconocidos en los cortes ya procesados; por tanto. no serán aplicables en el procesado del corte inicial (conveniencia de aplicación 0). Por ello se

añade la condición de que el corte que se esté analizando no sea el inicial.

La medida de aumento de proximidad de área se introduce para incluir todas las regiones que corresponden a hueso cortical muy estrecho y poco denso conectado a hueso trabecular de baja densidad. En estos casos podría ocurrir que ninguna de las demás reglas decidiese unirlas al hueso cuando en realidad son parte de él. Por tanto, a la condición de aumento de proximidad en área (AA) hay que añadirle la de solape con la estructura correspondiente (tibia o peroné) del corte previo (IS); en otro caso se podrían incluir regiones externas. Esta unidad de condición puede aumentar o disminuir el grado de certeza de la pertenencia a la clase hueso. Aquí de nuevo las medidas son relativas y lo que se hace es considerar la distancia en área con la máscara del corte previo antes (D_{A_1}) y después (D_{A_2}) de considerar la región actual. La medida de distancia en área ya fue definida en el apartado 3.4.4. A partir de las medidas de distancia en área, definimos la medida de acercamiento en área (AA) del siguiente modo:

$$AA_i = 1.0 - \frac{D_{A_2}}{D_{A_1} + D_{A_2}} \tag{3.63}$$

También se determina en qué medida solapa esta región con la máscara del corte previo:

$$IS_i = \frac{SOLAPE_i}{A_i} \tag{3.64}$$

donde $SOLAPE_i$ representa el área del resultado de aplicar el operador lógico AND entre la región i-ésima y la máscara del modelo correspondiente. A_i representa el área de la región. El factor de contribución es para los dos predicados igual (0.5) y ha de darse que $AA > 0.5$ para que se produzca un acercamiento en área con el modelo e $IS > 0.5$ para no incluir como parte del hueso regiones externas al mismo. Los factores de verdad (FV) se indican en la figura 3.28. El grado de certeza se define del siguiente modo:

$$GC = \begin{cases} 1 & \text{si } VC \geq 0.97 \\ \min\{2(IS - 0.5), 2(AA - 0.5)\} & \text{otro caso} \end{cases} \tag{3.65}$$

Como se deduce de la expresión anterior, ahora el grado de certeza puede tomar valores positivos y negativos. De esta forma, si una región aumenta la distancia en área se tiende a rechazarla como parte del hueso bajo análisis.

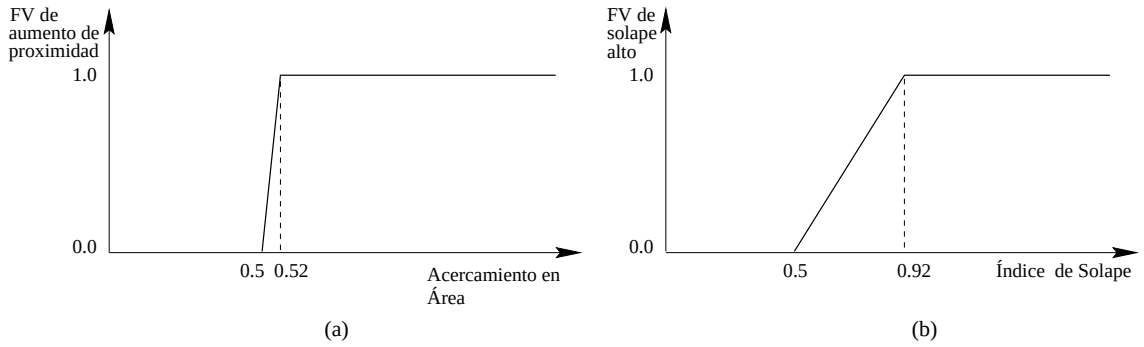


Figura 3.28: Factores de verdad para: (a) acercamiento en área y, (b) índice se solape.

La siguiente medida relativa a la similaridad es la de aumento de acercamiento en momentos, que se obtiene mediante el cálculo de distancias a los segundos momentos del agregado previo. La distancia en momentos también había sido definida en el apartado 3.4.4. Los diversos factores toman los mismos valores que para el caso anterior, considerando que ahora el aumento de semejanza de forma se corresponde con el aumento de acercamiento en momentos. Aquí también se incluye la condición de solape, y también pueda variar el grado de certeza en los dos sentidos.

La última medida de similaridad considerada es la aproximación del contorno del corte actual al del corte previo. Ésta sólo se puede utilizar para aumentar el grado de certeza, pero no dice nada a cerca de su disminución, puesto que el corte actual puede estar por esa parte expandido respecto del corte previo. Entonces seguimos con la filosofía apuntada de no tomar decisiones demasiado precipitadas. A parte de buscar la proximidad del contorno, hay que ver si la región solapa en su mayor parte con la del corte previo; en otro caso, podríamos incluir regiones externas al hueso cuando éste se contrae en esa parte con respecto al corte previo. La aproximación al contorno del corte previo se mide determinando el tanto por uno de puntos del contorno del corte previo que tienen próximo, dentro de un determinado radio, un punto de contorno en el agregado de áreas actual. Esta medida se evalúa para el estado actual del agregado (match_1) y considerando que la región actual forma parte del agregado (match_2). La medida de aproximación será:

$$AC_i = 1.0 - \frac{1.0 - match_2}{2.0 - match_2 - match_1} \tag{3.66}$$

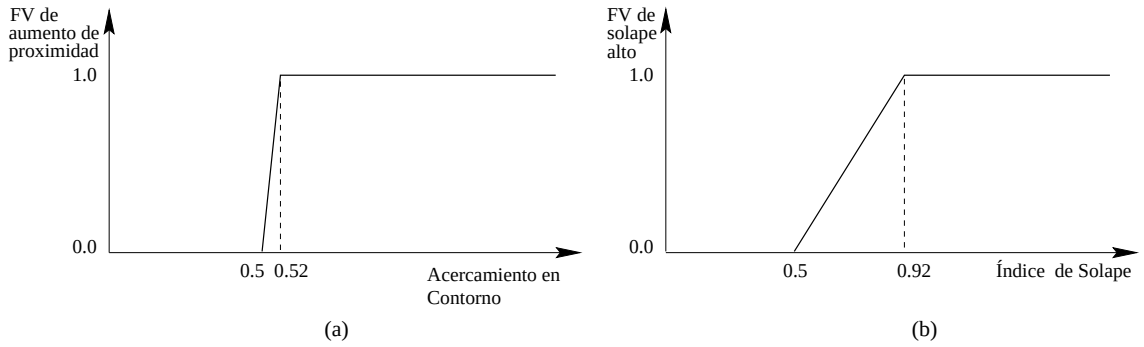


Figura 3.29: Factores de verdad para: (a) acercamiento en contorno y, (b) índice se solape.

Si esta medida es mayor que 0.5 es que la aproximación aumenta y la condición se cumple. También se determina en que medida solapa esta región con la máscara del corte previo. El factor de contribución es para los dos predicados igual (0.5) y los factores de verdad son los indicados en la figura 3.29. Finalmente, el grado de certeza se define como:

$$GC = \begin{cases} 1 & \text{si } VC \geq 0.90 \\ \min\{0, 2(IS - 0.5), 2(AC - 0.5)\} & \text{otro caso} \end{cases} \tag{3.67}$$

De esta forma, el valor de certeza (C) no puede disminuir como resultado de aplicación de esta regla.

Cada vez que la evaluación de una regla origina un grado de certeza (GC) menor que 1, se actualiza la certeza de la parte de acción (C). Esta certeza la habíamos llamado grado certeza de pertenencia a hueso. Después de haber chequeado todas las reglas anteriores, se ejecuta una nueva regla que chequea este valor, de tal forma que si éste es alto (> 0.75), se une dicha región al agregado óseo.

Cuando se trate del corte inicial no se dispondrá de ningún modelo y el conjunto de reglas se simplifica. En regiones alejadas de la parte proximal de la tibia, los contornos de tibia y peroné van a quedar perfectamente delimitados por el contorno exterior de las regiones de la clase más densa. De cualquier modo, incluimos dos reglas para el caso de que forme parte del hueso cortical alguna región que no sea de la clase más alta. Además, la unión de una región a hueso se hará únicamente si es adyacente a un núcleo. En esta situación no están definidos los agregados. Se impone la condición del perímetro para

evitar unir regiones alargadas que definen la separación entre hueso cortical y músculo. De esta forma, al no incluir estas regiones, tampoco se incluirá ninguna otra de músculo, ya que las uniones se hacen a núcleos adyacentes.

SI:
(1) CORTE INICIAL
(2) La COMPACTACIÓN de la UNIÓN de la REGIÓN a un NÚCLEO aumenta la COMPACTACIÓN
(2) El PERÍMETRO de la REGIÓN es menor que 2 veces su ÁREA
ENTONCES:
(1) UNIÓN a NÚCLEO
(3) ACTUALIZA GRAFO de ADYACENCIAS

Incorporación al hueso de aquellas regiones que actúen de nexo entre dos núcleos

SI:
(1) CORTE INICIAL
(2) (2) El PERÍMETRO de la REGIÓN es menor que 2 veces su ÁREA
(2) REGIÓN ADYACENTE a dos NÚCLEOS de HUESO
ENTONCES:
(1) UNIÓN a NÚCLEO de HUESO
(3) ACTUALIZA GRAFO de ADYACENCIAS

Para el caso de cortes diferentes al inicial, cada núcleo tiene asignada la etiqueta de tibia y/o de peroné según la región de soporte a la que pertenezca, de tal forma que esta clasificación se va pasando de corte a corte. En el proceso de evaluación se decidirá cuáles de los núcleos se quedan con las etiquetas de tibia y peroné en función de la mayor proximidad con el modelo. De esta forma se va pasando la información de un corte al siguiente.

En el caso del corte inicial, no existen agregados óseos y por lo tanto no hay clasificación inicial de los núcleos de crecimiento. Por tanto, una vez finalizado el proceso de unión se tiene que identificar tibia y peroné. La identificación es muy simple: de las dos regiones de mayor área, la tibia será la más próxima al centro de la imagen, y el peroné será la otra.

Reglas de análisis de agregados El módulo ANALIZADOR DE AGREGADOS ejecuta tareas relacionadas con la evaluación de la segmentación una vez finalizados los procesos de unión. Así, una vez terminado el proceso de análisis de regiones dentro de cada subárea determinada por las reglas de foco de atención, el sistema de control ha de evaluar los resultados, y en base a ellos determinar las siguientes reglas a evaluar. Uno

de los posibles procesos a ejecutar a continuación tiene que ver con el análisis de los agregados de regiones tras el proceso de crecimiento de los núcleos óseos. El ANALIZADOR de AGREGADOS es el encargado de analizar posibles uniones entre agregados de hueso diferentes. Los posibles agregados diferentes pueden ser tibia y peroné, o diferentes partes de la tibia, que se ha podido bifurcar por la aparición de lesiones. Este módulo será el encargado de determinar si se han conectado, o no, estructuras diferentes. Si se ha producido unión, las separará en ciertos casos simples. Finalmente, informará, al sistema de control de si el número de regiones se corresponde con el real o no. Esta información, junto con las medidas de similaridad, será utilizada por el sistema de control para decidir sobre la necesidad de realizar procesos de división.

Las dos características clave que debe reunir cualquier segmentación de una escena son las siguientes:

$$\begin{aligned} I &= \cup_{i=1}^{N} R_i \\ R_i \cap R_j &= \emptyset, \quad si \quad i \neq j \end{aligned} \tag{3.68}$$

La primera de las condiciones se va a verificar siempre. Sin embargo, cabe la posibilidad de que la segunda no; es decir, puede ocurrir que una región se incorpore a dos agregados diferentes. Esta situación puede darse cuando dos estructuras óseas estén muy próximas y que la región, o las regiones que las separan, cumplan la parte de condición de las reglas para los dos agregados. Cuando se de esta situación, quedarán formando parte del RAG las regiones que se incorporaron a más de una estructura. En este caso, las parte de etiqueta y de área de soporte del marco correspondiente a una de esas regiones sería del siguiente tipo:

(REGIÓN
(IDENTIFICADOR (VALOR 1234))
(ÁREA_DE_SOPORTE (VALOR 00000011))
(ETIQUETA(TIBIA(VALOR(1.0)) PERONÉ(VALOR(1.0)) FONDO(
VALOR(0))))
...

Otra posibilidad es que la presegmentación no consiguiese separar los dos objetos buscados, de tal forma que uniese en una única región dos regiones de objetos reales diferentes, y/o parte del fondo que las separa.

En el caso de que la región de unión sea de tamaño pequeño en comparación con la de los agregados (menor del 10% del área menor), se puede incorporar, sin más, al fondo, y de esta forma ya habremos conseguido eliminar el problema de clases no disjuntas. En el caso de tratar con regiones de tamaño grande, es muy probable que no corresponda principalmente al fondo, sino que incluya parte de éste y parte de alguno de los objetos de interés. En este caso, es más adecuado recurrir a procesos de división de regiones basados en técnicas de análisis de bordes.

Las reglas asociadas con este análisis son las que se describen a continuación:

Recorrido del RAG para incorporar al fondo las regiones pertenecientes a más de una clase:

SI:
(1) REGIÓN PERTENECE a más de un AGREGADO
(2) ÁREA de REGIÓN PEQUEÑA
ENTONCES:
(1) ELIMINAR REGIÓN de los AGREGADOS
(2) ACTUALIZAR GRAFO de ADYACENCIAS

Recorrido del RAG para separar objetos pegados:

SI:
(1) NÚCLEO ADYACENTE a NÚCLEO de DISTINTO AGREGADO
(2) CORRESPONDENCIA de NÚCLEOS con MODELO (μ) ALTA
(3) NO EXISTE REGIÓN PUENTE
ENTONCES:
(1) EROSIÓN de las DOS REGIONES por separado

Recorrido del RAG para determinar si quedan uniones o no:

SI:
(1) NÚCLEO ADYACENTE a NÚCLEO de DISTINTO AGREGADO
ENTONCES:
(1) Devolver SEÑAL de disparar información de BORDES

SI:
(1) REGIÓN perteneciente a NÚCLEOS de DISTINTO AGREGADO
ENTONCES:
(2) Devolver SEÑAL de disparar información de BORDES

Reglas de análisis de bordes El módulo ANALIZADOR de BORDES entra en funcionamiento sólo en el que caso de que, tras finalizar el proceso de análisis de agregados, se detecte la unión entre diferentes estructuras. Ante esta situación, el sistema de control echará mano de información de bordes que permita eliminar estos problemas de infrasegmentación.

En el caso de que se requiera la intervención del módulo ANALIZADOR de BORDES, es preciso ejecutar el procesado para la obtención de información de bordes ya comentado en el bloque de bajo nivel.

El resultado de la extracción de contornos preliminar es un conjunto de bordes que hay que unir para obtener contornos cerrados, o con el mayor grado de conexión posible, que definan las estructuras de interés. Por tanto, partiremos de un mapa de contornos que intentaremos unir en base a la proximidad de sus extremos y por prolongación de éstos.

Se comienza por etiquetar los contornos, según su proximidad a alguno de los presentes en el corte previo: para cada uno de sus pixels se busca en un radio ρ_b, y cuando se encuentra un contorno del modelo, se incrementa el contador de proximidad a ese contorno en 1. En caso de no encontrar ningún contorno en ese radio, se incrementa el contador de pertenencia a fondo. Al final se comparan los contadores de proximidad a los dos contornos más próximos y si no hay una relación suficientemente grande ($\sim$ 4:1), el contorno se etiqueta en principio como perteneciente al fondo; en otro caso recibirá la etiqueta del corte más próximo. Si el contorno se etiquetó como fondo, aún puede recibir otra etiqueta siempre que los vecinos más próximos a sus extremos hayan recibido la misma etiqueta, y ésta sea distinta de la del fondo. En otro caso, el contorno se eliminará.

El siguiente proceso consiste en unir los extremos de los contornos. Para ello se buscarán los contornos más próximos a sus extremos de entre los pertenecientes a la misma clase, y se unen por prolongación de sus extremos.

El último paso consiste en calcular la máscara formada por todas las regiones internas a los contornos obtenidos, y a partir de esa máscara obtener el contorno externo. A esta máscara le añadimos las regiones encerradas por los contornos obtenidos tras la etapa de preprocesado, y ya estamos en condiciones de integrar la información de contornos con la del crecimiento de regiones.

Las reglas del análisis de bordes para el cierre de los mismos son las siguientes:

Reglas de selección de cierre de contornos:

SI:
(1) GRADO de PROXIMIDAD a uno de los CONTORNOS del MODELO ALTO
ENTONCES:
(1) ETIQUETAR BORDE según MODELO PRÓXIMO

SI:
(1) GRADO de PROXIMIDAD a uno de los CONTORNOS del MODELO BAJO
ENTONCES:
(1) ETIQUETAR BORDE como FONDO

SI:
(1) BORDE ETIQUETADO como FONDO
(2) VECINOS MÁS PRÓXIMOS con MISMA ETIQUETA
(2) MISMA ETIQUETA DISTINTA DE FONDO
ENTONCES:
(1) ETIQUETAR BORDE con MISMA ETIQUETA

SI:
(1) BORDE ETIQUETADO como FONDO
(2) VECINOS MÁS PRÓXIMOS con DISTINTA ETIQUETA
ENTONCES:
(1) ELIMINAR BORDE

SI:
(1) BORDE ETIQUETADO como FONDO
(2) VECINOS MÁS PRÓXIMOS con ETIQUETA DE FONDO
ENTONCES:
(1) ELIMINAR BORDE

Regla de cierre:

SI:
(1) EXISTE CONTORNO MÁS PRÓXIMO de la MISMA CLASE
ENTONCES:
(1) UNIR con CONTORNO MÁS PRÓXIMO por VÉRTICE MÁS PRÓXIMO

Una vez que se han conectado el mayor número posible de bordes, se inicia el proceso de división. Como ya hemos indicado, este módulo sólo se ejecutará si falla el reconocimiento de regiones tras el proceso de unión. Lo que se pretende es dividir las regiones resultantes de la unión por aquellas partes donde haya evidencia de borde. Existe una fase de pre-procesado consistente en clasificar las regiones encerradas por los contornos obtenidos en

el módulo de bordes como continuación de alguna de las estructuras del corte previo. El análisis de bordes se efectuará en una zona localizada de la imagen indicada por el sistema de control. A este mapa de regiones se le añade la máscara obtenida en esa zona por crecimiento de regiones: para cada pixel de la máscara del crecimiento de regiones se mira si solapa con alguna región interior a los contornos; si es así, se le asigna la misma etiqueta, y en caso contrario se le asigna una nueva. Todos los puntos que no solapan llevarán la misma etiqueta. Si una de estas regiones une dos o más regiones con etiqueta diferente, entonces se elimina. Con ello se produce la separación entre las dos estructuras diferenciadas en el corte previo. En caso contrario, la región se une a la clase adyacente, siempre y cuando comparta con ella una parte mínima de su contorno.

Reglas de división:

SI:
(1) REGIÓN perteneciente a AGREGADO
(2) REGIÓN NO SOLAPANTE con MÁSCARA DE BORDES
(3) REGIÓN ADYACENTE a DOS o más CLASES
ENTONCES:
(1) ELIMINACIÓN de REGIÓN

SI:
(1) REGIÓN perteneciente a AGREGADO
(2) REGIÓN ADYACENTE a REGIONES de UNA única CLASE
(3) COMPARTICIÓN de PERÍMETRO ALTA
ENTONCES:
(1) UNIÓN de REGIÓN con CLASE

SI:
(1) REGIÓN perteneciente a AGREGADO
(2) REGIÓN ADYACENTE a REGIONES de UNA única CLASE
(3) COMPARTICIÓN de PERÍMETRO BAJA
ENTONCES:
(1) ELIMINACIÓN de REGIÓN

En todas estas reglas los factores de contribución son idénticos para todas las premisas. Se considera que el grado de compartición del perímetro es alto ($FV = 1$) si supera el 35% del perímetro total de la región considerada; en otro caso, el factor de verdad será 0. El grado de certeza será 1 si el factor de verificación de la condición es 1 y cero en otro caso. Con este proceso de división se tiene un camino de realimentación en el proceso de crecimiento de regiones. De todas formas, cuando la información de gradientes es incompleta, la división de regiones puede no resultar totalmente correcta. Por esta razón

proponemos la integración de la información del sistema de crecimiento de regiones con la resultante de la segmentación mediante modelos deformables. En secciones posteriores describiremos tanto el módulo basado en modelos deformables como el de integración.

Metareglas El módulo SUPERVISOR es el que realiza todas las tareas de control que llevamos comentando hasta el momento. Este módulo hace uso de las metareglas para coordinar la actividad de los demás módulos de procesado. Las condiciones de estas reglas se refieren a la historia del procesado y no a los datos de la segmentación. Sus acciones establecen el flujo en el proceso de control, especificando el siguiente proceso a activar.

Metaregla para la inicialización de regiones:

SI:
(1) LISTA de REGIONES sin INICIALIZAR
ENTONCES:
(1) INICIALIZAR REGIONES

Metaregla para la generación de áreas de foco de atención:

SI:
(1) ÁREAS de SOPORTE sin INICIALIZAR
(2) LISTA de REGIONES INICIALIZADA
ENTONCES:
(1) Llamada al módulo de FOCO de ATENCIÓN(nivel 0)

Metaregla para la determinación de la siguiente área de soporte a analizar y de sus núcleos de crecimiento:

SI:
(1) LISTA de ÁREAS de SOPORTE CREADA
(2) LISTA de ÁREAS de SOPORTE NO VACÍA
ENTONCES:
(1) Llamada al módulo de FOCO de ATENCIÓN(nivel 1)

Metaregla para la selección de región a analizar:

SI:
(2) LISTA de REGIONES en área de soporte NO VACÍA
ENTONCES:
(1) Llamada al módulo de FOCO de ATENCIÓN(nivel 2)

Metaregla para la llamada al módulo de análisis de regiones:

SI:
(1) El PROCESO anterior fue FOCO de ATENCIÓN
(2) CORTE INICIAL
ENTONCES:
(1) Llamada al módulo de ANÁLISIS de REGIONES(sin modelo)

SI:
(1) El PROCESO anterior fue FOCO de ATENCIÓN
(2) CORTE NO INICIAL
ENTONCES:
(1) Llamada al módulo de análisis de REGIONES(con modelo)

Metaregla para la llamada al módulo de foco de atención:

SI:
(1) El PROCESO anterior fue análisis de REGIONES
(2) LISTA de REGIONES NO VACÍA
ENTONCES:
(1) Llamada al módulo de FOCO de ATENCIÓN(nivel 2)

Metareglas para la parada del crecimiento en el área de soporte actual:

SI:
(1) El PROCESO anterior fue análisis de REGIONES
(2) LISTA de REGIONES VACÍA
ENTONCES:
(1) PARADA ANÁLISIS DE REGIONES

SI:
(1) El PROCESO anterior fue análisis de REGIONES
(2) CORRESPONDENCIA de NÚCLEO con MODELO
ENTONCES:
(1) PARADA ANÁLISIS DE REGIONES

SI:
(1) PARADA PROCESADO 2D DEL CORTE ACTUAL ALCANZADA
ENTONCES:
(1) ACTIVAR FOCO de ATENCIÓN(nivel 3)

SI:
(1) FINALIZADO FOCO de ATENCIÓN(nivel 3)
ENTONCES:
(1) ACTIVAR módulo de ANÁLISIS DE AGREGADOS

Metaregla para suavización:

SI:
(1) PARADA de ANÁLISIS de AGREGADOS ALCANZADA
ENTONCES:
(1) REALIZAR CIERRE MORFOLÓGICO

Metaregla para división:

SI:
(1) CIERRE MORFOLÓGICO REALIZADO
(2) SEÑAL de incorporación de información de BORDES activada
ENTONCES:
(1) ACTIVAR ANALIZADOR de BORDES

Metareglas para medidas de similaridad:

SI:
(1) PARADA de ANÁLISIS de AGREGADOS ALCANZADA
(2) CORTE NO INICIAL
ENTONCES:
(1) ACTIVAR Medidas de SIMILARIDAD

SI:
(1) Medidas de SIMILARIDAD ALTAS
ENTONCES:
(1) IDENTIFICACIÓN FINAL de OBJETOS
(2) PARADA PROCESADO 2D DEL CORTE ACTUAL

SI:
(1) Medidas de SIMILARIDAD BAJAS
(2) OBJETOS ENCONTRADOS igual en número a OBJETOS BUSCADOS
ENTONCES:
(1) PARADA PROCESADO 2D DEL CORTE ACTUAL

SI:
(1) Medidas de SIMILARIDAD BAJAS
(2) SOLAPE de estructuras
ENTONCES:
(1) Llamada a ANALIZADOR DE BORDES(ventana)
(2) PARADA PROCESADO 2D DEL CORTE ACTUAL

Posibilidad de existencia de bifurcación:

SI:
(1) Medidas de SIMILARIDAD BAJAS
(2) OBJETOS ENCONTRADOS MAYORES en número a OBJETOS BUSCADOS
ENTONCES:
(1) Realizar Medidas de SIMILARIDAD para BIFURCACIÓN
(2) PARADA PROCESADO 2D DEL CORTE ACTUAL

Comienzo de análisis de un nuevo corte:

SI:
(1) PARADA PROCESADO 2D DEL CORTE ACTUAL ALCANZADA
(2) CORTES SIN PROCESAR
ENTONCES:
(1) ACTIVAR FOCO de ATENCIÓN(nivel 1)

Metaregla para parada del sistema:

SI:
(1) PARADA PROCESADO 2D DEL CORTE ACTUAL ALCANZADA
(2) TODOS LOS CORTES PROCESADOS
ENTONCES:
(1) PARADA DE SISTEMA

El módulo SUPERVISOR es el encargado de guiar todo el proceso ordenando la entrada en funcionamiento de los diferentes módulos y de dar una etiqueta definitiva a las diferentes regiones. Entre sus funciones está también la determinación de la zona localizada de la imagen sobre la que se introduce información de bordes. Si la llamada se hace cuando se detecta solape entre regiones (caso más habitual), esta zona se define a partir del corte previo y será el mínimo rectángulo que incluya la zona de la imagen incorporada a hueso en el corte actual pero no en el corte previo, y que además conecte dos estructuras en el corte previo. Cuando la llamada se hace al detectar más regiones de las esperadas (posible bifurcación) la ventana se define como el mínimo rectángulo que incluye todas estas regiones

Otra función a realizar por este módulo es el cálculo de las medidas de similaridad. Cuando se dispone de un modelo para tibia y otro para peroné se efectúan las medidas de similaridad teniendo en cuenta el modelo. Cuando se dispone sólo de modelo de tibia, porque el peroné ya desapareció y haya más de un núcleo, y no se alcance correspondencia con el modelo, se deducirá la posibilidad de bifurcación. Por tanto, se aplican las medias de similaridad relativas al punto de bifurcación. Finalizado el procesado de cada corte, y según el valor de la medida de similaridad, se hará la integración con la segmentación resultante de la aplicación de los modelos deformables de forma diferente. Si la similaridad es alta (> 0.95), se realiza ajuste de contornos; en otro caso, se hará la integración de la forma que se indicará en la siguiente sección. Finalizada la integración, se decidirá que ha finalizado el peroné cuando se pasa de un corte con las dos estructuras a otro que sólo presenta tibia. También se confirmará la existencia de la bifurcación en tibia cuando se pase de la existencia de un modelo de tibia a 2 ó más.

3.4.5 Resultados

Una muestra de los resultados proporcionados por este sistema de crecimiento es la que se ilustra en las figuras 3.23 y 3.30. En la figura 3.23 se mostraba la imagen original,

la presegmentación, el modelo y los núcleos iniciales para el crecimiento de tibia. En la figura 3.30 se muestra el estado de la segmentación en diferentes etapas, considerando tanto tibia como peroné. El crecimiento de regiones parte del mapa de regiones proporcionado por la presegmentación inicial (figura 3.23.b) y del modelo proporcionado por el resultado de la segmentación del corte previo (figura 3.23.c). Se determinan los núcleos para el crecimiento de tibia (regiones a nivel 255 en la figura 3.23.d) y se procede con el crecimiento de regiones (figuras 3.30.e-g). La figura 3.30.h muestra el resultado final de la segmentación superpuesto a la imagen original; aquí se observa como el sistema identifica las dos estructuras de interés (tibia y peroné).

En general, el sistema no presenta problemas para separar tibia y peroné siempre y cuando exista entre las dos una información completa relativa a gradientes fuertes o regiones con contraste suficiente. Otro ejemplo más complejo resuelto con éxito por el sistema es el representado en la figura 3.31.a. En otras circunstancias puede ocurrir que el proceso de unión conecte las dos zonas óseas y la información de gradiente, por incompleta, no permita una separación precisa entre las dos estructuras. Este tipo de problemas surgen, principalmente, cuando la distancia entre cortes es muy grande en una zona con gran variación de forma debida sobre todo a lesiones o patologías. En estos casos el proceso de interpolación de imágenes entre cortes no suele ser lo suficientemente buena como apoyar de forma conveniente el proceso de segmentación. Un ejemplo de este tipo es el que se muestra en la figura 3.31.b. Este ejemplo corresponde a una secuencia de corte de CT con una distancia entre cortes excesivamente grande como para poder hacer una buena reconstrucción de la estructura 3D. Debido a esta enorme separación, el proceso de generación de imágenes intermedias mediante interpolación baja enormemente su rendimiento y la segmentación resultante dista mucho de la real.

En la figura 3.32.a se muestra el resultado de la segmentación mediante crecimiento de regiones de la imagen cuya presegmentación se mostró en la figura 3.14. En esta figura se observa como los dos objetos de interés (tibia y peroné) son identificados, pero los contornos no están localizados con precisión. Este falta de precisión es debida a que la imagen ha sido sometida a un proceso de realce de tal forma que el contraste entre hueso y músculo se ha incrementado notablemente. Como la presegmentación se basa en características del entorno del pixel, se han incluido de partida en la misma región pixels que no corresponden a hueso, por lo que el sistema de crecimiento de regiones expande el contorno más allá de sus límites reales. La mejora de la localización de estos contornos

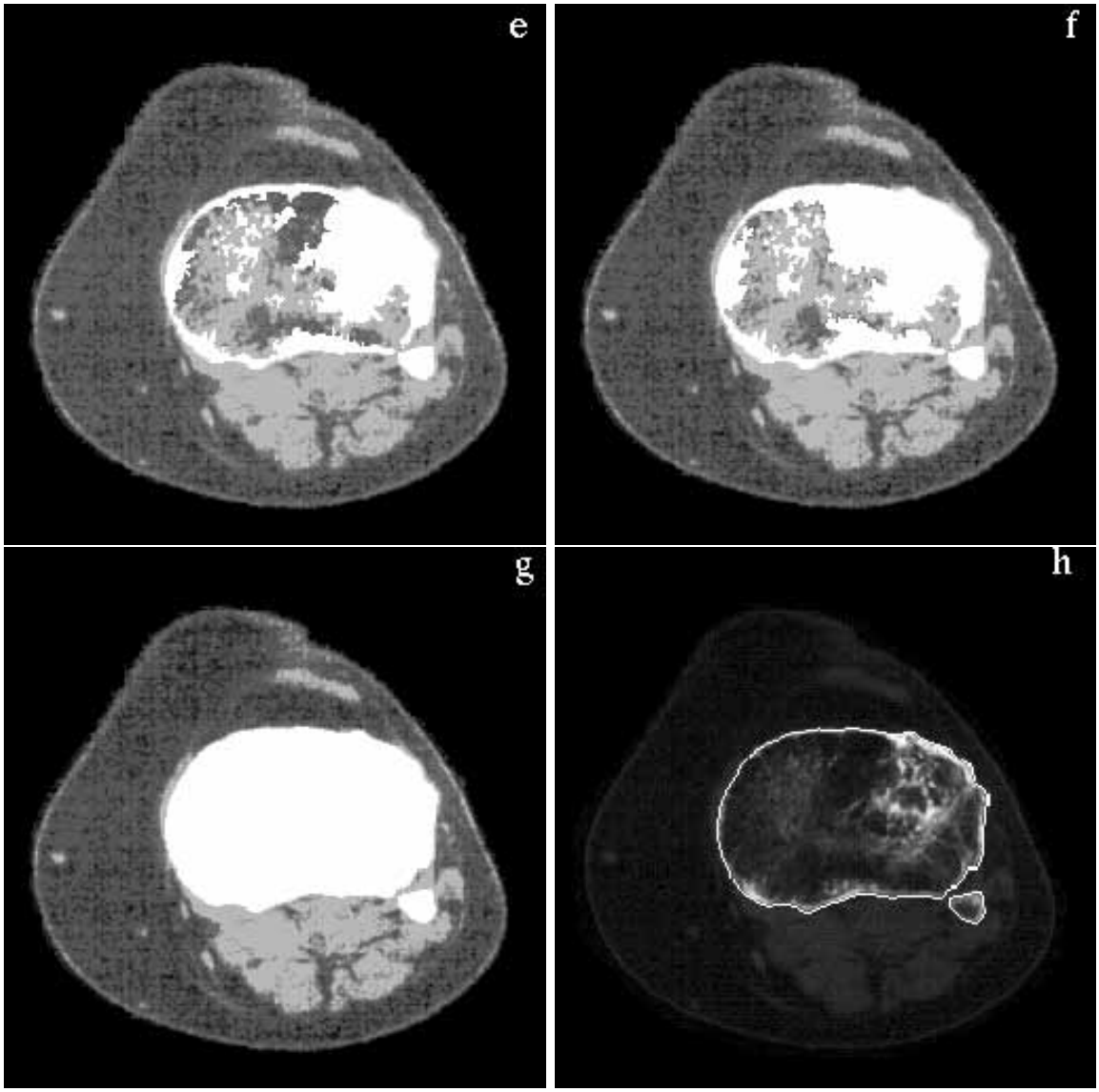


Figura 3.30: Continuación de La figura 3.23: (e)-(g) estados intermedios del crecimiento de regiones; (h) segmentación final superpuesta a la imagen original.

se podrá conseguir mediante la aplicación de modelos deformables, los cuales ofrecen en general una mayor precisión en la localización de los contornos siempre que el modelo inicial esté próximo al contorno final deseado. La figura 3.32.b muestra el resultado de la segmentación para la imagen de la figura 3.16. En esta imagen, aunque el hueso cortical

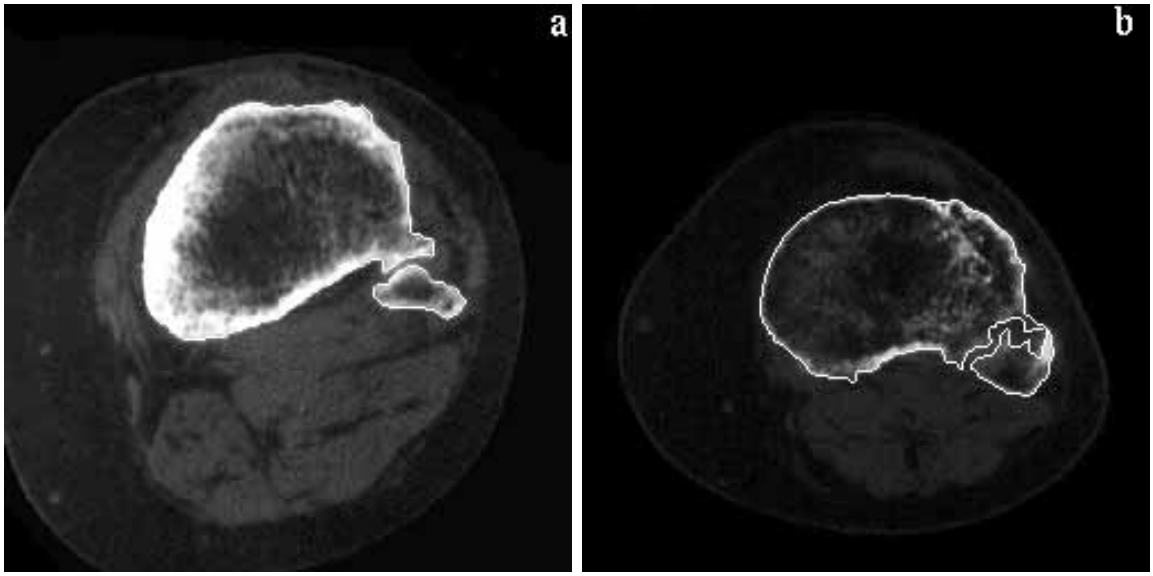


Figura 3.31: (a) Ejemplo de segmentación compleja resuelta con éxito; (b) caso de segmentación errónea.

es muy estrecho y de baja densidad, se consigue en general una buena segmentación. De todas formas, algunos puntos del contorno están mal localizados. El ajuste mediante modelos deformables proporcionaría una localización más precisa.

En la figura 3.33 se muestra la segmentación para el caso de la figura 3.15. Como se parte de un caso de infrasegmentación, se acaban uniendo tibia y peroné de la forma que se muestra en la figura 3.33.a. Es preciso recurrir a información de bordes para separarlas de la forma que se muestra en la figura 3.33.b. De esta forma se obtiene una segmentación correcta.

Las valoraciones que hemos hecho sobre la bondad de la segmentación son de tipo subjetivo y cualitativo proporcionadas por un operador humano. Sin embargo, si se pretende realizar un procesado automático, no deberíamos usar este operador; por tanto, es preciso introducir medidas objetivas de tipo cuantitativo que puedan dar una estimación de la calidad de la segmentación. Las medidas de rendimiento son un factor clave para la proporción de un camino de realimentación mediante el cual un sistema pueda modificar su estrategia durante el procesado [Levine y Nazif, 1985]. Para sistemas con metas bien definidas, los parámetros de rendimiento son los que miden la distancia a la meta

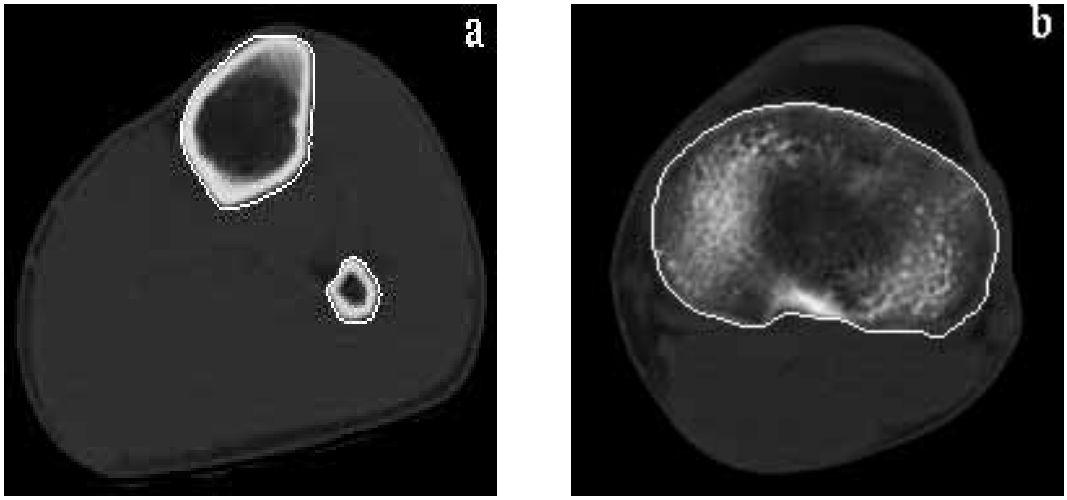


Figura 3.32: Ejemplo de segmentación con necesidad de ajuste de localización de bordes.

en cualquier instante de tiempo. Desafortunadamente, los sistemas de segmentación de imágenes no caen dentro de esta categoría, al menos en el procesado de bajo nivel. La meta sólo podrá definirse de forma precisa en términos de alto nivel. Hay tres factores que hacen que el procesado de bajo nivel sea difícil: el primero es que los pixels correspondientes al mismo tejido pueden exhibir diferentes intensidades entre cortes, o incluso dentro del mismo corte; en segundo lugar está el problema de la distinta resolución entre cortes respecto de la existente en el plano; y en tercer lugar, que los contornos de los tejidos son borrosos debido a las grandes dimensiones de los voxels (artefacto de volumen parcial). Por ello, en las etapas de bajo nivel sólo se pueden esperar definiciones borrosas de las metas (uniformidad, contraste).

El problema es pues encontrar un conjunto de medidas que puedan ser calculadas y que simulen la evaluación humana de la misma segmentación. La mayoría de los métodos establecidos requieren de una segmentación de referencia. Sin embargo en un procesado automático el elemento humano no está generalmente disponible, lo cual los hace inútiles en la utilización en tiempo real. Los métodos de consenso y coherencia de bordes surgieron por la necesidad de eliminar tal dependencia [Levine y Nazif, 1985]. Sin embargo, no proporcionan medidas completas que puedan funcionar sin el conocimiento de la correcta segmentación u otra referencia precalculada.

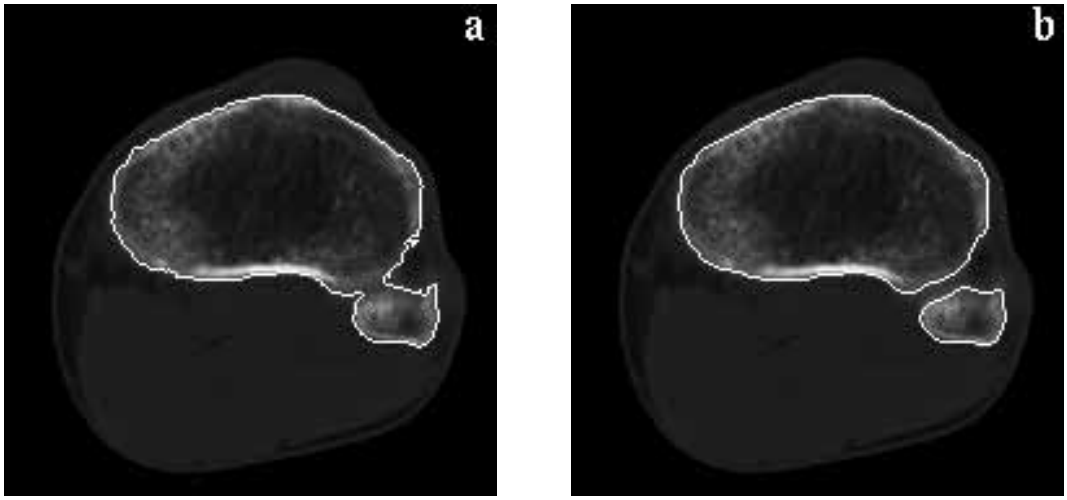


Figura 3.33: Segmentación con operación de división.

Las medidas de calidad de la segmentación en la etapa de bajo nivel no son suficientes por cuatro razones básicas: variabilidad de formas debido a presencia de alta componente de ruido, lesiones y patologías; la heterogeneidad de las estructuras óseas, el bajo contenido en información estructural que por separado aportan las regiones resultantes de la presegmentación, y la no unicidad de solución en la segmentación de bajo nivel. Por lo tanto, es preciso recurrir a medidas que se evalúen en la etapa de alto nivel, tras las operaciones de unión. Éstas serán medidas que decidirán si el resultado de la segmentación está de acuerdo con lo esperado, y que permitirán reconocer los objetos de interés en la escena. Estas medidas, en nuestro caso, se corresponden con las medidas de similaridad que ya hemos descrito y que conducen el proceso. Así, las estrategias y módulos de procesado previos tienden a la búsqueda de formas compactas y cerradas, dejando aún flexibilidad sobre la forma de las estructuras, ya que hay que contar con la presencia de posibles deformaciones (lesiones, patologías) que alteren la forma normal de las estructuras buscadas. De todos modos la valoración final de los resultados ofrecidos por el sistema es subjetiva al ser realizada por el operador humano. Consideramos que esta valoración es más significativa que cualquier otra que se pudiese obtener como distancia a la segmentación real ya que esta última tampoco estaría determinada de forma objetiva. Cuando se trabaja con imágenes médicas de individuos vivos se considera como segmentación real aquella que definiría el usuario experto. Por tanto, esta segmentación

real es de por si subjetiva y, en la mayoría de los casos, la segmentación trazada de forma manual resultará peor de la que ofrezca el propio sistema.

3.5 Segmentación mediante modelos deformables

Un planteamiento diferente y complementario a la partición de una imagen basada en características de homogeneidad (regiones) lo constituye el análisis basado en la detección de transiciones en los niveles de gris (bordes). En este apartado describiremos nuestro módulo de segmentación basado en contornos, comparando sus resultados con los obtenidos anteriormente mediante la unión y división de regiones.

Uno de los problemas que surgen al intentar segmentar una imagen mediante operadores de detección de bordes consiste en que los puntos de borde detectados pueden no corresponder con los contornos del objeto de interés. Con la excepción de imágenes de muy alta calidad en entornos controlados, los detectores de bordes producen bordes espúreos y discontinuidades. Por este motivo, el uso de los detectores de bordes es limitado en general y nulo en imágenes de pobre calidad [Staib y Duncan, 1992]. Las limitaciones de los detectores de bordes se deben, en parte, a su completa confianza sobre cálculos hechos directamente en un entorno local de los pixels de la imagen. Esto ignora tanto la información basada en modelos como organizaciones de la imagen de orden superior.

Se han propuesto distintos métodos de postprocesado de los segmentos extraidos para lograr la detección de contornos de objetos. En el capítulo 2 ya se hizo una revisión de estos métodos. Allí se referencian métodos de *programación dinámica* y *búsqueda heurística*, que han sido métodos habituales para la resolución del problema de optimización, pero presentan el inconveniente de carencia de robustez [Efford, 1994]. Otros métodos como la *transformada de Hough*, presentan problemas de almacenamiento y alto coste computacional cuando se trata con deformaciones.

Actualmente, los métodos basados en contornos que más interés despiertan en el dominio médico son los contornos deformables [Kass y col., 1988]. Los métodos basados en contornos deformables (*snakes*) parten de una hipótesis de contorno que, mediante procedimientos de mínimización de energía, es atraído hacia características de la imagen tales como líneas y bordes, mientras que fuerzas internas imponen condiciones de suavidad. En el capítulo anterior, en la sección 2.2.4 describimos sus características.

En este sentido, nosotros utilizaremos un método basado en snakes para mejorar la localización de bordes de interés. Las técnicas basadas en modelos deformables han mostrado su efectividad en la segmentación de estructuras anatómicas gracias a su capacidad para explotar ligaduras derivadas de los datos junto con el conocimiento a priori sobre localización, tamaño y forma de esas estructuras[McInerney y Terzopoulos, 1996].

Como habíamos comentado, una *snake* es una curva elástica aproximada que colocada sobre una imagen empieza a deformarse con el fin de ajustarse a las características notables de la escena. Esta deformación se produce por efecto de fuerzas externas que atraen la *snake* hacia características salientes de la imagen, y de fuerzas internas que intentan mantener la condición de suavidad en la forma de la curva. La solución final viene dada por el mínimo en la energía total de la snake, dada por la ecuación

$$E_{snake} = \int_0^1 [E_{int}(u(s)) + E_{ext}(u(s))]ds \tag{3.69}$$

donde E_{int} y E_{ext} son, respectivamente, la energía interna y la energía externa de la snake. La energía E_{int} viene dada por la suma de la energía de membrana, que expresa el *estiramiento* de la snake, y la energía de deformación plana, que expresa la *flexión* de la snake:

$$E_{int}(u(s)) = \frac{\alpha(s)|u_s(s)|^2 + \beta(s)|u_{ss}(s)|^2}{2} \tag{3.70}$$

donde $u(s) = (x(s), y(s))$ es la curva de la snake, s es la longitud de arco de la curva, y $u_s(s)$ y $u_{ss}(s)$ indican la primera y la segunda derivadas respectivamente. Los parámetros de elasticidad α y β controlan la suavidad de la curva.

La energía externa *empuja* la snake hacia las características de la imagen que estamos buscando. Términos típicos en la expresión de la energía externa son [Cohen y Cohen, 1993, Radeva y Serrat, 1993]: $E_I = \gamma I(u(s))$ que atrae la curva a puntos de intensidad alta (o baja); $E_G = -\delta|\nabla(G(u(s)) \times I(u(s)))|$ para la atracción de la curva hacia puntos de gradiente alto de la imagen convolucionada con un filtro Gaussiano para eliminación de ruido; y $E_E = -\zeta e^{-d(u(s))^2}$, donde $d(u(s)$ es la distancia al punto de borde más próximo, que hace que el contorno tienda a aproximarse a puntos de un mapa de bordes.

La solución al problema viene dado por la minimización del funcional de energía, proceso que se ha realizado por métodos diversos: diferencias finitas [Kass y col., 1988],

elementos finitos [Cohen y Cohen, 1993], programación dinámica [Amini y col., 1990], temple simulado (*simulated annealing*) [Storvik, 1994] o redes neuronales [Zhu y Yang, 1997]. Vilariño y col. [1997, 1998a, 1998b, 1998c] han propuesto métodos alternativos de posicionamiento final del contorno basados en redes neuronales celulares.

Nosotros hemos añadido nuevos términos a las expresiones de la energía interna y a la energía externa que mejoran el comportamiento de los modelos deformables en nuestro dominio de aplicación. Al mismo tiempo, hemos desarrollado un método de minimización de energía más rápido que los anteriores, que se aprovecha de las características concretas del proceso de segmentación [Pardo y col., 1998a]. En las siguientes subsecciones introduciremos todos estos nuevos elementos.

3.5.1 Nuevo término de energía interna

La naturaleza discreta de las snakes exige discretizar las derivadas numéricas en la expresión de la energía interna. Esto conduce a una expresión del tipo:

$$E_{int} = \sum_{i \in v_i} \frac{\alpha(i)(v_i - v_{i-1})^2 + \beta(i)(v_{i-1} - 2v_i + v_{i+1})^2}{2} \tag{3.71}$$

Tomando la derivada parcial de la ecuación 3.71 e igualando a cero se obtiene la siguiente expresión:

$$\begin{aligned} \sum_{i \in vi} \quad & \alpha_i \quad (v_i - v_{i-1}) + \alpha_{i+1}(v_{i+1} - v_i) + \\ & \beta_{i-1} \quad (v_{i-2} - 2v_{i-1} + v_i) + \beta_{i+1}(v_{i+2} - 2v_{i+1} + v_i) = 0 \end{aligned} \tag{3.72}$$

Considerando los coeficientes de elasticidad constantes, $\alpha_i = \alpha, \beta_i = \beta, \forall i$, resulta,

$$\sum_{i \in vi} \alpha(v_{i-1} - 2v_{i-1} + v_{i+1}) + \beta(-v_{i-2} + 2v_{i-1} - 2v_{i+1} + v_{i-2}) = 0 \tag{3.73}$$

Como podemos observar, el centroide de los puntos del contorno minimiza los términos en α y β [Chu and Aggarwal, 1993], por lo que, en ausencia de fuerzas externas importantes, la snake tenderá a la forma circular y a contraerse en un punto. Cuando se está procesando un conjunto de cortes consecutivos, la información que se transfiere de un corte a otro

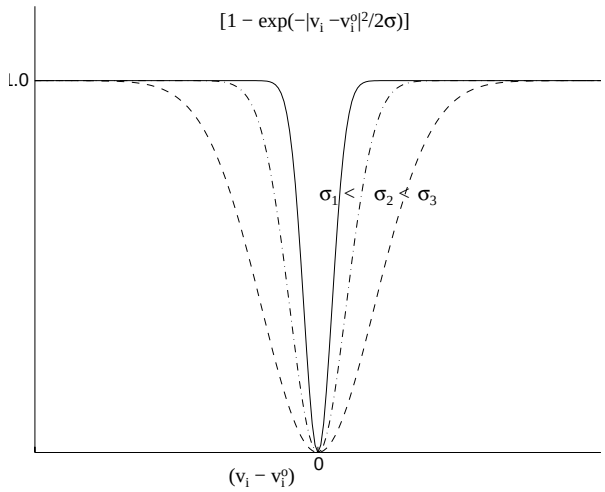


Figura 3.34: Comportamiento del nuevo término de energía interna, controlado por σ.

consiste en una forma y localización aproximadas de los contornos, e inherentemente, un tamaño aproximado de éste. Con la expresión clásica de la energía interna no se impone ninguna restricción basada en este conocimiento, de tal forma que éste puede no ser totalmente aprovechado. Nosotros consideramos que, para los casos en los que se dispone de un buen modelo inicial, es conveniente añadir una expresión del siguiente tipo para la energía interna:

$$\sum_{i \in v_i} \kappa(i) \left[1 - e^{-|v_i - v_i^0|^2/2\sigma}\right] \tag{3.74}$$

donde v_i^0 representa la posición del punto i-ésimo del modelo deformable inicial y σ controla la *flexibilidad* en la desviación respecto a esa posición inicial. Esta expresión *transporta*, de forma explícita, restricciones en cuanto a la desviación permitida y, de forma implícita, información relativa a la longitud del contorno. El parámetro σ indica la fuerza de la restricción, tal como se indica en la figura 3.34. Valores bajos de σ penalizan severamente las desviaciones grandes con respecto a la posición inicial de la curva. Mediante los parámetros $\kappa(i)$ controlamos el peso de este término en la expresión de la energía interna. La idoneidad de este término se puede observar en los resultados que presentaremos en las siguientes secciones.

El nuevo término de energía interna transporta información a cerca de la aproximación

del modelo. Si la aproximación es burda, debido por ejemplo a una gran separación entre cortes, σ deberá tomar un valor alto para permitir mayores variaciones de la *snake*. Si la separación es pequeña, entonces, σ deberá tomar un valor pequeño, evitando de esta forma que el contorno pueda verse atraído por texturas o ruido en posiciones de mala definición de los límites del objeto real, en donde la forma obtenida para cortes contiguos puede guiar mejor el posicionamiento del mismo. Por otra parte, cuando el objeto es pequeño o presenta bordes agudos, o el contorno presenta abundantes recovecos, el cálculo del gradiente en unos segmentos del contorno se ve afectado por la proximidad de otros segmentos de contorno. Esto origina un efecto de suavización en el mapa de gradientes y una eliminación de puntos de borde o su desplazamiento hacia el interior del objeto, lo que lleva a una reducción del tamaño detectado. Esto efecta a los términos de la energía externa. El nuevo término de energía interna que nosotros añadimos pretende contrarrestar el efecto anterior.

3.5.2 Nuevo término en el potencial externo

Uno de los problemas con el que han de enfrentarse los métodos basados en contornos deformables es el de las malas inicializaciones, que pueden provocar la *caida* de la snake en mínimos locales. Situaciones de este tipo se alcanzan, por ejemplo, cuando al tratar de determinar el contorno externo a un hueso, el modelo inicial se sitúa por la parte interior del hueso cortical, o cuando el modelo inicial se sitúa en la parte externa, pero más próximo al contorno de otra estructura. En un intento de mitigar estos problemas de mala inicialización proponemos un nuevo término de la energía externa que añadir a los clásicos de gradiente, intensidad y proximidad a puntos en un mapa de bordes. Así, la expresión del potencial que proponemos incluye los tres términos ya clásicos y un cuarto que introduce una fuerza normal al contorno. La expresión es:

$$\begin{aligned} P(u(s)) = & - \gamma\Big|\nabla[G(u(s)) * I(u(s))]\Big| - \delta I(u(s)) \\ & - \zeta e^{-d(u(s))^2} - \eta\frac{\partial}{\partial \overrightarrow{n}} I(u(s)) \end{aligned} \tag{3.75}$$

donde $d(u(s))$ denota la distancia entre el punto $u(s)$ y el borde más próximo. γ, δ, ζ son constantes positivas. El cuarto término provoca que la snake se desplace hacia el contorno externo, para el caso de objetos con doble contorno (hueso cortical), y hacia el contorno

del objeto más *próximo*, en términos de su centroide, en el caso de la proximidad de otras estructuras. Aquí $\overrightarrow{n}$ representa la normal a la snake. Cuando la estructura presenta un doble contorno, como por ejemplo, la tibia debido al hueso cortical, este nuevo término de potencial presenta el valor mínimo justo en la posición del contorno externo visto desde el centro de masas de la curva. Cuando el modelo inicial se sitúa entre dos estructuras

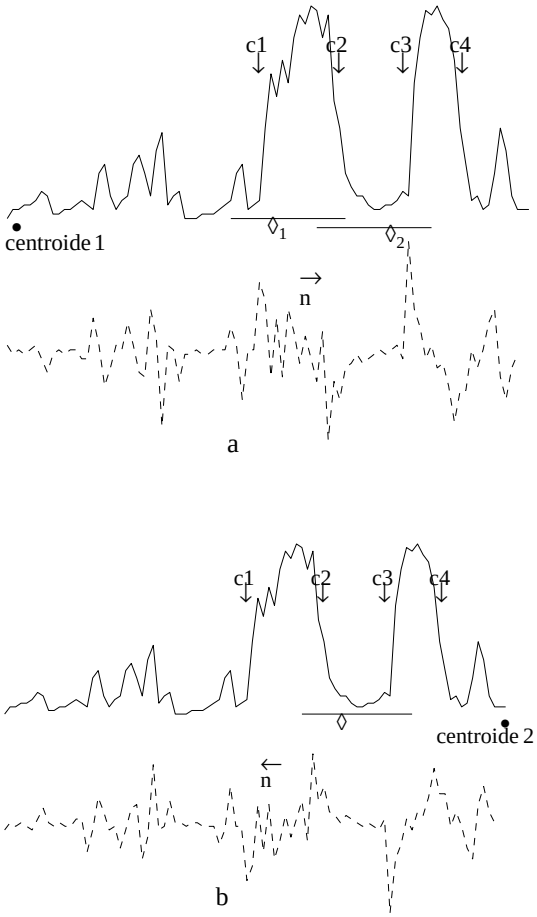


Figura 3.35: Ejemplo del efecto de la derivada normal. El perfil de la imagen aparece en trazado continuo y la derivada en trazado discontinuo. Con el $\diamond$ se representan las posibles localizaciones iniciales.

próximas, por ejemplo, entre tibia y peroné, este término de potencial atrae a la curva hacia el contorno más próximo al centro de masas de la snake; es decir, hacia el contorno

externo de la tibia si el modelo inicial corresponde al contorno de la tibia del corte previo, y hacia el contorno externo del peroné si el modelo inicial corresponde al del peroné del corte previo. La figura 3.35 ilustra el efecto del nuevo término de potencial. La figura corresponde a un perfil de imagen que comprende cuatro bordes fuertes y próximos. Estos son los contornos interno y externo del hueso cortical de la tibia (c1,c2) y del peroné (c3,c4). Se consideran dos casos: (a) la snake tiene el centro de masas en *centroide1*, rodeando de esta forma la tibia, y (b) la snake tiene el centro de masas en *centroide2* y rodea el peroné. Se representa también el valor de la derivada a lo largo de este perfil para los dos casos: en (a) la normal presenta dirección hacia a fuera de la tibia, y en (b) hacia a fuera del peroné. La dirección del vector normal viene indicado por un $\overrightarrow{n}$. Dos posibles posiciones iniciales de la snake en este perfil vienen marcadas por un diamante y los segmentos centrados en tales posiciones indican las zonas de desplazamiento permitido para la snake. Consideremos primero la figura 3.35.a, que responde al caso en el que el centroide de la snake es el marcado como 1 y la normal corresponde con el sentido indicado en la figura. Sin el efecto del nuevo término de potencial, la snake quedaría atrapada en el contorno fuerte más cercano, $c1$, para la situación inicial indicada por $\diamond_1$, y $c3$ para la indicada por $\diamond_2$. Sin embargo, el potencial normal empujará a la curva hacia la posición indicada por $c2$, tanto si la posición inicial es interna al borde ($\diamond_1$), como si se sitúa entre los dos bordes, ($\diamond_2$). Esa posición corresponde al mínimo de energía. Consideremos ahora la figura 3.35.b; ahora el vector normal tiene dirección opuesta y la curva de la derivada del perfil tiene polaridad opuesta. En esta nueva situación, la snake rodea al peroné, su centroide es el marcado como *centroide 2*, y con el efecto del nuevo término de potencial el mínimo de energía corresponde a $c3$, el contorno externo del peroné. Sin el efecto del nuevo término de potencial, el contorno externo del peroné quedaría atrapado en el contorno externo de la tibia (c2).

Cohen y Cohen [1993] presentaron una fuerza de presión interna normal ($k\overrightarrow{n}(v_i)$) a la superificie de la snake ($k\overrightarrow{n}$) intentando hacer el resultado final menos sensible a las condiciones iniciales. Sin embargo esta fuerza sólo es adecuada para modelos iniciales de contornos interiores al objeto, y además, en caso de bordes débiles podría hacer que la curva pasase sobre el contorno. La fuerza asociada al potencial que nosotros definimos empuja la curva hacia la parte externa del contorno, independientemente de la posición relativa (interna o externa) a éste. La presión de Cohen y Cohen siempre empuja en la misma dirección. Nuestro término de potencial decide la dirección en base a los datos

de la imagen. Tagare [1997], y Radeva y Martí [1995] proponen métodos basados en el ángulo formado entre la derivada normal a la snake con la orientación del gradiente en ese punto. Nuestro método presenta ante estos dos las ventajas de menor sensibilidad a mínimos locales y mayor velocidad de convergencia hacia el mínimo.

3.5.3 Método simplificado de minimización de energía

La solución al problema, detección del contorno, vendrá dada por la minimización de la función de la energía total de la snake. Amini y col. [1990] propusieron el uso de la programación dinámica para resolver problemas variacionales en visión por sus características de optimización global de la solución, estabilidad numérica y la posibilidad de imponer ligaduras fuertes en el comportamiento de la solución de una forma natural y directa. El suyo era un algoritmo iterativo donde en cada paso se aplicaba la programación dinámica para mejorar un poco el contorno. Sin embargo, los métodos de programación dinámica son lentos y requieren gran cantidad de memoria. Geiger y col. [1995] proponen un método más simplificado basado en el análisis del espacio de escalas y en la restricción de la zona de variación de la snake. Nosotros haremos uso de esta idea de restricción en la zona de variación para acelerar el proceso de convergencia.

En nuestro caso el modelo para el contorno deformable inicial (U^0) lo constituye el contorno correspondiente extraido para el corte previo, por lo que habrá, en general, un buen contorno aproximado inicial. Utilizaremos pues esta buena propiedad para reducir la complejidad del proceso de minimización.

De entre los puntos del contorno del corte previo, U^0, elegimos un conjunto de K de puntos de control, V^0. La elección de estos puntos puede hacerse de tal forma que sean equidistantes en el parámetro de la curva, s, ($u^0(s) \rightarrow v_i^0$, tal que $|v_i^0 - v_{i+1}^0| = \Delta s$), o bien determinar los puntos dominantes de la curva para que éstos aparezcan entre el conjunto de puntos de control. Nosotros optamos por el primer método por razones de simplificación de cálculo y porque la aplicación final es el modelado mediante elementos finitos, en donde es deseable poder generalizar a cerca de la enumeración de los elementos y generar parcelas de superficie no muy distorsionadas. Esto exige que en el modelado mediante elementos finitos la elección más adecuada de los puntos de control sobre cada contorno se haga empezando por el punto de contorno que sea la intersección del vector trazado desde el centro de masas del contorno y con la misma pendiente para todos los

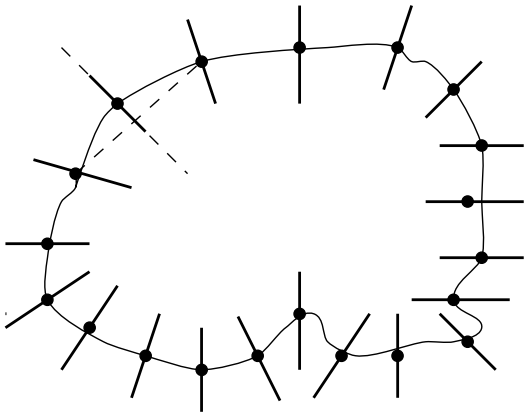


Figura 3.36: Ejemplo de contorno con sus puntos de control y los segmentos de posible variación de la curva.

cortes [Kang y col., 1993] . Por esta razón no tendría demasiado sentido el considerar puntos dominantes.

Elegimos, por tanto, un conjunto de puntos de control equidistantes en el parámetro de la curva, y definimos los segmentos de posible desplazamiento de los vértices en la perpendicular al segmento que une sus dos vértices vecinos, como se indica en la figura 3.36. Estos segmentos de variación permitida estarán centrados sobre cada uno de los puntos de control. A continuación se parte de un vértice arbitrario que consideramos inicialmente fijo y utilizamos programación dinámica para determinar la posición de los demás puntos de control del contorno para el mínimo de energía. Cambiando el vértice de comienzo en los sucesivos pasos del algoritmo se elimina la ligadura de pasar a través de puntos inicial y final fijos [Geiger y col., 1995]. Se repite el proceso hasta alcanzar un mínimo en la energía global de la snake. En cada iteración se recalcula el segmento de variación de los puntos de control para dar mayor flexibilidad a la forma de la snake. El segmento se determina a partir de la normal a la curva, cálculo que se aprovecha para determinar el nuevo término de potencial normal de la energía externa. El proceso, aunque subóptimo, es muy rápido y presenta resultados aceptables. La forma final completa se obtiene conectando los puntos de control finales de la snake mediante splines.

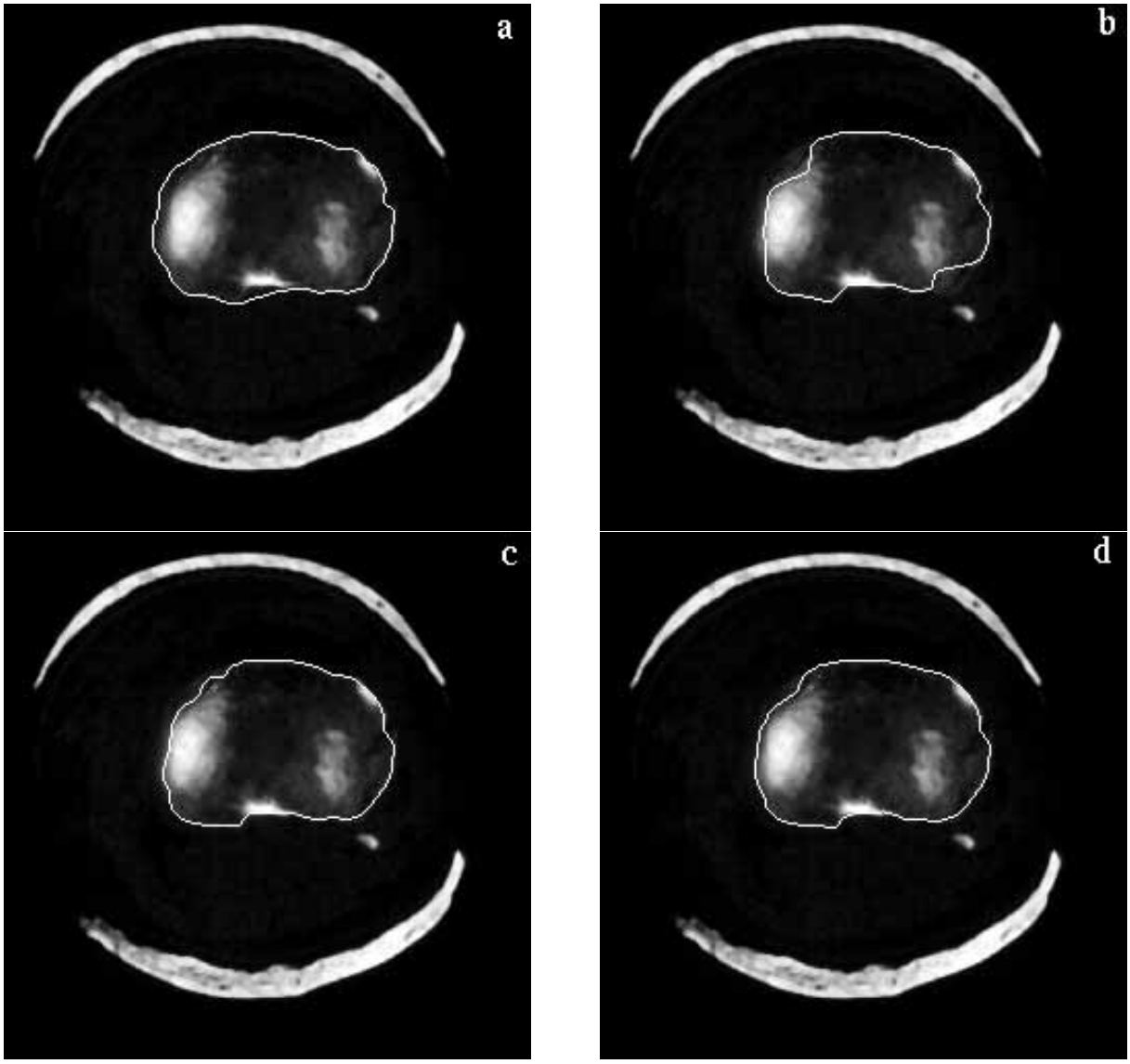


Figura 3.37: (a) Contorno inicial; (b) contorno final con las expresiones clásicas de E_{int} y E_{ext}; (c) contorno obtenido con la nueva expresión de E_{int} y la expresión clásica de E_{ext}; y (d) contorno obtenido con las nuevas expresiones de E_{int} y E_{ext}.

3.5.4 Resultados

El efecto del nuevo término de energía interna lo podemos observar en la figura 3.37. En los puntos de mala definición del contorno (parte superior derecha de la tibia) se observa

que, con la definición clásica de energía interna, la snake se desplaza hacia texturas internas (hueso trabecular) de la tibia. Se podría pensar en el aumento de valor en los parámetros de elascticidad de la snake, pero eso en general tendría el efecto negativo de una pérdida de flexibilidad de forma en las restantes partes de la snake. Otro comportamiento incorrecto de posicionado de contorno se observa en el peroné, en donde, al tratarse de una estructura pequeña, el valor del gradiente y la localización de los cruces por cero se ven muy afectados por la proximidad de otros puntos con gradientes importantes. El nuevo término de energía interna mejora la localización del contorno final al penalizar un desplazamiento excesivo con respecto al obtenido para el corte previo.

El efecto del nuevo término de potencial se puede observar en la figura 3.38. En la figura 3.38.b se observa como el contorno es atraido unas veces por el contorno interno y otras por el externo al hueso, dependiendo de la posición inicial de la snake y de la *fuerza* de las características de la imagen, junto con las condiciones de suavidad de la energía interna. En la figura 3.38.c se observa que, al incluir el nuevo término en $P(u(s))$, el contorno es atraído hacia la parte exterior de los dobles contornos. Aporta, por tanto, un criterio importante ante la presencia de puntos de borde fuertes que puedan desviar a la snake de su posición más idónea.

La segmentación mediante modelos deformables presenta una mayor eficiencia con respecto al sistema basado en regiones en lo que se refiere a la separación entre estructuras diferentes. En la figura 3.39 se muestra un caso en el que la segmentación por crecimiento de regiones, figura 3.39.a, no logra separar adecuadamente tibia y peroné debido a una incompleta información de bordes entre ambas estructuras. Los modelos deformables si consiguen separar ambas estructuras, figura 3.39.b. En general, los contornos suelen tener mayor precisión en su localización. Por contra, estos sistemas son más sensibles a ruido y texturas finas, por lo que aún existe la posibilidad de que el contorno se quede atrapado en texturas internas a la estructura de interés. Esta característica es opuesta a la presentada por el sistema basado en regiones, que en general presenta una cierta tendencia a la expansión de la segmentación del hueso por fuera de sus límites reales. Un ejemplo donde el crecimiento de regiones y los modelos deformables dan resultados comparables es el indicado en la figura 3.40.

Otra ventaja del sistema basado en crecimiento de regiones respecto del sistema basado en modelos deformables es la menor dependencia de los primeros con el modelo inicial, como se puede observar en la figura 3.41. Aquí se han tomado al azar dos cortes de

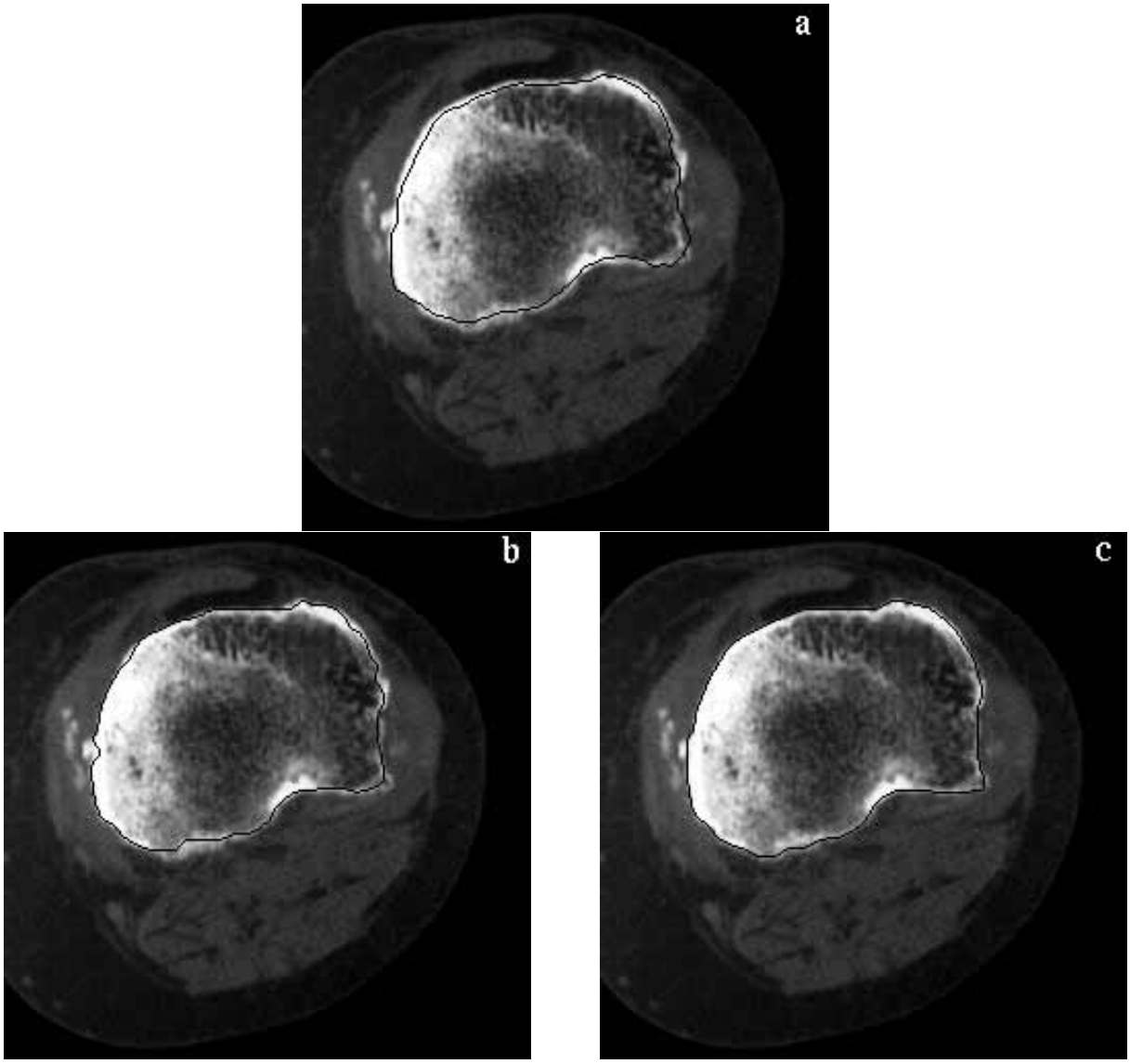


Figura 3.38: (a) Modelo inicial; (b) posicionamiento sin potencial normal; (c) posicionamiento con potencial normal.

una secuenca de imágenes CT, y se efectuarán varias segmentaciones considerando como modelo los contornos extraidos para las mismas estructuras en cortes adyacentes. Los índices representan las posiciones relativas de los cortes cuyos contornos finales se toman como modelos respecto del corte que se segmenta. Los errores se representan en uno de

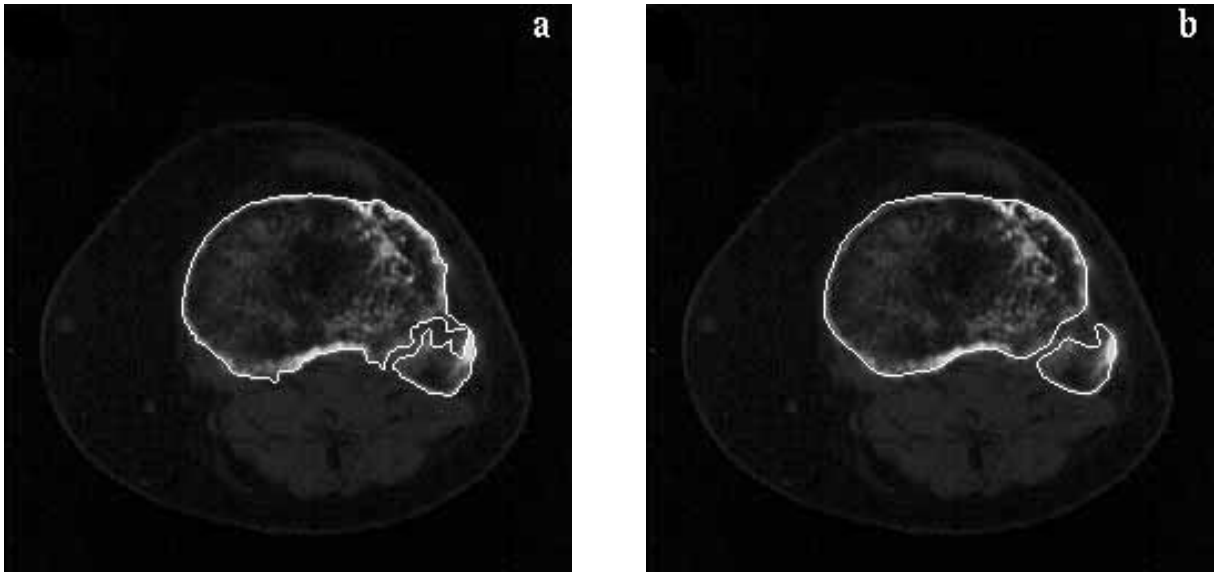


Figura 3.39: Segmentación de estructuras próximas: (a): por crecimiento de regiones; (b): mediante modelos deformables.

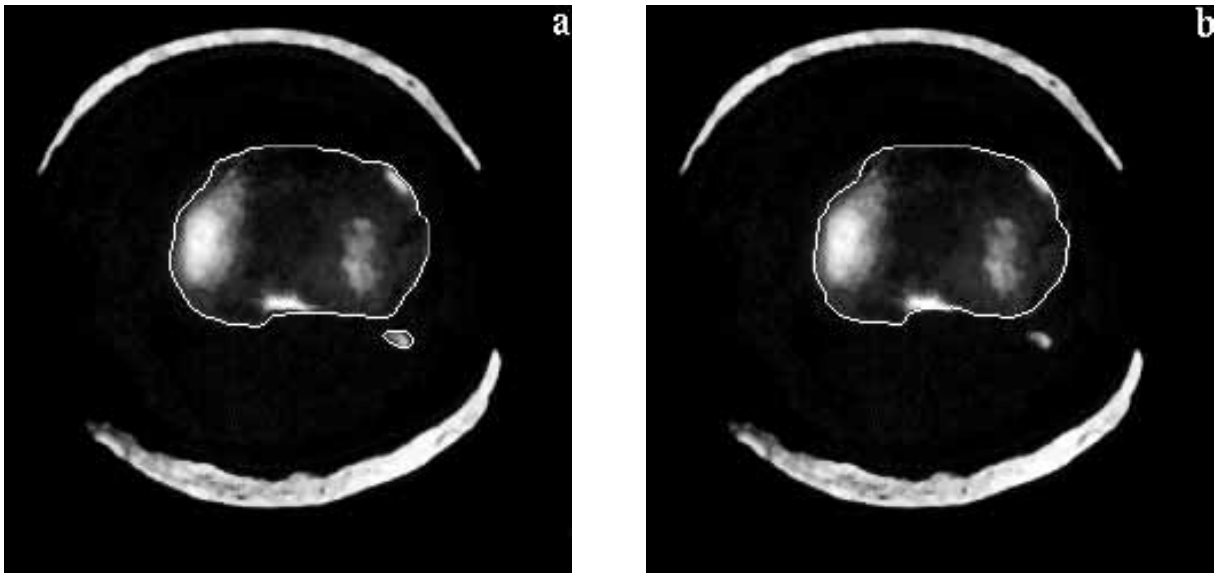
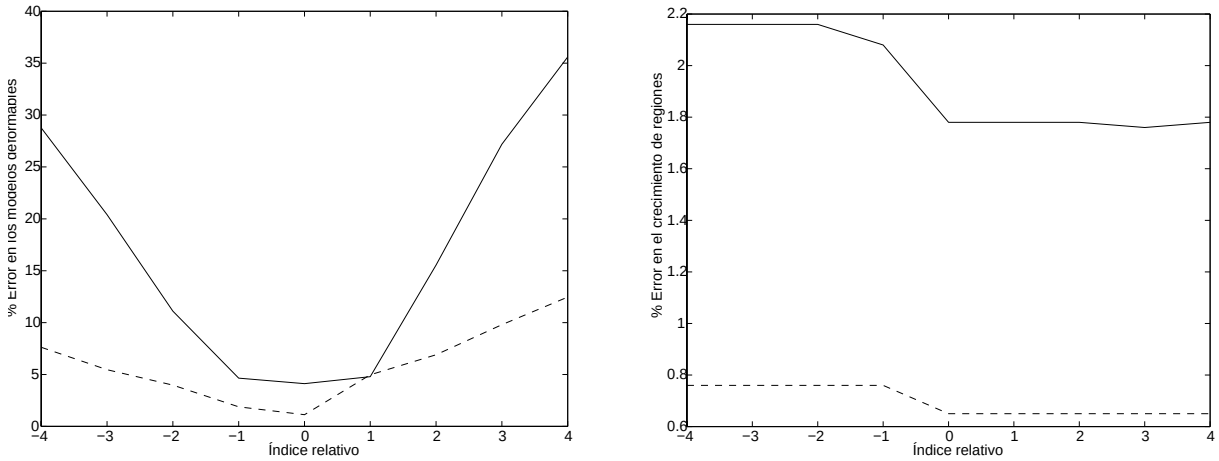


Figura 3.40: (a) Segmentación por crecimiento de regiones; (b) segmentación mediante modelos deformables.



(a) (b)

Figura 3.41: Dependencia de los modelos deformables, (a), y del crecimiento de regiones, (b), con el contorno inicial, para dos casos elegidos al azar. El eje de coordenadas representa la posición relativa respecto del corte actual de los cortes cuyos contornos finales se han tomado como contornos iniciales. El eje de abcisas representa el porcentaje de error respecto de una segmentación manual aproximada.

los casos con trazo continuo, y en el otro con trazos discontinuos. Para calcularlos se han comparado los resultados con una segmentación manual, y expresándose el error en % de pixels, con respecto al área de la segmentación manual, que aparecen sólo en una de las máscaras. En la figura se observa como los modelos deformables tienen una dependencia muy alta con el modelo inicial, lo cual no ocurre con el crecimiento de regiones.

La figura 3.42 muestra un caso donde se obtiene mejor resultado en la segmentación basada en crecimiento de regiones que mediante modelos deformables. Esto es debido a que la diferencia en forma de la estructura del corte actual con el previo es grande y los contornos se ven atraidos por texturas internas, afectando negativamente al comportamiento de los contornos activos.

A la vista de los resultados mostrados, se pone de manifiesto que un sistema automático no puede basarse únicamente en sólo uno de los métodos anteriores. Para obtener un sistema robusto es más adecuado añadir un nuevo módulo en el que se integren los resultados de la segmentación obtenidos por crecimiento de regiones y mediante modelos deformables.

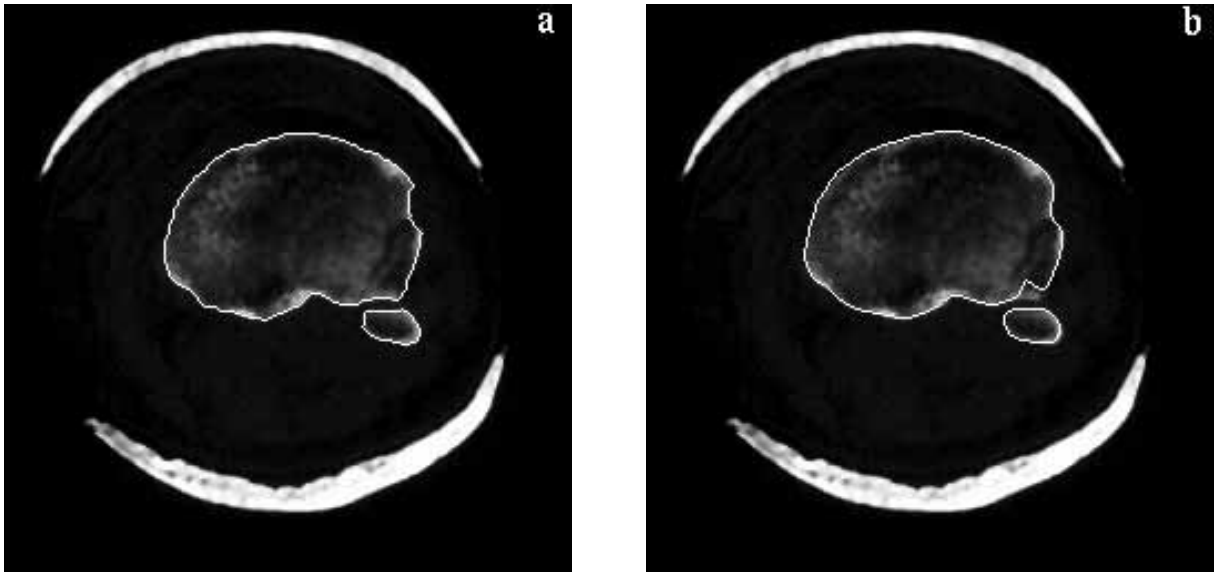


Figura 3.42: (a) Segmentación por crecimiento de regiones; (b) segmentación mediante modelos deformables.

La forma de llevar a cabo esta integración se describe en la siguiente sección.

3.6 Módulo de integración

En el procesado de imágenes CT de rodilla el método de crecimiento de regiones basado en conocimiento y el método basado en modelos deformables proporcionan en general buenos resultados con calidad equiparable. Sin embargo, en determinadas imágenes en las que se aprecia baja calidad de hueso cortical o lesiones, se generan segmentaciones defectuosas. Con el objetivo de minimizar el error de la segmentación automática es para lo que proponemos el esquema de integración de ambas metodologías.

Como ya hemos comentado en otras partes de la memoria, la segmentación mediante modelos deformables presenta, en general, con respecto al sistema basado en regiones, una mayor precisión en la localización de los contornos; la figura 3.42 ilustró este hecho. Por contra, son más sensibles al ruido y texturas finas; existe siempre la posibilidad de que el contorno se quede atrapado en texturas internas a la estructura de interés. Un ejemplo

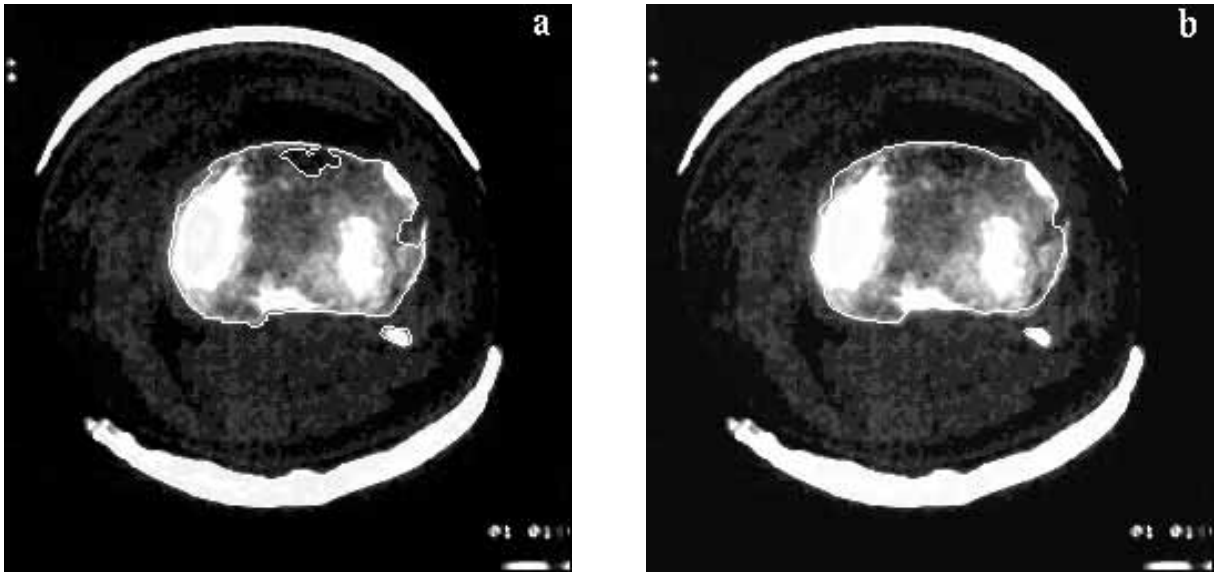


Figura 3.43: (a) Segmentación por crecimiento de regiones; (b) segmentación mediante modelos deformables.

se presenta en la figura 3.43.b. El método basado en crecimiento de regiones se basa en propiedades del entorno del pixel, más que en las características locales de éste; por ello tiende a la expansión de la segmentación del hueso por fuera de sus límites reales, como pudimos observar en la figura 3.44.a, y a obviar regiones estrechas y/o poco densas, como se ilustra en la figura 3.43.a.

Por otra parte, el sistema basado en crecimiento de regiones, respecto del sistema basado en modelos deformables, es menos sensible al ruido y presencia de texturas, y tiene menor dependencia con el modelo inicial. Esta última característica se puso de manifiesto en la figura 3.41, donde se escogieron dos cortes al azar y se realizó la segmentación tomando como modelos los contornos reales de las mismas estructuras en varios cortes adyacentes. Cuando existe una diferencia importante entre dos cortes, la segmentación mediante contornos deformables resulta poco fiable, como ilustra la figura 3.42. Es por esta razón por la que, a pesar de sus buenas propiedades, una segmentación robusta no puede basarse exclusivamente en ellos.

Como se desprende de lo anterior, los problemas no afectan de igual modo a ambos métodos, y como apuntan algunos autores, de la integración de la información propor-

cionada por diferentes metodologías podría resultar una optimización de la segmentación [Pavlidis y Liow, 1990; Chu y Aggarwal, 1993; Pankanti y Jain, 1995]. Sin embargo, esta integración casi siempre ha dado lugar a sistemas complejos o con gran dependencia con los algoritmos de segmentación empleados y sólo efectivos en el caso de estructuras homogéneas con respecto de alguna propiedad, no para el caso de estructuras constituidas por varias regiones [Haddon y Boyce, 1990; Pavlidis y Liow, 1990; Chakraborty y col., 1996; Tek y Kimia, 1997; Cho y Meer, 1997]. Nosotros proponemos la integración de los resultados obtenidos por separado por ambos métodos.

En esta sección abordaremos el problema de la integración de la información proporcionada por los dos esquemas de segmentación que hemos descrito; el objetivo final es obtener una segmentación más precisa. La integración de los métodos basados en regiones y en contornos activos se realizará en dos situaciones diferentes. La primera de ellas se corresponde con la que ilustra la figura 3.44, en la que se se realiza el ajuste de los contornos obtenidos mediante el crecimiento de regiones utilizando el método de los modelos deformables. En este caso, el contorno inicial será el resultante del crecimiento. Este modo de integración se realizará cuando se obtiene una buena medida de similaridad de forma ($\mu \geq 0.95$) entre objetos reconocidos y modelo.

La segunda surge cuando existen problemas de infrasegmentación en el sistema de crecimiento de regiones, de tal modo que es preciso recurrir a la división de regiones, o bien siempre que no se obtenga una correspondencia con los modelos aproximados. En estos casos recurriremos siempre a la integración de los resultados obtenidos de forma independientes por ambos métodos. Previamente, para la aplicación del sistema basado en modelos deformables usaremos como contorno inicial el obtenido por interpolación de forma entre la máscara resultado del crecimiento de regiones, y la máscara obtenida en el procesado del corte previo [Raya y Udupa, 1990]. Con esta interpolación de forma se pretende proporcionar un modelo lo más aproximado posible. Ejemplos de esta situación son los mostrados en las figuras 3.31.b ó 3.42.a.

A continuación describiremos dos esquemas de integración. El primero de ellos hace una integración de información usando reglas que resumen conocimiento sobre el comportamiento de ambos métodos sobre el tipo de imágenes propias de este dominio [Pardo y col., 1997a]. El segundo se basa en un esquema más formal, como es el de los Campos Aleatorios de Markov. Los resultados proporcionados por ambos esquemas son equiparables cuando las discrepancias entre los resultados proporcionados por los dos métodos de

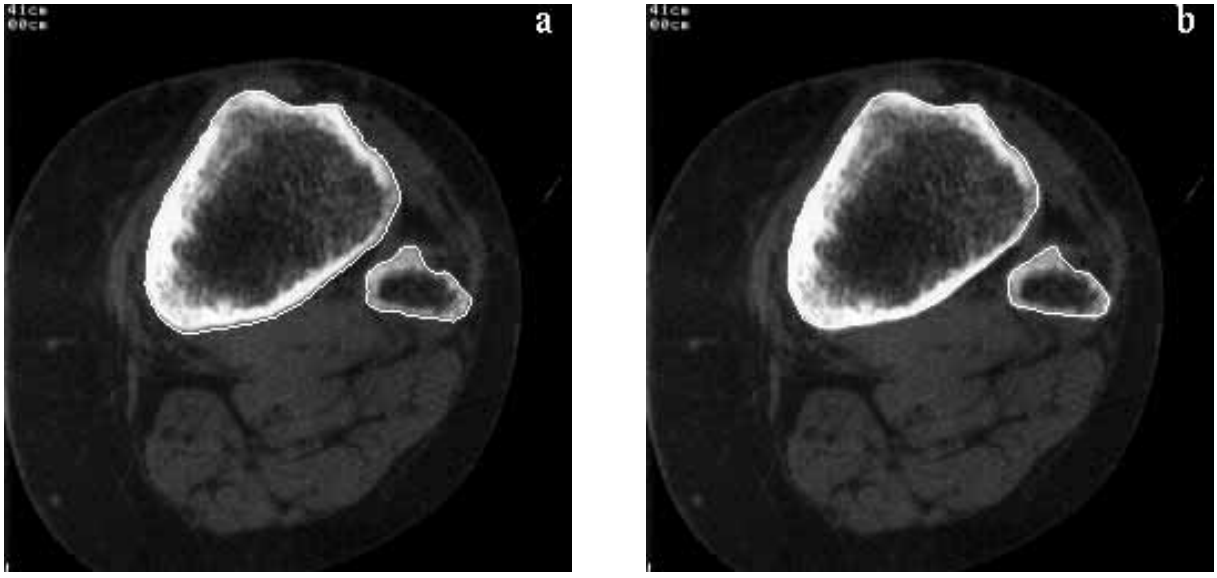


Figura 3.44: (a) Segmentación resultante del crecimiento de regiones. (b) Ajuste mediante modelos deformables usando como contorno inicial el mostrado en (a).

segmentación no son demasiado grandes. El primero de ellos quizás deje más clara la forma de realizar la integración y es más rápido, pero el segundo es un esquema más adecuado por presentar menor dependencia con parámetros, mayor facilidad en su extensibilidad y adecuación a nuevas situaciones, y ofrece mejores resultados cuando la discrepancia entre los dos métodos es grande.

3.6.1 Integración de información mediante heurísticas

Los módulos definidos con anterioridad intentan optimizar su rendimiento mediante el uso de heurísticas y conocimiento específico sobre el dominio. En éste se intenta integrar la información aportada por ambos mediante el uso de conocimiento del dominio y del comportamiento de los métodos de segmentación previos sobre este tipo de imágenes [Pardo y col., 1997a]. Para ello se parte de las regiones encerradas por los contornos obtenidos en el detector de bordes, que se clasifican en hueso (tibia, peroné) o fondo, según la estructura del corte previo. Las regiones resultantes de una operación AND entre las máscaras resultantes del crecimiento de regiones y éstas se confirman como

hueso, conservando su etiqueta. Sin embargo, aquellas obtenidas como resultado de la función EXOR se les asigna la etiqueta de región a analizar. En este punto hay que hacer constar que, en los puntos de bifurcación, los modelos de contorno a priori utilizados tanto en el crecimiento de regiones como en el ajuste de bordes serán menores en número que los objetos reales. Esto puede llevar a la unión de objetos diferentes en el módulo de crecimiento de regiones, no así en el de detección de contornos. Si el número de objetos finalmente detectados en este corte es distinto que en el anterior se habrá producido una bifurcación, de tal modo que a partir de ese momento habrá más de una estructura con la misma etiqueta, pero separadas.

Por otra parte, cada región R del EXOR llevan una etiqueta que las identifica como procedente de la detección de bordes R^e, o como clusters iniciales del crecimiento de regiones R^{rg}. Las siguientes reglas, que resumen conocimiento heurístico del comportamiento de los métodos de segmentación sobre imágenes de CT de estructuras óseas, decidirán sobre la unión o eliminación de las regiones a analizar:

SI:
(1) R_i NO es ADYACENTE a HUESO
(2) R_i NO ES un HUESO
ENTONCES:
(1) UNIR R_i a FONDO

SI:
(1) R_i es ADYACENTE a R_j
(2) R_j ES un HUESO_l
(3) R_i es ADYACENTE a R_k
(4) R_k ES un HUESO_m
(5) $l \neq m$
ENTONCES:
(1) UNIR R_i a FONDO

SI:
(1) R_i es ADYACENTE a R_j
(2) R_j ES un HUESO_l
(3) $\frac{\text{PERÍMETRO}_{(R_i \cap R_j)}}{\text{PERÍMETRO}_{(R_i)}} > TH$
ENTONCES:
(1) UNIR R_i a HUESO_l

SI:
(1) R_i es ADYACENTE a R_j
(2) R_j ES un HUESO_l
(3 La COMPACTACIÓN(R_i) ES BAJA
(4) La ELONGACIÓN(R_i) ES BAJA

(5) La DENSIDAD(R_i) ES BAJA
ENTONCES:
(1) UNIR R_i a HUESO$_l$

SI:
(1) R_i es ADYACENTE a R_j
(2) R_j ES un HUESO$_l$
(3) La COMPACTACIÓN(R_i) es BAJA
(4) La ELONGACIÓN(R_i) es BAJA
(5) La DENSIDAD(R_i) es BAJA
ENTONCES:
(1) UNIR R_i a FONDO

SI:
(1) R_i^{rg} es ADYACENTE a R_j
(2) R_j ES un HUESO$_l$
(3) La COMPACTACIÓN(R_i^{rg}) es BAJA
(4) La ELONGACIÓN(R_i^{rg}) es BAJA
ENTONCES:
(1) UNIR R_i^{rg} a FONDO

SI:
(1) R_i^e es ADYACENTE a R_j
(2) R_j ES un HUESO$_l$
(3) La COMPACTACIÓN(R_i^e) es BAJA
(4) La ELONGACIÓN(R_i^e) es ALTA
ENTONCES:
(1) UNIR R_i^e a HUESO$_l$

En la figura 3.45 se muestra el resultado de la integración heurística de las segmentaciones de la figura 3.43. Aquí se observa como la integración ofrece un mejor resultado que los dos métodos de segmentación por separado. Hay que recordar en este punto que, en general, los métodos de segmentación ofrecen buenos resultados; la integración sólo va a ser necesaria en un número limitado de casos conflictivos. En el caso de la figura 3.43 aparecen dos lesiones consistentes en el hundimiento de la meseta tibial. Estas lesiones corresponden a las zonas con nivel de gris más alto dentro de la tibia. Este hundiminto fue provocado por una fractura causada por un efecto de pérdida ósea. La baja calidad del hueso cortical hace que las transiciones de niveles de gris entre tibia y músculo sean mucho mas suaves que las que existen alrededor de las lesiones. Como consecuencia, la segmentación basada en regiones no detecta en algunos casos la presencia de hueso cortical debido a la estrechez del mismo y su baja densidad, mientras que en otros, los contornos deformables son atraidos en mayor medida por el contorno de una de las lesiones más

Figura 3.45: Resultado de la integración heurística de las segmentaciones de la figura 3.43.

que por el contorno entre hueso y músculo. El esquema de integración conjuga ambos resultados para ofrecer un mejor resultado, como se muestra en la figura 3.45.

Un ejemplo de integración con este método, que presenta un error, se muestra en la figura 3.46. En ella se muestran resultados intermedios y la segmentación final de uno de los cortes de la secuencia. Las figuras 3.46.a y 3.46.b muestran, respectivamente, la imagen del corte original y los dos modelos de estructuras a identificar. En las figuras 3.46.c y 3.46.d se indican, superpuestos a la imagen original, los resultados obtenidos por los dos métodos de segmentación de forma separada. En la figura 3.46.e se muestra, con un nivel de gris mayor que 0 y menor que 255, las regiones que evaluará el sistema de integración. Las regiones con el nivel de gris más alto (255) ya serán consideradas desde el principio como hueso. Como podemos observar en la figura 3.46.f, esta estrategia de integración falla en la parte superior del peroné. Aquí existe un hueso cortical muy poco denso y estrecho que delimita hueso trabecular de muy baja densidad. El esquema basado en los Campos Aleatorios de Markov resolverá todos los problemas de segmentación que se muestran en la figura 3.46.c y 3.46.d, dando como resultado una aproximación más precisa en la delimitación de los contornos de ambas estructuras anatómicas.

Mostraremos más ejemplos de este tipo de integración junto con los de la integración

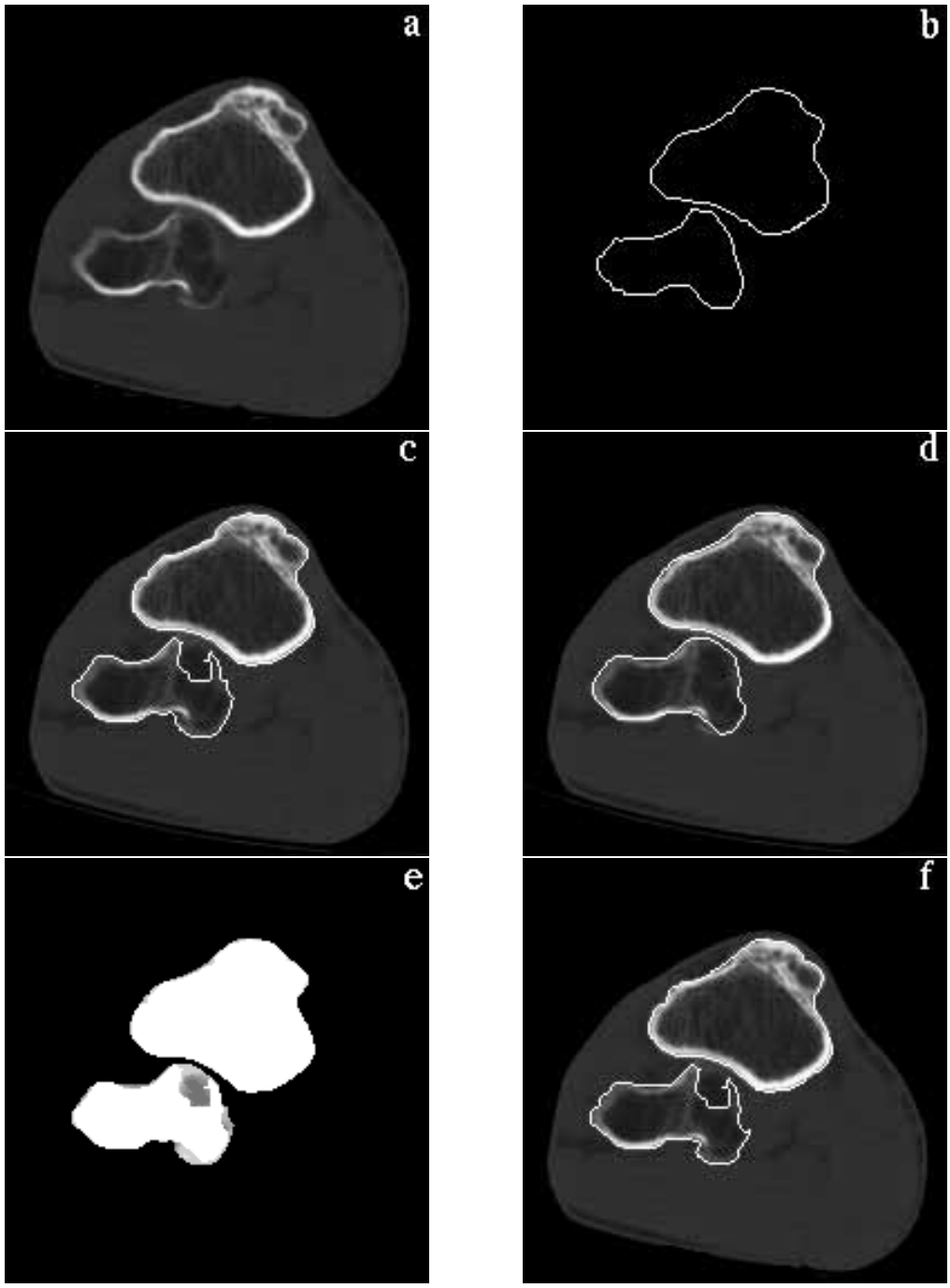


Figura 3.46: (a) Imagen original; (b) modelo; (c)segmentación de regiones; (d) segmentación de bordes; (e) entrada para integración; (f) resultado de integración.

mediante Campos Aleatorios de Markov con el fin de establecer una comparación.

3.6.2 Integración de información mediante MRF

Aunque el método presentado con anterioridad puede resolver el problema de integración de la información de regiones y bordes, consideramos que una forma más elegante y a la vez más cómoda para tratar de integrar todos los casos posibles es mediante un modelo de Campos Aleatorios de Markov.

Campos aleatorios de Markov y distribuciones de Gibbs

Con frecuencia, los problemas de visión se formulan en términos de la optimización de algún criterio, ya sea de forma implícita o explícita. La utilización de los principios de optimización se debe a las incertidumbres presentes en los procesos de visión (ruido, ambigüedades en la interpretación, etc.). En general no existe la solución *exacta*. En su lugar se busca la solución óptima. La teoría de los Campos Aleatorios de Markov (*Markov Random Fields* (MRF)) proporciona un fundamento matemático para resolver el problema de hacer inferencias globales usando información local. Se usa en problemas de etiquetado para establecer distribuciones probabilísticas de etiquetas interactuantes [Li, 1995].

Muchos problemas en visión se pueden plantear como problemas de etiquetado en los que la solución al problema es un conjunto de etiquetas asignados a pixels u otras características en la imagen. El problema de etiquetado se especifica en términos de un conjunto de *sitios* y un conjunto de etiquetas. Sea $\mathcal{S}$ un conjunto discreto de sitios,

$$\mathcal{S} = \{1, \ldots, m\} \tag{3.76}$$

en donde $1, \ldots, m$ son índices. Un sitio es generalmente un punto o una región en el espacio Euclídeo, tal como un pixel, un borde, o una región homogénea. Una etiqueta es un evento que puede ocurrir en un sitio. Sea L el conjunto de etiquetas, que puede ser continuo o discreto. En el caso discreto, las etiquetas asumen valores discretos en un conjunto de M etiquetas,

$$\mathcal{L} = \{1, \ldots, M\} \tag{3.77}$$

En la detección de bordes, por ejemplo, el conjunto de etiquetas sería $\mathcal{L}$ ={borde, no–borde} $\equiv$ {1,0}. Cuando las etiquetas representan valores cuantizados, se habla de conjunto de etiquetas ordenado, y en ellas se puede definir una medida de similaridad.

El problema de etiquetado consiste en asignar una etiqueta de un conjunto $\mathcal{L}$ a cada uno de los sitios en $\mathcal{S}$. El conjunto

$$f = f_1, \ldots, f_m \tag{3.78}$$

se llama *etiquetado* de los sitios de $\mathcal{S}$ en términos de las etiquetas de $\mathcal{L}$. Cuando a cada sitio se le asigna una única etiqueta, $f_i = f(i)$ se puede considerar como una función con dominio $\mathcal{S}$ e imagen $\mathcal{L}$:

$$f : \mathcal{S} \rightarrow \mathcal{L} \tag{3.79}$$

En visión, un etiquetado o *configuración* puede corresponder a una imagen, un mapa de bordes, una interpretación de características de la imagen en función de características de un modelo, etc.

Sea $\mathcal{F} = F_1, \ldots, F_m$ una familia de variables aleatorias definidas en el conjunto $\mathcal{S}$, en el que cada variable F_i toma un valor f_i en $\mathcal{L}$. La familia $\mathcal{F}$ se llama campo aleatorio. Para un conjunto $\mathcal{L}$ discreto, la probabilidad de que la variable aleatoria F_i tome el valor f_i se denota por $P(F_i = f_i) \equiv P(f_i)$, y la probabilidad conjunta se denota $P(\mathcal{F} = f) = P(F_1 = f1, \ldots, F_m = f_m) \equiv P(f)$. Cuando todos los sitios tienen el mismo conjunto de etiquetas $\mathcal{L}$, el conjunto de todos los posibles etiquetados o el espacio de configuracione, es el siguiente producto Cartesiano

$$\mathcal{F} = \mathcal{L} \times \mathcal{L} \times \ldots \times \mathcal{L} = \mathcal{L}^m \tag{3.80}$$

Para un conjunto discreto de m sitios y M etiquetas existen un total de M^m configuraciones.

El etiquetado con ligaduras contextuales es indispensable para la interpretación de imágenes [Haralick y Shapiro, 1992]. En términos de probabilidades, las ligaduras contextuales se pueden expresar localmente en términos de las probabilidades condicionales $P(f_i|\{f_{i'}\})$, donde $\{f_{i'}\}$ denota el conjunto de etiquetas en los sitios $i \neq i^{'}$, o globalmente como la probabilidad conjunta $P(f)$. Puesto que la información local se observa

más directamente, es normal que la inferencia global se haga basándose en propiedades locales.

En situaciones donde las etiquetas son independientes unas de otras (no hay contexto), la probabilidad conjunta es el producto de las probabilidades locales

$$P(f) = \prod_{i \in \mathcal{S}} P(f_i); \tag{3.81}$$

es decir, no hay dependencia condicional:

$$P(f_i|\{f_{i'}\}) = \prod_{i \in \mathcal{S}} P(f_i) \quad i^{'} \neq i \tag{3.82}$$

Sin embargo, en la presencia de contexto las etiquetas son mutuamente dependientes, por lo que las relaciones anteriores no se mantienen. La teoría de los campos aleatorios de Markov es una rama de la teoría de la probabilidad que permite analizar dependencias contextuales en problemas físicos. Así, los sitios en $\mathcal{S}$ están relacionados unos con otros a través de un sistema de vecindad, el cual se define como

$$\mathcal{N} = \{\mathcal{N}_i | \forall i \in \mathcal{S}\} \tag{3.83}$$

donde $\mathcal{N}_i$ es el conjunto de vecinos de i. La relación de vecindad tienen las siguientes propiedades:

1. un sitio no es vecino de si mismo: $i \notin \mathcal{N}_i$;
2. la relación de vecindad es mutua: $i \in \mathcal{N}_{i'} \iff i^{'} \in \mathcal{N}_i$;

Para un conjunto $\mathcal{S}$ regular, el conjunto vecino de i se define generalmente como el conjunto de sitios dentro de un radio r

$$\mathcal{N}_i = \{i^{'} \in \mathcal{S} / ||pixel_{i'}, pixel_i||^2 \leq r, i^{'} \neq i\} \tag{3.84}$$

donde $||\cdot, \cdot||$ denota la distancia Euclídea. Para un conjunto $\mathcal{S}$ irregular, el conjunto vecino de i se define, análogamente, como el conjunto de sitios dentro de un radio r

$$\mathcal{N}_i = \{i^{'} \in \mathcal{S} / [dist(\text{característica}_{i'}, \text{característica}_i)]^2 \leq r, i^{'} \neq i\} \tag{3.85}$$

donde, ahora $dist(\cdot,\cdot)$ debe definirse adecuadamente.

El par $(\mathcal{S},\mathcal{N})$ constituye un grafo $\mathcal{G}$, donde $\mathcal{S}$ contiene los nodos y $\mathcal{N}$ determina los arcos. Se define un *clique* c, para $(\mathcal{S},\mathcal{N})$, como un subconjunto $\mathcal{S}$, consistente de un único sitio $c=\{i\}$, o, en general, colecciones de n sitios vecinos. Por ejemplo

$$\mathcal{C}_1 = \{i/i\in\mathcal{S}\} \tag{3.86}$$
$$\mathcal{C}_2 = \{\{i,i^{'}\}/i^{'}\in\mathcal{N}_i, i\in\mathcal{S}\} \tag{3.87}$$
$$\mathcal{C}_2 = \{\{i,i^{'},i^{''}\}/i^{''},i^{'},i\in\mathcal{S} \text{ son vecinos entre si}\} \tag{3.88}$$

Los sitios en un clique están ordenados, por lo que $\{i,i^{'}\}\neq\{i^{'},i\}$. La colección de todos los cliques para $(\mathcal{S},\mathcal{N})$ es

$$\mathcal{C}_1\cup\mathcal{C}_2\cup\mathcal{C}_3\cup\ldots \tag{3.89}$$

Se dice que $\mathcal{F}$ es un **campo aleatorio de Markov** sobre $\mathcal{S}$ con respecto al sistema de vecindad $\mathcal{N}$ si y sólo si se satisfacen las siguientes condiciones:

$$\begin{array}{ll} P(f)>0, \quad \forall f\in\mathcal{F} & (positividad) \\ P(f_i|f_{\mathcal{S}-\{i\}})=P(f_i|f_{\mathcal{N}_i}) & (Markovianidad) \end{array} \tag{3.90}$$

La positividad se asume por razones técnicas y se puede satisfacer generalmente en la práctica. Por ejemplo, cuando se satisface la condición de positividad, la probabilidad conjunta $P(f)$ de cualquier campo aleatorio se determina de forma única a partir de las probabilidades condicionales locales [Li, 1995]. La Markovianidad representa las características locales de $\mathcal{F}$. La etiqueta de un sitio depende sólo de los vecinos directos. Esta condición se puede satisfacer siempre con tal de seleccionar un $\mathcal{N}_i$ suficientemente grande (en el caso extremo todo $\mathcal{S}$).

Un MRF puede tener otras características tales como homogeneidad e isotropía. La homogeneidad significa que $P(f_i|f_{\mathcal{N}_i})$ es independiente de la posición relativa de i en $\mathcal{S}$. La isotropía implica independencia respecto de la orientación de los cliques.

La equivalencia entre los campos aleatorios de Markov y las distribuciones de Gibbs, demostrada por Hammersley y Clifford en 1971, proporciona una forma matemática tratable de especificar la probabilidad conjunta de un MRF.

Un conjunto de variables $\mathcal{F}$ se dice que es un **campo aleatorio de Gibbs** (GRF) en $\mathcal{S}$ con respecto a $\mathcal{N}$ si y sólo si obedece a una distribución de Gibbs:

$$P(f) = Z^{-1} \times e^{-\frac{1}{T}U(f)} \tag{3.91}$$

$$Z = \sum_{f \in \mathcal{F}} e^{-\frac{1}{T}U(f)} \tag{3.92}$$

$$U(f) = \sum_{c \in \mathcal{C}} V_c(f) \tag{3.93}$$

donde Z es una constante de normalización llamada *función de partición*, T es una constante llamada *temperatura* y $U(f)$ es la *función de energía* definida como suma de los potenciales de clique $V_c(f)$, los cuales dependen únicamente de la configuración local del clique c.

El teorema Hammersley–Clifford establece que $\mathcal{F}$ es un MRF en $\mathcal{S}$ con respecto a $\mathcal{N}$ si y sólo si $\mathcal{F}$ es un GRF en $\mathcal{S}$ con respecto a $\mathcal{N}$. La equivalencia entre los campos de Markov y Gibbs proporciona una forma simple de especificar la probabilidad conjunta. Se puede especificar la probabilidad conjunta $P(\mathcal{F} = f)$ especificando las funciones potenciales de clique, las cuales se definirán en función del comportamiento deseado. De esta forma se codifica el conocimiento a priori.

Para calcular la probabilidad conjunta de un MRF, que sea una distribución de Gibbs, es necesario evaluar Z. Como esto implica la suma de todas las configuraciones posibles, el cálculo es generalmente intratable. La evaluación explícita se puede soslayar en modelos de visión basados en la máxima probabilidad.

Optimización

La optimización ha jugado un papel esencial en visión computacional. La mayoría de los problemas de visión se formulan para optimizar algún criterio, ya sea de forma implícita o explícita. la utilización de los principios de optimización se debe a las incertidumbres presentes en los procesos de visión (ruido, ambigüedades en la interpretación, etc.). En general no existen solución *exacta*. En su lugar se busca la solución óptima. Se pueden citar varios ejemplos: transformada de Hough [Duda y Hart, 1973], restauración de imágenes (Geman y Geman, 1984), detección de bordes [Torre y Poggio, 1986],

reconocimiento e interpretación de objetos [Modestino y Zhang, 1992; Friedland y Rosenfeld, 1992], segmentación y etiquetado [Kim y Yang, 1996], etc.

Hay tres elementos básicos en la optimización: representación del problema, función objetivo, y algoritmos de optimización. En la segmentación se puede usar una cadena de puntos del contorno para representar la solución, o bien usar un mapa de regiones. Para MRF el mapa de regiones es la representación más natural [Li, 1995]. El segundo elemento es como formular una función objetivo para la optimización. La formulación determina como se codifican las ligaduras dentro de la función. El tercero es como optimizar el objetivo, e. d., como buscar la solución óptima en el espacio admisible. Aquí hay que considerar (1) el problema de los mínimos locales y (2) la eficiencia de los algoritmos en el espacio y en el tiempo. Estos problemas presentan, en general, soluciones encontradas. Los tres elementos están relacionados unos con otros: el esquema de representación influye en la formulación de la función objetivo, y estos dos influyen en el diseño del algoritmo de búsqueda.

Para el reconocimiento de patrones hay dos planteamientos básicos en la formulación de la función de energía: paramétrica y no paramétrica. En la paramétrica se conocen los tipos de las distribuciones, y se parametrizan con unos pocos parámetros. Por tanto, se puede obtener la forma funcional de la energía, y ésta queda completamente definida cuando se especifican los parámetros.

En el planteamiento no paramétrico no se hacen suposiciones a cerca de las distribuciones. Aquí, o bien se estima la distribución a partir de los datos, o se aproxima mediante funciones base con parámetros desconocidos. Puesto que los parámetros son parte de la formulación de la función de energía, la solución no está completamente definida si los parámetros no se especifican de alguna forma.

En los modelos formales, en oposición a los heurísticos, la función de energía se formula basándose en un criterio establecido. Cuando existe conocimiento a cerca de la distribución de los datos, pero no información a priori sobre la cantidad a estimar, se puede usar el criterio de máxima verosimilitud (*Maximum Likelihood*, ML). Cuando la situación es la opuesta, se puede usar el criterio de máxima entropía. Sin embargo ninguno de estos métodos es adecuado cuando se conocen las distribuciones a priori y la condicional (*likelihood*). En este caso lo mejor es maximizar un criterio de Bayes. Estos criterios son los más populares en visión computacional, y entre ellos, el de máxima probabilidad a

posteriori (MAP) es el más popular en el modelado mediante MRF's [Geman y Geman, 1984] .

En la estimación de Bayes se minimiza un riesgo para obtener la estimación óptima:

$$R(f^*) = \int_{f \in \mathcal{F}} C(f^*, f) P(f|d) df \tag{3.94}$$

donde d representa los datos de la imagen, $C(f^*, f)$ es una función de coste y $P(f|d)$ es la distribución posterior. Esta última se calcula de acuerdo con la regla de Bayes:

$$P(f|d) = \frac{p(d|f)P(f)}{p(d)} \tag{3.95}$$

donde $P(f)$ es la probabilidad a priori de los etiquetados, $p(d|f)$ es la f.d.p. condicional, y $p(d)$ es la densidad de d, la cual es constante.

La función de coste $C(f^*, f)$ determina el coste de la estimación f cuando f^* es la exacta. Una de las expresiones más populares de esta función es la siguiente:

$$C(f^*, f) = \begin{cases} 0 & \text{si} \quad ||f^* - f|| \leq \delta \\ 1 & \text{otro caso} \end{cases} \tag{3.96}$$

donde $\delta > 0$ es una constante pequeña. Para esta función, el riesgo de Bayes es:

$$R(f^*) = \int_{f:||f^*-f||>\delta} C(f^*, f) P(f|d) df = 1 - \int_{f:||f^*-f||\leq\delta} C(f^*, f) P(f|d) df \tag{3.97}$$

que, cuando $\delta \to 0$, se aproxima por

$$R(f^*) = 1 - \kappa P(f|d) \tag{3.98}$$

donde κ es el volumen del espacio que contiene todos los puntos f para los que $||f^* - f|| \leq \delta$. La minimización de este riesgo es equivalente a maximizar la probabilidad a posteriori [Li, 1995]. Por tanto el mínimo riesgo es

$$f^* = arg \max_{f \in \mathcal{F}} P(f|d) = arg \max_{f \in \mathcal{F}} p(d|f) P(f) \tag{3.99}$$

En el etiquetado MAP–MRF esto resulta en

$$f^* = arg \min_{f} U(f|d) = arg \min_{f} \{U(d|f) + U(f)\} \tag{3.100}$$

Modelos MRF de alto nivel

En el análisis de imágenes de alto nivel se trata con características de la imagen que son más abstractas que los pixels. Una escena se representará mediante propiedades de estas características y relaciones entre ellas. Para esta representación se puede usar un grafo de la forma $\mathcal{G} = (\mathcal{S}, \mathcal{N}, d)$, donde d representa el conjunto de vectores de propiedades asignados a las características.

Para realizar el reconocimiento (matching) mediante MRF's es preciso definir los potenciales de los cliques. Como ejemplo podemos presentar una generalización de los potenciales a priori de Geman y Geman [1984]. Para cliques de primer orden, el potencial se define como

$$V_1(f_i) = \begin{cases} v_{10} & \text{si} \quad f_i = 0 \\ o & \text{otro caso} \end{cases} \tag{3.101}$$

donde v_{10} es una constante. Si f_i es la etiqueta NULA se incurre en una penalización. Esta etiqueta representa el no reconocimiento de la característica observada.

Para los cliques de segundo orden el potencial se define como

$$V_1(f_i, f_{i'}) = \begin{cases} v_{20} & \text{si} \quad f_i = 0 \quad \text{ó} \quad f_{i'} = 0 \\ 0 & \text{otro caso} \end{cases} \tag{3.102}$$

donde v_{20} es una constante. Si f_i ó $f_{i'}$ son la etiqueta NULA se incurre en una penalización. Estos potenciales de clique definen la energía a priori,

$$U(f) = \sum_{i \in \mathcal{S}} V_1(f_i) + \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{N}_i} V_2(f_i, f_{i'}) \tag{3.103}$$

La f.d.p. condicional de los datos observados, d, vistos como una función de f tienen las siguientes características:

1. Se condiciona para etiquetados no nulos $f_i \neq 0$.
2. Es independiente del sistema de vecindad $\mathcal{N}$.
3. Depende de como el objeto modelo es observado en la escena.

Asumiendo que la escena verdadera está corrompida por un ruido Gaussiano de media nula, la f.d.p. condicional es una distribución de Gibbs con la energía

$$U(d|f) = \sum_{i \in \mathcal{S}, f_i \neq 0} V_1(d_1(i)|f_i) + \sum_{i \in \mathcal{S}, f_i \neq 0} \sum_{i' \in \mathcal{S} - \{I\} i, f_{i'} \neq 0} V_2(d_2(i, i')|f_i, f_{i'}) \tag{3.104}$$

donde

$$V_1(d_1(i)|f_i) = \begin{cases} \sum_{k=1}^{K_1} \frac{[d_1^k(i) - D_1^k(f_i)]^2}{2[\sigma_1^k]^2} & \text{si } f_i \neq 0 \\ 0 & \text{en otro caso} \end{cases} \tag{3.105}$$

$$V_2(d_2(i, i')|f_i, f_{i'}) = \begin{cases} \sum_{k=1}^{K_2} \frac{[d_2^k(i,i') - D_1^k(f_i, f_{i'})]^2}{2[\sigma_2^k]^2} & \text{si } i' \neq i \ \& \ f_i \neq 0 \ \& \ f_{i'} \neq 0 \\ 0 & \text{en otro caso} \end{cases} \tag{3.106}$$

donde $[\sigma_l^m]^2$ es la varianza de una componente de ruido, los vectores $D_1(f_i)$ y $D_2(f_i, f_{i'})$ son los vectores media condicionados por f_i y $f_{i'}$, de los vectores aleatorios $d_1(i)$ y $d_2(i, i')$ respectivamente, condicionados por la configuración f.

Usando $U(f|d) = U(f) + U(d|f)$, se obtiene la expresión de la energía a posteriori

$$\begin{aligned} U(f|d) = & \sum_{i \in \mathcal{S}} V_1(f_i) + \\ & \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{N}_i} V_2(f_i, f_{i'}) + \\ & \sum_{i \in \mathcal{S}: f_i \neq 0} V_1(d_1(i)|f_i) + \\ & \sum_{i \in \mathcal{S}: f_i \neq 0} \sum_{i' \in \mathcal{S} - \{i\}: f_{i'} \neq 0} V_2(d_2(i, i')|f_i, f_{i'}) \end{aligned} \tag{3.107}$$

De todos los parámetros (varianzas del ruido $[\sigma_l^m]^2$ y penalizaciones a priori v_{n0}), sólo importan sus valores relativos ya que la solución no cambia si se multiplica la energía por un factor.

Algoritmo de Temple Simulado (Simulated Annealing)

El algoritmo de Temple Simulado introducido por Kirkpatrick y col. [1983] es un algoritmo estocástico para optimización combinacional o multivariable. Simula el proceso

físico en el cual se funde una sustancia y se enfría lentamente en busca de una configuración de baja energía. Para encontrar la siguiente configuración, en cada paso se usa un método aleatorio tal como el algoritmo de Metropolis, o el muestreador de Gibbs [Geman y Geman, 1984]. Esta búsqueda aleatoria se controla mediante una secuencia de temperaturas decrecientes. A T alta se pueden aceptar grandes incrementos en la energía (E); a T más baja sólo se pueden aceptar pequeños incrementos, y cerca de la T de congelación no se aceptan incrementos de la energía. Esto posibilita que el algoritmo escape de los mínimos locales.

El algoritmo de Metropolis genera una secuencia de configuraciones, para una temperatura T, usando un procedimiento de Monte Carlo, del modo que se indica a continuación:

```
inicializa f;
repite
        genera f' ∈ N(f);
        ΔE ← E(f') − E(f);
        P = min{1, e^{−ΔE/T}};
        si Randon[0,1) < P entonces f ← f';
hasta que (se alcance equilibrio);
devolver f;
```

Se propuso originalmente para simular el comportamiento de un sistema de partículas en equilibrio térmico a una temperatura T. En cada paso se escoge la siguiente configuración $f^{'}$ de forma aleatoria de $\mathcal{N}(f)$ aplicando una perturbación a f. $Random[0,1)$ es un número aleatorio de distribución uniforme en $[0,1)$. Si es más pequeño que P se acepta el cambio de configuración. Por tanto, se aceptará siempre la nueva configuración $f^{'}$ para la que $\Delta E \leq 0$, mientras que aquella para la que $\Delta E > 0$ se acepta con probabilidad $P = e^{-\Delta E/T}$.

Considérese un sistema en el cual cualquier f del espacio de configuraciones $\mathcal{F}$ tiene probabilidad

$$P_T(f) = [P(f)]^{1/T} \tag{3.108}$$

donde $T > 0$ es la temperatura del sistema. En el límite $T \to \infty$, $P_T(f)$ es una distribución uniforme en $\mathcal{F}$; para $T = 1$, $P_T(f) = P(f)$; cuando $T \to 0$, $P_T(f)$ se concentra en los picos

de $P(f)$. Lo mismo sucede con $P(f|d)$. El algoritmo de Temple Simulado aplica una algoritmo de muestreo, como el de Metropolis, en valores sucesivos y decrecientes de T:

```
inicializa T y f;
repite
            muestrear aleatoriamente f de N(f) bajo T;
            decrementar T;
hasta que (T → 0);
devolver f;
```

Inicialmente, T es muy alto y f es una configuración aleatoria. A una T fija, el muestreo es acorde a una distribución de Gibbs. Después de la convergencia del algoritmo de Metropolis a la temperatura actual T, ésta se decrementa y se repite el proceso. La iteración del algoritmo termina cuando T está próximo a 0, y el sistema está congelado cerca del mínimo de $E(f)$.

Geman y Geman [1984] probaron dos teoremas de convergencia. El primero hace referencia al algoritmo de Metropolis y dice que si cada configuración se visita con frecuencia infinita, la distribución de las configuraciones generadas convergen a una distribución de Gibbs. El segundo es acerca del Temple Simulado. Este establece que si la secuencia de temperaturas decrecientes satisface

$$\lim_{t\to\infty} T^{(t)} = 0, \quad \text{y} \tag{3.109}$$

$$T^{(t)} \geq \frac{m \times \Delta}{\ln(1+t)} \tag{3.110}$$

donde $\Delta = \max_f E(f) - \min E(f)$, el sistema converge al mínimo global independientemente de la configuración inicial.

Estas condiciones son suficientes, pero no necesarias para la convergencia. En la práctica se usan aproximaciones. Kirkpatrick y col. [1983] escogieron $T^{(t)} = \kappa T^{(t-1)}$, donde κ toma valores entre 0.8 y 0.9. La temperatura inicial se hace suficientemente grande para que todos los cambios de configuración sean aceptados. A cada temperatura se intentan un suficiente número de configuraciones, que serán o bien $10m$ transiciones aceptadas, siendo m el número de sitios, o el número de pruebas excede $100m$. El sistema

se congela y el algoritmo termina si el número de aceptaciones deseados no se alcanza en tres temperaturas sucesivas.

Integración mediante MRF

Como formalismo para el esquema de integración empleamos los Markov Random Fields (MRF), los cuales están recibiendo especial atención como esquemas de integración de ligaduras para la interpretación de imágenes [Modestino y Zhang, 1992; Kim y Yang, 1996]. Aunque la mayoría de los trabajos relacionados con el modelo MRF tratan con estructuras de las imágenes a nivel de pixel como sitios a etiquetar, se puede alcanzar la extensión natural a estructuras de las imágenes con un nivel más abstracto definiendo un modelo MRF sobre regiones.

El sistema de integración considerará como zonas óseas seguras aquellas partes de la imagen original que fueron calificadas como hueso a la vez por el sistema basado en regiones y el sistema basado en bordes. También considerará como zona de fondo segura la que en ambos casos recibió esa misma clasificación. El sistema de integración no modificará la consideración de tales regiones; trabajará sólo sobre las que han recibido distinta clasificación en cada sistema de segmentación. Las presentes en la máscara del crecimiento de regiones se volverán a dividir en las componentes resultantes de la presegmentación. El esquema basado en los MRF's decidirá sobre la incorporación de estas regiones al hueso o al fondo. En la figura 3.47 se muestra la entrada al sistema de integración para las dos segmentaciones de la figura 3.43. El nivel 255 representa zona ósea segura, el 0 representa fondo, y los niveles intermedios representan regiones a etiquetar. Las regiones consideradas como hueso seguro recibirán como etiqueta el valor 2 para tibia y el valor 3 para peroné. Las regiones de fondo seguro recibirán la etiqueta con valor 0. Al resto de regiones se les asignará inicialmente una etiqueta de pertenencia a hueso o a fondo de forma aleatoria. El etiquetado óptimo del conjunto de regiones se estima a través de una función de energía unificada.

En nuestra aplicación, los sitios corresponderán al conjunto de N regiones $\mathcal{R} = \{R_1, \ldots, R_N\}$ antes comentado. A partir de ellas se define un grafo $\mathcal{G} = \{\mathcal{R}, \mathcal{E}, \mathcal{D}\}$, donde $\mathcal{E}$ representa el conjunto de bordes que conectan las regiones, y $\mathcal{D} = \{\lceil_1, \lceil_2\}$ representa el conjunto de propiedades: de la propia región ($\lceil_1$), o de relación con regiones vecinas($\lceil_2$). Sobre este grafo se define un sistema de vecindad $\mathcal{N} = \{N_i | \forall i \in \mathcal{R}\}$, donde N_i es el

Figura 3.47: Mapa de regiones de entrada al sistema de integración extraida de las imágenes de la figura 3.43. El sistema decidirá sobre el etiquetado de las regiones con nivel de gris mayor que 0 y menor que 255.

conjunto de vecinos de R_i. Dos sitios (regiones) son vecinos si comparten algún borde. El conjunto de etiquetas será $\mathcal{L} = \{3, 2, 1, 0, -1\} = \{$tibia, peroné, probable hueso, fondo, probable fondo$\}$. El problema de integración se formula como el problema de asignar una etiqueta $\mathcal{L}$ a cada uno de los sitios $\mathcal{R}$.

$$\mathcal{I} : \mathcal{R} \rightarrow \mathcal{L} \tag{3.111}$$

o, equivalentemente, como un conjunto de variables aleatorias $\mathcal{I} = \{I_1, \ldots, I_N\}$ definidas sobre $\mathcal{R}$. $\mathcal{I}$ es el campo aleatorio, e $I_i \in \{3, 2, 1, 0, -1\}$ es la variable aleatoria asociada con R_i. $\mathcal{I}$ será el MRF sobre $\mathcal{G}$ con respecto al sistema de vecindad $\mathcal{N}$.

La segmentación (configuración) óptima se evaluará sobre la función de energía que se define a continuación:

$$U(I) = \sum_{i \in \mathcal{R}} V_1(I_i) + \sum_{i \in \mathcal{R}} \sum_{i' \in \mathcal{N}_i} V_2(I_i, I_{i'}) \tag{3.112}$$

$$+ \sum_{i \in \mathcal{R}} V_1(d_1 | I_i) + \sum_{i \in \mathcal{R}} \sum_{i' \in \mathcal{N}_i} V_2(d_2(i, i') | I_i, I_{i'}) \tag{3.113}$$

donde V_1 hace referencia a funciones de energía en la que sólo se tienen en cuenta etiquetas o propiedades (d_1) de regiones individuales, y V_2 representa las funciones de energía dependientes de la relación entre dos regiones entre etiquetas u otro tipo de propiedades (d_2).

Tabla 3.4: DEFINICIÓN DE PROPIEDADES: (a) propiedades unarias de las regiones, (b) propiedades de relación entre regiones.

(a)

Propiedad (d_1)	Definición
Área	A_i = *Tamaño de la región* R_i
Perímetro	P_i = *Longitud del perímetro de la región* R_i
Nivel de gris medio	$\rho = \frac{1}{A_i} \sum_{(x,y) \in R_i} I(x,y)$
Varianza en los niveles de gris	$\sigma = [\frac{1}{A_i} \sum_{(x,y) \in R_i} (I(x,y) - \rho_i)^2]^{1/2}$

(b)

Propiedad (d_2)	Definición
Adyacencia al fondo	$Is\_Ady\_B(i) = \begin{cases} 1 & \text{si } \exists i' \in N_i \text{ tal que } I_{i'} = 0 \\ 0 & \text{otro caso} \end{cases}$
Conectar estructuras diferentes	$Con\_D\_St(i) = \begin{cases} 1 & \text{si } \exists j,k \in N_i \text{ tal que } I_j = 2 \text{ y } I_k = 3 \\ 0 & \text{otro caso} \end{cases}$
Perímetro compartido	Per_{ij} = *Perímetro compartido entre las regiones* R_i *y* R_j
Número de vecinos	NN = *Número de regiones adyacentes*

En nuestra aplicación las etiquetas con valor 0, 2 y 3 no se cambiarán. Se incluirán en el grafo de adyacencias de regiones y sólo se utilizarán para evaluar ciertas propiedades de relación de las regiones a analizar, como se indicará más adelante.

Los potenciales los definimos de tal forma que se evite la conexión entre estructuras diferentes (tibia y peroné), se favorezca la forma regular del hueso y se tienda a etiquetar como fondo aquellas regiones alejadas de zona ósea segura, ya que estas regiones fueron generadas por el crecimiento de regiones, y aquí las zonas de transición aparecen representadas por una serie de franjas paralelas y estrechas. También habrá que eliminar huecos dentro del hueso, favoreciendo la etiqueta de hueso si todos sus vecinos la tienen. Y también se tendrá en cuenta el nivel de gris para guiar la integración. Las propiedades (d_1) de las regiones que aquí se emplean son: perímetro (P), nivel de gris medio (ρ) y su varianza (σ). Las propiedades de relación que se utilizan son: ser adyacente al fondo (Is_Ady_B $\in \{1,0\}$), conectar estructuras diferentes (Con_D_St $\in \{1.0\}$), perímetro compartido (Per_i) con cada vecino ($R_i \in N_i$), y número de vecinos (NN). Estas carac-

terísticas se representan en la tabla 3.4. Los potenciales de clique deben representar las ligaduras antes comentadas. Nosotros hemos construido una función de energía con las siguientes expresiones para los potenciales de clique:

$$V_1(I_i) = 0 \tag{3.114}$$

$$V_2(I_i, I_{i'}) = 0 \tag{3.115}$$

$$V_1(d_1(i)|I_i) = \Lambda I_i \mathrm{Con\_D\_St}(i) - \lambda I_i \frac{\rho_i + \sqrt{\sigma_i} - \rho_b}{\rho_b} \tag{3.116}$$

$$V_2(d_2(i,i')|I_i, I_{i'}) = \chi I_i \Theta_U(\frac{(I_{i'}+1)\mathrm{Per}_{ii'}}{\mathrm{P}_i} + \frac{\epsilon_1}{NN}) \tag{3.117}$$

$$- \psi I_i \mathrm{Is\_Ady\_B}(i) \Theta_{\epsilon_2}(\frac{(I_{i'}-1)\mathrm{Per}_{ii'}}{\mathrm{P}_i} + \frac{\epsilon_3}{NN}) \tag{3.118}$$

$$- \xi I_i(\frac{(I_{i'}+1)\mathrm{Per}_{ii'}}{2\mathrm{P}_i} - P_{th}) \tag{3.119}$$

La función Θ_U es la función de Heavisade unitaria centrada en el cero; es decir, devuelve valor 1 para números mayores que cero y 0 para los demás. La función $\Theta_{\epsilon_2}(\theta)$ devolverá 1 si el número es mayor o igual que ϵ_2 y 0 en otro caso. $\Lambda, \lambda, \chi, \psi, \xi$ son constantes positivas. El término de Λ pretende evitar la unión entre estructuras diferentes. El término en λ favorece la inclusión como parte del hueso de las regiones con densidad mayor que un umbral ρ_b, y como parte del fondo en caso contrario. El término en χ favorece la inclusión como parte del hueso de regiones en la que la mayor parte de sus vecinos tienen esta etiqueta. El término en ψ tiende a eliminar de la máscara de hueso aquellas regiones adyacentes al fondo y sin vecinos con etiquetas de hueso o con los que comparte muy poco contorno. Finalmente, el término en ξ tiende a eliminar concavidades cuya adyacencia con el hueso supere un umbral Per_{th}. En la Tabla 3.5 se especifica el significado de cada término de energía y los valores dados a a todos los parámetros del sistema.

Para la optimización del etiquetado se utilizará el criterio MAP, que se implementará mediante el algoritmo de Temple Simulado. Las regiones que tras el proceso de optimización tengan valor 0 ó -1 serán consideradas fondo, y las que tengan un valor positivo serán consideradas finalmente hueso.

El proceso de integración se efectúa para mejorar la localización de los contornos exteriores de las estructuras de interés. Para ello se parte de las segmentaciones proporcionadas por el sistema de crecimiento de regiones y el sistema basado en contornos deformables y se integra la información. En la figura 3.48.a se muestra el resultado de la integración

Tabla 3.5: DEFINICIÓN DE: (a) términos de energía, (b) funciones, (c) valores de parámetros asumidos en nuestra implementación.

(a)

Pesos	Significado de Términos Ponderados	Grado de Magnitud
Λ	Evitar conexión entre regiones de estructuras diferentes	∞
λ	Favorecer etiqueta de hueso para densidades altas, y de fondo para densidades bajas	1.0
χ	Favorecer homogeneidad del etiquetado	1.0
ψ	Favorecer incorporación al fondo de regiones alejadas de regiones óseas	1.0
ξ	Evita la formación de concavidades en el hueso	1.0

(b)

Función	Definición
Θ_U	$\Theta_U(\theta) = \begin{cases} 1 & \text{si } \theta \geq 1 \\ 0 & \text{otro caso} \end{cases}$
Θ_{ϵ_2}	$\Theta_{\epsilon_2}(\theta) = \begin{cases} 1 & \text{si } \theta \geq \epsilon_2 \\ 0 & \text{otro caso} \end{cases}$

(c)

Parámetro	Valor
ϵ_1	0
ϵ_2	0.1
ϵ_3	0
P_{th}	0.6

para las segmentaciones de la figura 3.39. En este caso, la segmentación por unión y división no había conseguido separar tibia de peroné, cosa que si se había logrado mediante contornos activos. Sin embargo, estos últimos por un efecto de suavización excesiva, no se ajustan correctamente al contorno. En el proceso de integración se consigue una mejor delimitación de las estructuras óseas. Otro ejemplo de la mejora introducida por nuestro sistema de integración es el que se muestra en la figura 3.48.b para las segmentaciones de la figura 3.43. En este caso la integración no difiere sustancialmente de la obtenida mediante métodos heurísticos y que se mostró en la figura 3.45.

La figura 3.49 muestra un ejemplo donde se ve claramente el mejor comportamiento de la integración basada en el modelo MRF respecto del método heurístico. Se consigue integrar la parte superior del peroné, que en el otro caso (ver figura 3.46.d) se perdía.

En partes alejadas de la parte proximal de rodilla, donde la distancia entre cortes es pequeña y el hueso cortical es ancho y denso, la integración provoca un leve ajuste de localización de contornos. A medida que nos acercamos a parte proximal se reduce

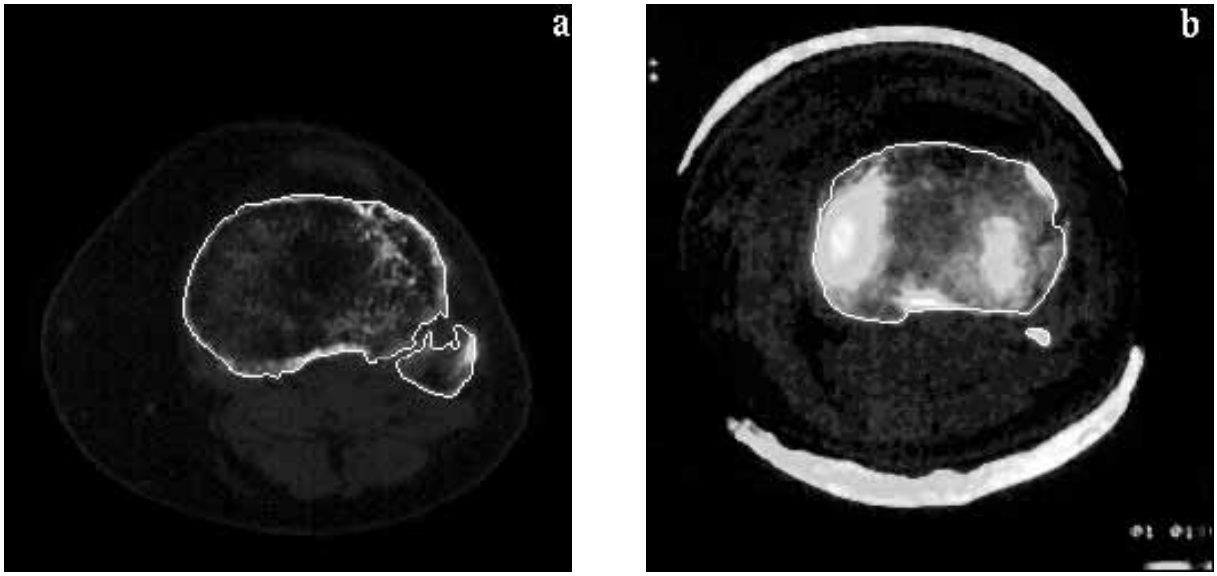


Figura 3.48: Integración mediante MRF: (a) para el caso de la figura 3.39, (b) para el caso de la figura 3.43.

la distancia entre tibia y peroné, disminuye el ancho del hueso cortical y aparecen las lesiones. En esta parte la segmentación se complica y la integración provoca mejoras más notables, como ya observamos en la figura 3.44. En la figura 3.50 podemos ver un nuevo ejemplo de la aplicación del sistema de integración en zona proximal, en el que se muestran los resultados con ambos esquemas. Las entradas al sistema son los resultados de las segmentaciones mostradas en la figura 3.42. Allí, el sistema basado en regiones ofrece un resultado adecuado después de haber efectuado un proceso de división entre las dos estructuras. Como este proceso de división no siempre ofrece una buena localización de contornos, siempre se recurre a la integración si el sistema basado en regiones tuvo que separar estructuras mediante información de bordes. La figura 3.50.a muestra el resultado para la integración heurística, y la figura 3.50.b muestra el resultado ofrecido por el esquema basado en los Campos Aleatorios de Markov. En este caso los resultados son análogos.

En la figura 3.51 se muestra un ejemplo de integración para un punto de bifurcación de la estructura de tibia. En la figura 3.51.b se muestra el resultado del crecimiento de regiones que no logra separar las dos bifurcaciones. La aplicación de los modelos

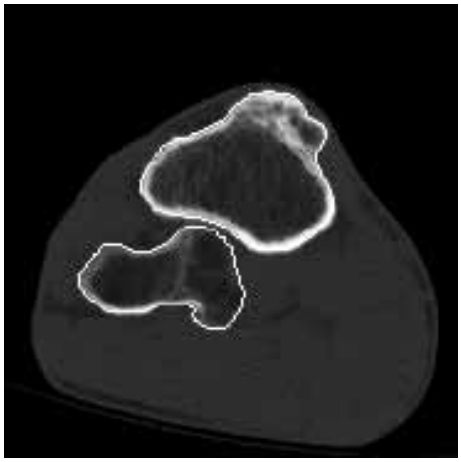


Figura 3.49: Integración mediante MRF para las segmentaciones de la figura 3.46.

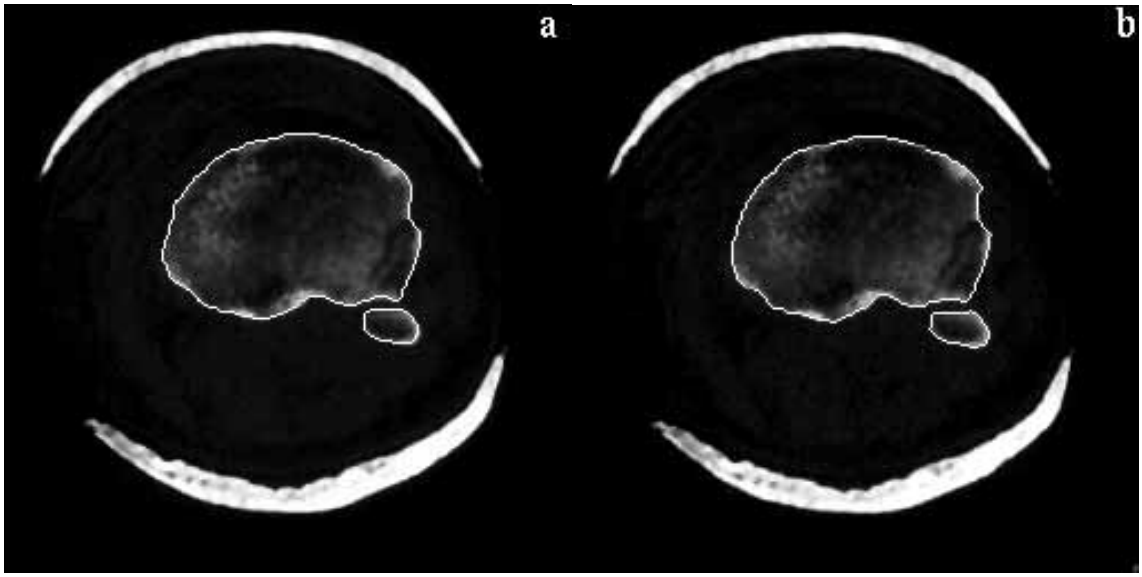


Figura 3.50: Integraciones para las segmentaciones de la figura 3.42: (a) heurística, (b) MRF.

deformables logra separar las dos estructuras al tiempo que proporciona una buena delimitación del hueso, como se observa en la figura 3.51.c. La figura 3.51.d muestra el resultado de la integración que se asemeja mucho a la segmentación obtenida mediante modelos deformables. La figura 3.52 muestra el resultado de la integración heurística

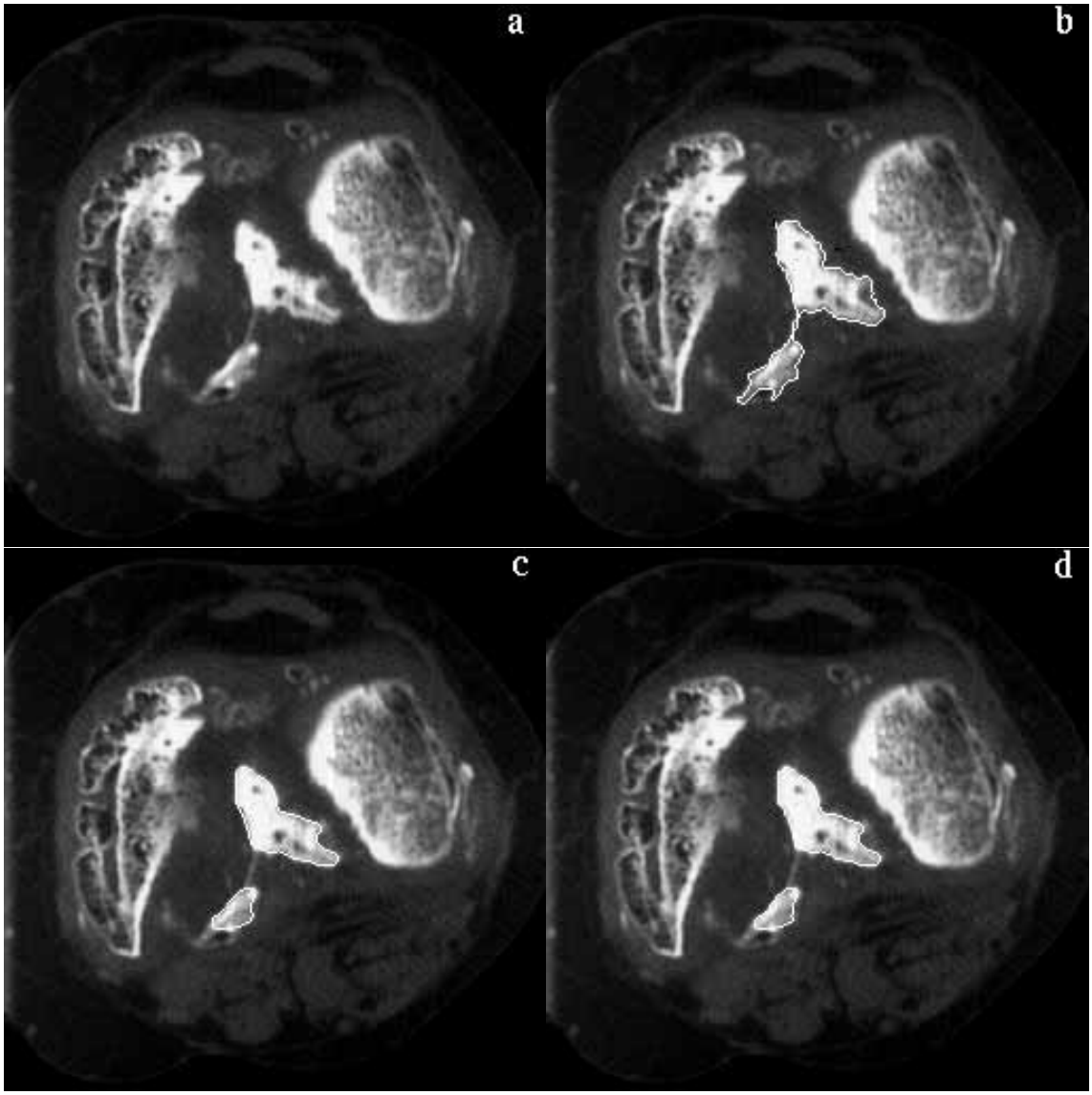


Figura 3.51: (a) Imagen original, (b) segmentación por crecimiento de regiones, (c) segmentación basada en modelos deformables, (d) integración.

para las mismas segmentaciones, ofreciendo, en general, un peor comportamiento que la integración anterior.

Como ya hemos comentado con anterioridad, tanto el método basado en crecimiento de regiones, como el basado en contornos deformables producen buenas segmentaciones en

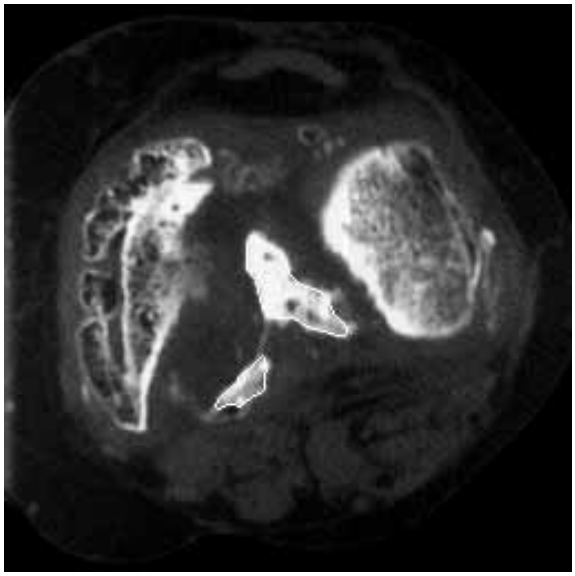


Figura 3.52: Integración de las segmentaciones de la figura 3.51 mediante el método heurístico.

la mayoría de los casos. Sin embargo, ninguno de estos métodos resulta invulnerable. Con nuestro sistema de integración superamos los inconvenientes presentados por los módulos individuales. De esta forma obtenemos un método de segmentación robusto. A su vez hemos probado dos esquemas de integración, uno heurístico y otro basado en los Campos Aleatorios de Markov. El primero resulta más rápido, pero el segundo es más robusto y presenta un esquema más formal que lo hace más fácilmente adaptable y extensible a nuevas situaciones.

Como resultado final del proceso de integración obtendremos un conjunto de contornos que delimitan las estructuras de interés en cada uno de los cortes.

3.6.3 Ejemplos

A lo largo de la exposición realizada hemos ido introduciendo distintas imágenes sobre las que se han mostrado las ventajas y los inconvenientes de los métodos de segmentación por separado, y como la integración final de las segmentaciones consigue eliminar los errores que se pudiesen presentar de forma individual. En las secciones más alejadas del extremo proximal de la tibia la segmentación se resuelve mediante crecimiento de regiones

y posterior ajuste de las localizaciones de estos contornos mediante modelos deformables. A medida que avanzamos en la secuencia de cortes hacia el extremo proximal se van acercando tibia y peroné y van apareciendo problemas de pérdida ósea y lesiones. En esta parte es necesaria la integración de las segmentaciones obtenidas por separado mediante crecimiento de regiones y contornos activos, para la obtención de una segmentación más precisa.

Las numerosas características indeterminadas de las imágenes reales hacen que resulte complicada una evaluación analítica de las segmentaciones obtenidas [Zhang, 1997]. Por lo que como regla general, los resultados de las segmentaciones sólo pueden ser juzgados o bien usando una segmentación manual como referencia o mediante la inspección visual [Pal y Pal, 1993; Zhang, 1997]. Muchos autores consideran estas evaluaciones más subjetivas que objetivas, por lo que proponen realizar las evaluaciones de los métodos sobre imágenes sintéticas con segmentación verdadera conocida [Zhang, 1997]. Otros autores, Hoover y col. [1996] entre ellos, consideran que la valoración del rendimiento de un algoritmo ha de hacerse sobre imágenes reales; consideran que las valoraciones hechas sobre imágenes sintéticas inspiran poca confianza. Por otra parte, indica también que el establecimiento de la segmentación real requiere de ciertas dosis de imaginación y resulta en general muy laboriosa. Como el establecimiento de la segmentación verdadera basada en la intensidad resulta más subjetiva que la que se basa en la geometría, decide hacer las valoraciones de los algoritmos sobre imágenes reales con objetos de geometría simple. A parte de este desacuerdo en qué tipo de imágenes son las más adecuadas, tampoco se han establecido métricas de error estándar. Esto ha llevado a que la evaluación visual haya evolucionado como la norma de actuación [Hoover y col., 1996].

Como ya hemos comentado al hablar de la valoración de los resultados proporcionados por el método de análisis de regiones, cuando se trata con imágenes adquiridas a partir de sujetos vivos no se puede establecer la segmentación ideal (real). En este caso, la valoración de los resultados del método ha de hacerse mediante inspección visual por parte del médico. De esta inspección se concluye que el método de segmentación automático que nosotros proponemos proporciona una segmentación válida. Por otra parte, si definimos la velocidad de la segmentación como el número de cortes de la misma secuencia procesados por minuto, los valores obtenidos están en el rango de 1 a 1.5 minutos por corte. Estos tiempos mejoran los de cualquier segmentación semiautomática. El sistema ofrece ventajas definitivas respecto de la segmentación semiautomática al proporcionar

resultados precisos, ser repetible y ahorrar cantidades de tiempo significativas.

Tal y como indicamos en la introducción, este proceso es necesario para obtener una reconstrucción 3D de la geometría del hueso que sirva como modelo de entrada a procesos tales como el análisis mediante elementos finitos de la distribución de tensiones en la interfaz de hueso y prótesis. Por otra parte, el sistema desarrollado proporciona al médico una herramienta útil para la visualización 3D de la estructura ósea. De esta forma se obtiene una representación volumétrica de la información que sirve de apoyo al proceso de diagnóstico al permitir la visualización directamente en 3D de información que de otra forma tendría que ser organizar de forma mental a partir de imágenes 2D. Además de esta visualización, se puede llevar a cabo la medición de parámetros volumétricos. En el capítulo final comentaremos el método de generación de las superficies 3D del hueso a partir del apilamiento de los contornos 2D que se obtienen como resultado de la segmentación.

Para ilustrar el tipo de resultados obtenidos, mostraremos a continuación los contornos de las segmentaciones obtenidas tras procesar distintos conjuntos de cortes de tomografía computerizada (CT) de la parte proximal de la tibia correspondientes a diferentes pacientes.

La figura 3.53 muestra el primer ejemplo; corresponde al procesado de una secuencia de 11 cortes contiguos de CT de la tibia proximal de un paciente que había sufrido una fractura de la meseta tibial un año antes. La resolución inicial del conjunto de imágenes era de $0.6 \times 0.46 \times 3\ mm$. Esta baja resolución en el eje Z hace que aparezcan variaciones grandes en forma, tamaño y, en ocasiones, de posición de centro de masas entre las estructuras óseas a segmentar de unos slices a otros. Para evitar esos problemas se generaron mediante interpolación 5 cortes intermedios entre cada par de ellos para obtener un volumen isótropo. En la figura se muestran los contornos resultantes de procesar todos los cortes de la secuencia. Aún cuando hasta ahora no hemos planteado el problema de identificar las zonas de lesión, en esa imagen, a parte de los contornos externos de tibia y peroné, se incluyen también los contornos internos de la tibia que delimitan la fractura que sufrió el paciente. Hasta ahora no hemos tratado el problema de la detección de lesiones, será en el próximo capítulo donde describiremos el proceso de detección de zonas de lesión. Si la lesión es interna, se obtendrá un conjunto de contornos adicional que la delimiten, y si es lateral se modificarán los contornos externos obtenidos previamente para poner de manifiesto su presencia.

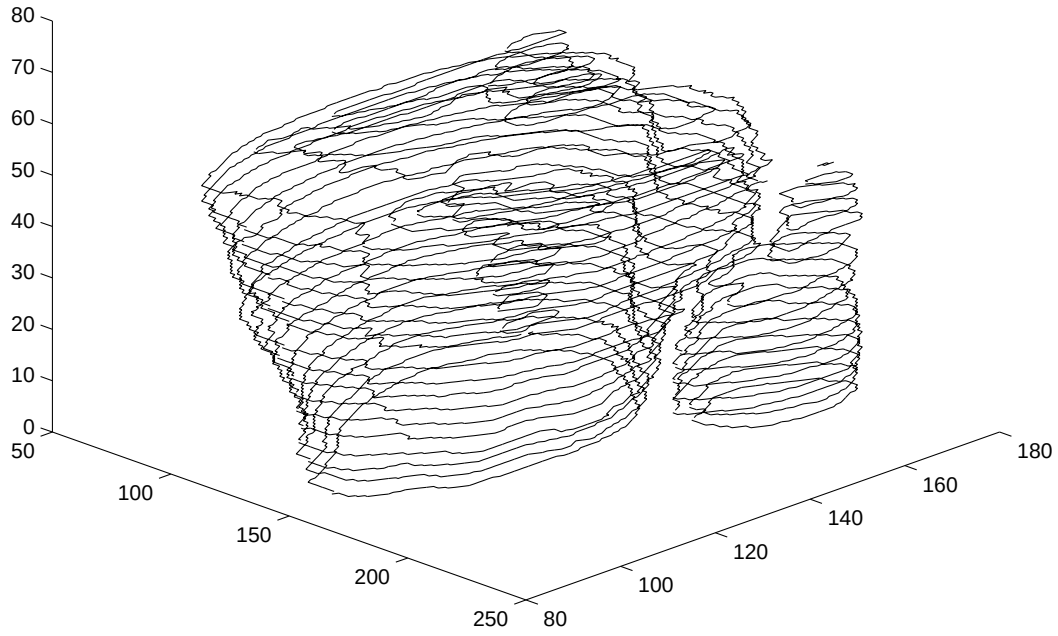


Figura 3.53: Apilamiento de contornos de tibia, peroné y lesión interna, por hundimiento de la meseta tibial, resultantes de la segmentación de una secuencia de 76 cortes de CT.

La figura 3.54 muestra la secuencia de contornos apilados de la tibia izquierda de un nuevo paciente para el que se dispuso de una secuencia de 24 cortes. En este caso el paciente presenta una fractura de la meseta tibial externa. El tamaño del pixel es, ahora, de 0.71 mm^2, y la distancia entre cortes es de 1 mm.

En la figura 3.55 se muestra el resultado de procesar una secuencia de 45 cortes de otro paciente, que sufre artrosis de rodilla. La resolución espacial en la imagen es de $0.8 \times 0.8 \times 1$ mm.

Finalmente, la figura 3.56 muestra los resultados correspondientes al procesado de una secuencia de imágenes obtenida a partir de la base de datos del Mallinckrodt Institute of Radiology. Según aparece en la documentación de dicha base de datos, las imágenes

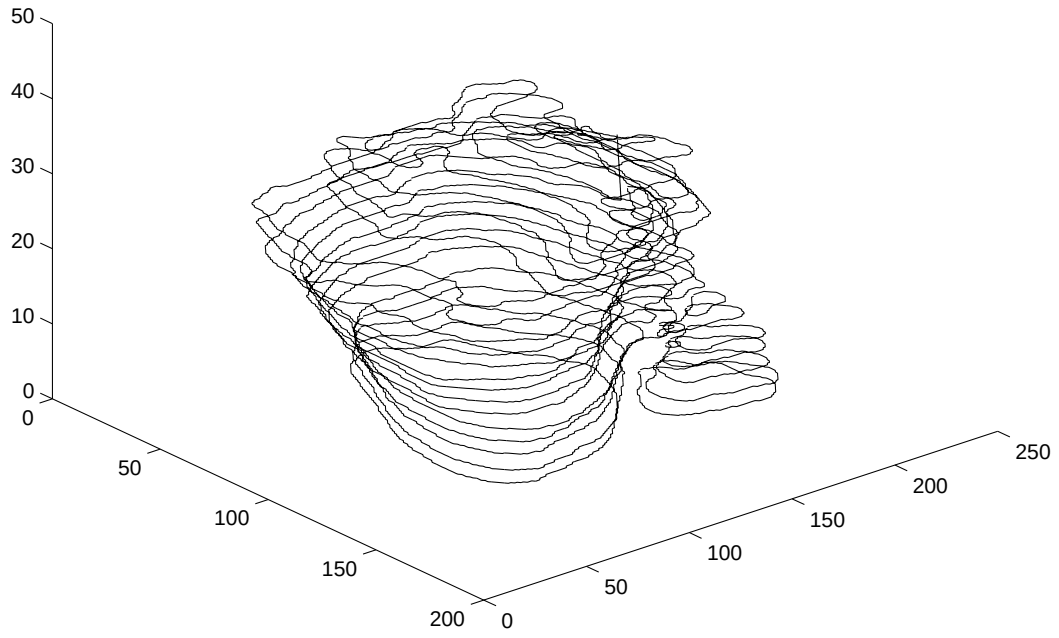


Figura 3.54: Apilamiento de contornos de tibia y peroné resultantes de la segmentación de una secuencia de 24 cortes de CT de un paciente que presentaba una fractura lateral de tibia.

iniciales fueron interpoladas y escaladas para obtener una resolución isótropa de 8-bits, obteniéndose un conjunto de 180 cortes. La secuencia original presentaba las siguientes dimensiones de voxel: $0.45 \times 0.45 \times 1.0$ pulgadas.

El apilamiento de los contornos, como ilustran la figuras 3.53-3.56, aún no define la superficie 3D de las estructuras óseas. En el capítulo 5 trataremos el problema de definir la superficie a partir de la secuencia de contornos. Por otra parte, el proceso de segmentación que hemos descrito hasta el momento aún no permite identificar lesiones en el hueso. Aquí hemos mostrado los resultados finales de la segmentación tanto para casos que necesitaban de un proceso de identificación de regiones como para los que no. En los capítulos siguientes abordaremos los problema de la detección de lesiones y de la definición de la superficie externa 3D a partir de los contornos 2D.

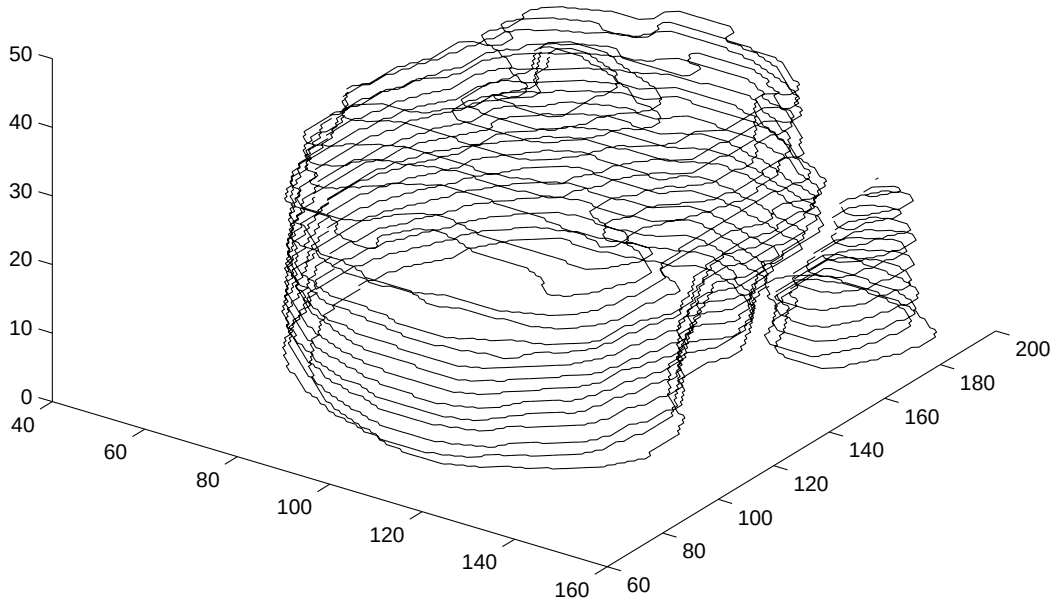


Figura 3.55: Apilamiento de contornos de tibia y peroné resultantes de la segmentación de una secuencia de 45 cortes de CT de la rodilla de un paciente con artrosis.

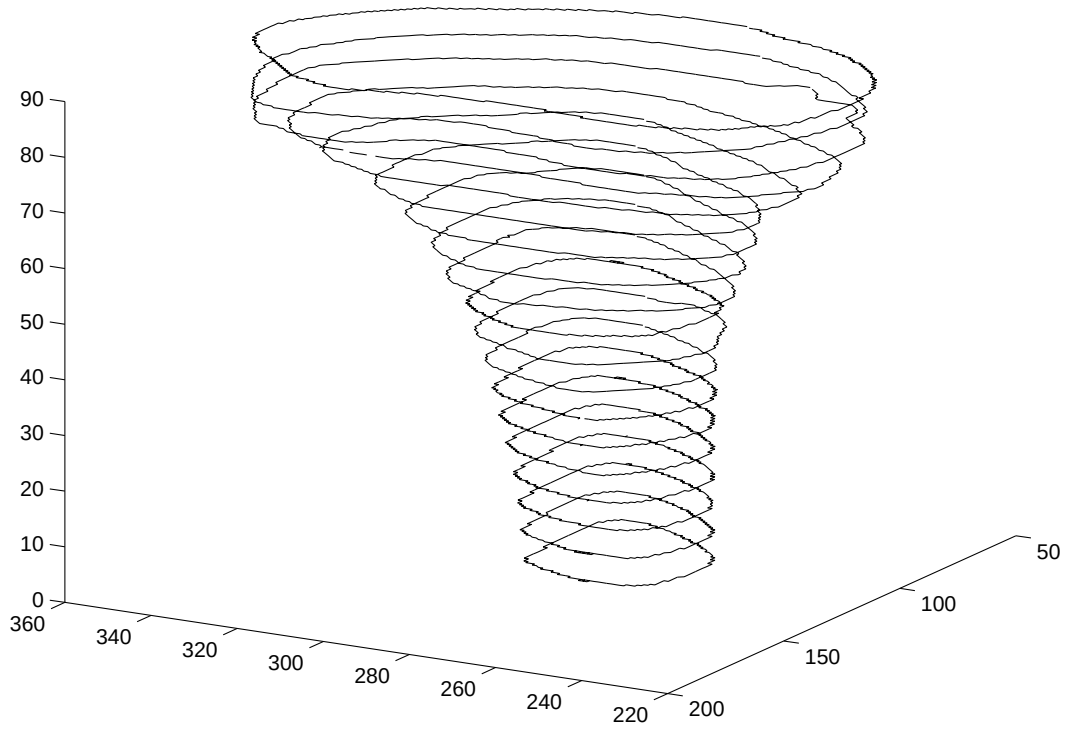


Figura 3.56: Apilamiento de contornos de tibia resultantes de la segmentación de una secuencia de 180 cortes de CT.

Capítulo 4

Detección de lesiones

4.1 Introducción

Una vez delimitados los contornos externos del hueso, el siguiente proceso que se ha de abordar es la identificación de las posibles regiones de lesión que se manifiesten como pérdida ósea. Para realizar esta tarea se parte de la presegmentación que se había efectuado inicialmente dentro del módulo de crecimiento de regiones, considerando únicamente las particiones interiores a las estructuras óseas finalmente identificadas. El resto de regiones serán parte del fondo de la imagen. Para la identificación de zonas de lesión se irá elaborando una hipótesis usando un esquema ascendente conducido por datos (*data-driven bottom-up*), que después se analizará y refinará usando un esquema descendente (*top-down*).

Para la clasificación de una región como lesión nos basamos en conocimiento a priori sobre las características generales de las regiones asociadas con lesión: nivel de gris (densidad) bajo, alta uniformidad, forma regular, carencia de huecos, área mayor de un umbral, continuidad 3D y fuerte contraste con sus regiones vecinas. Este conjunto de características permiten diferenciar entre lesión (pérdida ósea) y regiones de trabécula, que en zonas de baja densidad pueden ser fáciles de confundir al tratarse de imágenes ruidosas. La zona de trabécula presenta en general características distintas. Así, ésta se manifestará tras la segmentación bien mediante regiones grandes e irregulares con gran número de huecos, que corresponden a las zonas de trabécula más densa, bien mediante

regiones pequeñas, que corresponden a hueso canceloso (*cancellous bone*), y que presentarán en general gran variación de forma de un corte a otro, o finalmente, como regiones de tamaño medio, que corresponderán a zonas de trabécula menos densa, con vecinos de trabécula más densa, y por lo tanto un contraste medio-bajo, y que van a contener huecos y/o presentar formas muy irregulares. La consideración de una región como lesión afecta, y estará afectada por las consideraciones que se hagan sobre sus vecinos. Así, por ejemplo:

- La consideración de dos regiones vecinas como lesión puede llevar tras su unión a una región de mayor tamaño que no cumpla las condiciones antes citadas. Así, por ejemplo, podría resultar una región con huecos, lo cual indica más bien la presencia de hueso trabecular que de lesión.
- Características tales como la de regularidad de forma se podrían conseguir mediante la unión de regiones que por separado no las cumplen.
- El fuerte contraste entre dos regiones puede estar enmascarado por la existencia de regiones de transición, por lo general irregulares en forma y/o muy estrechas.

Como ya hemos comentado en el capítulo anterior, un esquema formal para la integración del conjunto de ligaduras y la solución del problema de interpretación lo proporciona el modelo de los Campos Aleatorios de Markov (MRF) [Geman y Geman, 1984; Modestino y Zhang, 1992; Kim y Yang, 1996]. Este modelo permite relacionar propiedades locales y globales, lo cual proporciona potenciales ventajas en la incorporación de conocimiento y en la optimización. En este capítulo hemos adaptado ese formalismo para implementar un sistema que identifique zonas de lesiones en la parte proximal de la tibia [Pardo y col., 1997b, 1998b]. Se comienza con el conjunto inicial de regiones segmentadas obtenidas a partir del módulo de sobresegmentación de bajo nivel descrito con anterioridad. La información correspondiente a esta sobresegmentación se organiza en un **grafo de adyacencias de regiones** (RAG), en el que se almacenan las propiedades individuales de cada región y mediante arcos se establecen las relaciones de vecindad. Estas relaciones se ponderan mediante la magnitud del contorno compartido. El proceso de identificación de lesiones se divide en dos partes. Se comienza con la identificación de posibles zonas de lesión en cada corte de la secuencia 3D, para lo que definimos un MRF y formulamos la resolución como un problema de estimación de la máxima probabilidad *a posteriori* (MAP). Finalmente, se hace un análisis 3D de las regiones candidatas

a lesión en base a su continuidad 3D, relación de tamaños, etc. Este último módulo lo implementamos como un sistema de producción que emplea un reducido conjunto de reglas.

Para eliminar regiones no significativas se realiza un proceso inicial de eliminación de regiones muy pequeñas (<10 pixels) para reducir el tiempo de procesado. No se deben eliminar regiones algo mayores porque eso implicaría eliminación de huecos significativos que pudiesen evidenciar la presencia de región trabecular y no lesión. Estas regiones pequeñas se unen con aquellos vecinos con los que presenten mayor similaridad. Esta similaridad se define del siguiente modo

$$S_{1,2} = \frac{1 + \max\{0, \min\{\rho_1 + \sigma_1, \rho_2 + \sigma_2\} + \max\{\rho_1 - \sigma_1, \rho_2 - \sigma_2\}\}}{1 + |\min\{\rho_1 - \sigma_1, \rho_2 - \sigma_2\} - \max\{\rho_1 + \sigma_1, \rho_2 + \sigma_2\}|} \tag{4.1}$$

De esta forma la similaridad será igual a 1 cuando valor medio y varianza coincidan, y disminuirá tendiendo a cero a medida que el solape sea menor y la distancia entre extremos sea mayor. El comportamiento de la medida de similaridad se ilustra en la figura 4.1. En esta figura cada uno de los segmentos representa el rango de intensidades de una región, comprendido entre su media menos su varianza y su media más la varianza. Cuanto mayor es el solape entre los segmentos asociados a dos regiones, mayor es su similaridad. Cuanto más alejados estén los segmentos, menor será la similaridad.

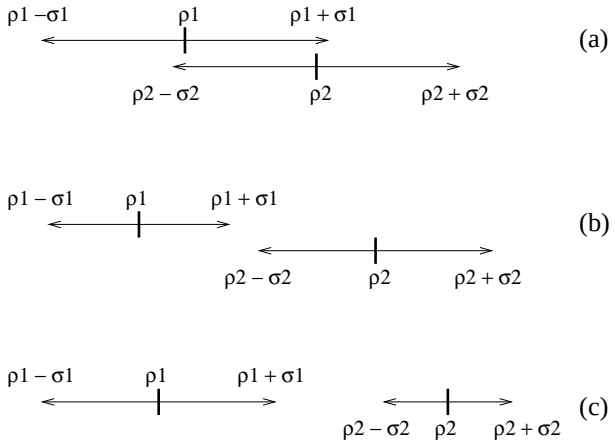


Figura 4.1: La similaridad disminuye de (a) a (c).

4.2 Etiquetado mediante MRF

En la identificación de lesiones se plantea un problema similar al comentado en la sección 3.6.2, en el que se trataba de integrar los resultados de los dos métodos de segmentación mediante un proceso de reetiquetado de regiones. Aquí vamos a usar los mismos formalismos para la identificación de lesiones, aunque variando el conjunto de etiquetas y las definiciones de los potenciales.

Sea un conjunto de r sitios (regiones), $\mathcal{R} = \{1, \ldots, r\}$, cada una de ellas con un conjunto de propiedades asociado y ligadas unas con otras a través de relaciones contextuales. Sea $\mathcal{L} = \{-1, 1\}$ el conjunto de etiquetas. Supongamos que la estructura de la escena se representa por $\mathcal{G} = (\mathcal{R}, d)$, y la del modelo por $\mathcal{G}^{'} = (\mathcal{L}, D)$, donde d y D hacen referencia a propiedades y relaciones entre las respectivas características.

Las características extraidas de la imagen, $d = \{d_1, d_2\}$, constan de dos fuentes de ligaduras: propiedades unarias d_1 de cada región, tal como área, perímetro, densidad media, varianza, segundos momentos centrales, etc., y relaciones binarias, d_2, tales como relaciones de adyacencia, hueco de, etc. El modelo consta a su vez de un conjunto de propiedades D, que incluyen ambos tipos de ligaduras.

Consideraremos que la observación d, bajo la configuración f, es una versión contaminada con ruido blanco Gaussiano n de las características D del modelo.

$$d_1(i) = D_1(f_i) + n(i), \quad d_2(i, i^{'}) = D_2(f_i, f_{i'}) + n(i, i^{'}) \tag{4.2}$$

por tanto, $d_1(i)$ y $d_2(i, i^{'})$ son distribuciones Gaussianas de medias $D_1(f_i)$ y $D_2(f_i, f_{i'})$ respectivamente.

La interpretación de la imagen se formula como un problema de estimación de máxima probabilidad a posteriori (MAP). En el modelado MRF esto se traduce en el problema de minimizar la función de energía definida para el problema. Como método de optimización utilizaremos el Temple Simulado, ya que determinará el mínimo global con un coste no excesivo.

Con potenciales de clique de hasta dos sitios, la energía a minimizar toma la forma

$$U(f|d) \quad = \quad \sum_{i \in \mathcal{R}} V_1(f_i) + \sum_{i \in \mathcal{R}} \sum_{i' \in \mathcal{N}_i} V_2(f_i, f_{i'}) + \tag{4.3}$$

$$\sum_{i\in\mathcal{R}} V_1(d_1|f_i) + \sum_{i\in\mathcal{R}}\sum_{i'\in\mathcal{N}_i} V_2(d_2(i,i')|f_i,f_{i'}) \tag{4.4}$$

donde $\mathcal{N}_i$ es el conjunto de vecinos de i, que se corresponden con las regiones adyacentes en el RAG. Esta expresión con cliques de hasta segundo orden es la forma más simple capaz de transportar información contextual. Los dos primeros términos hacen referencia a la probabilidad conjunta a priori, y los dos últimos hacen referencia a la probabilidad condicional de la observación d dada la etiqueta f.

4.2.1 Potenciales de clique

Para el caso de detección de lesiones en tibia el conjunto de etiquetas será de la forma $\mathcal{L} = \{$ lesión, hueso $\} = \{-1, 1\}$. La relación de vecindad vendrá ponderada por el contorno compartido. Las zonas de lesión dentro del hueso se caracterizarán por su baja densidad, homogeneidad y forma compacta, frente a la característica de alta densidad media, propia del hueso cortical, o alta varianza, propia del hueso trabecular. Por otra parte, está el problema en distinguir lesión de médula en zona trabecular. Nosotros hemos elegido los siguientes potenciales de clique para el modelo de MRF de forma que nos permitan identificar las zonas de lesión. Para la energía a priori tenemos que:

$$V_1(f_i) = \alpha_i \tag{4.5}$$

$$V_2(f_i,f_{i'}) = \begin{cases} \beta(f_i-1)f_{i'} + \beta_{ii'} & \text{si} \quad \begin{cases} ((f_i=-1,\ f_{i'}=1) \\ \ \& \ (N_{i,i'}=P_{i'} \mid N(i,i')=P_i)) \end{cases} \\ \beta(f_i-1)f_{i'} & \text{otro caso} \end{cases} \tag{4.6}$$

donde $N_{i,i'}$ denota el contorno compartido o grado de vecindad entre dos regiones y P_i representa el perímetro de la región i. La condición $(N_{i,i'} = P_{i'} \mid N_{i,i'} = P_i)$ quiere decir que una de las regiones es hueco de la otra. Con estos potenciales de clique se define la energía a priori sin más que hacer la suma sobre todos los cliques en $\mathcal{R}$. El potencial para cliques de primer orden penaliza la etiqueta con mayor valor relativo de α. El potencial para cliques de segundo orden penaliza regiones etiquetadas como lesiones que tengan huecos con cualquier etiqueta, y regiones etiquetadas como hueso que sean hueco de regiones etiquetadas como lesión. La penalización varía según la etiqueta del

hueco y de la región que lo contiene. Obviamente, una región etiquetada finalmente como lesión no podrá tener huecos que sean hueso, pero si huecos que sean lesión, ya que al final se unirían como una única región de lesión. Finalmente, se penaliza el etiquetado de lesiones contiguas.

Para la energía condicional tenemos las siguientes expresiones para las funciones de clique:

$$V_1(d_1(i)|f_i) = \begin{cases} \min\{1.0, \frac{[\rho_i+\sigma_i]^2}{2\sigma^2_{\rho\sigma}}\} & \text{si} \quad f_i = -1 \\ \min\{1.0, \frac{[\min\{0,(\rho_i+\sigma_i)-\rho_b\}]^2}{2\sigma^2_{\rho\sigma}}\} & \text{si} \quad f_i = 1 \end{cases} \tag{4.7}$$

$$V_2(d_2(i,i')|f_i, f_{i'}) = \begin{cases} \gamma(1-2f_i)\frac{N_{i,i'}[(\rho_i+\sigma_i)-(\rho_{i'}-\sigma_i)]}{P_i(\rho_i+\sigma_i)} & \text{si} \quad f_i \neq f_{i'} \\ 0 & \text{otro caso} \end{cases} \tag{4.8}$$

Estos potenciales definen la función de distribución de probabilidad condicional de los datos observados (características y relaciones entre regiones). A cada región se le asigna un potencial de clique de primer orden diferente dependiendo de la etiqueta que posea. Para el caso de lesión se penalizan las densidades y las varianzas altas. Para el caso de hueso la situación es más compleja, ya que los tejidos óseos pueden presentar o bien densidad alta y varianza baja (hueso cortical), o varianza alta y densidad media media–baja. Para considerar estos dos casos se define el potencial de clique como distancia al cuadrado entre valor medio (ρ_i) más varianza (σ_i) y (1) densidad nula en el caso de lesión, y (2) un valor alto típico de hueso (ρ_b) para el caso de etiqueta de hueso. Las dos medidas están acotadas entre 0 y 1 adoptando la forma del potencial de Blake y Zisserman [Li, 1995]. El potencial de clique de segundo orden hace referencia al contraste entre regiones. Afianza una región como lesión cuando es menos densa que sus vecinos de hueso y la penaliza en otro caso. Lo mismo ocurre con las etiquetas de hueso. El término está ponderado por la longitud del contorno compartido entre las dos regiones.

La función de distribución de probabilidad no queda completamente caracterizada dando sólo la forma de los potenciales de los cliques. Es preciso determinar también el valor de los parámetros que aparecen en las expresiones de dichos potenciales.

Los valores medios y las varianzas se pueden determinar de forma sencilla a partir de la imagen. Para determinar los valores medios y las varianzas en las expresiones de los potenciales de clique de primer orden para la función de distribución de probabilidad condicional de los datos observados hemos evaluado la densidad media $\overline{\rho}$ y la varianza

media $\overline{\sigma}$ en las regiones de tejido muscular adyacentes al hueso cortical. Escogemos $\rho_b = 3(\overline{\rho} + \overline{\sigma})$ y $\sigma^2_{\rho\sigma} = 4.5(\overline{\rho} + \overline{\sigma})^2$. Como lo que importa no son los valores absolutos de los parámetros, sino su valor relativo, consideraremos los anteriores como fijos, y nos quedarán por fijar los parámetros $\alpha_i, \beta, \beta_{ii'}, \gamma$. Como el aprendizaje de los parámetros es difícil de llevar a cabo, optamos por un ajuste manual de parámetros que permita reconocer por exceso un número de regiones de posible lesión. Más tarde, en la aplicación de la segunda etapa del etiquetado se discriminarán las lesiones verdaderas de las falsas mediante la aplicación de reglas relativas a continuidad 3D.

Esta etapa que comentamos parte de un etiquetado aleatorio de las regiones resultantes de la presegmentación y se le aplica un método de minimización global. Como resultado de este etiquetado resultará un mapa de regiones con etiquetas de lesión y hueso. Habrá regiones adyacentes etiquetadas como lesión, por lo que un paso inmediato consiste en unir estas regiones en una sola. El etiquetado obtenido en ese momento aún no será en general correcto debido a que muchas regiones pequeñas correspondientes a huecos en la región trabecular pudieron haber sido etiquetadas como lesiones. Además, ciertas regiones no pequeñas, pero de baja densidad, también pueden estar etiquetadas como lesión. En un filtrado inicial pueden eliminarse aquellas muy pequeñas (área < 40 pixels) y aquellas de forma muy irregular; por ejemplo, aquellos para los que el área de la elipse de dispersión sea mayor que 2 veces su área. Esta eliminación consiste en el cambio de etiqueta, que en el procesado posterior se podría recuperan si la información 3D así lo demandase. Se trata pues de un proceso revocable.

4.3 Análisis 3D

Una vez analizadas las imágenes 2D se obtiene para cada corte un conjunto de regiones con etiqueta de lesión y el resto con etiqueta de hueso. La baja relación señal ruido de las imágenes, así como una estructura trabecular poco densa, pueden dar lugar en muchas ocasiones a etiquetados erróneos. El etiquetado puede mejorarse si ahora analizamos las configuraciones resultantes mediante un sistema basado en reglas en el que se incluyan restricciones tales como continuidad 3D, bifurcaciones, relaciones de área entre secciones 2D de la misma lesión 3D, conexión de la lesión 3D con el exterior, etc.

El análisis 3D consta de tres fases que describimos a continuación. En la primera se

realiza un análisis de la continuidad 3D de las regiones 2D etiquetadas como lesión en la etapa anterior. Aquéllas que no presenten continuidad en los dos cortes vecinos se etiquetarán finalmente como hueso, mientras que las que tengan continuidad sólo en uno de sus vecinos recibirán una etiqueta nula. Sobre el etiquetado final de estas regiones se decide en la segunda fase, la cual trata de unir regiones volumétricas de lesión separadas como máximo por dos cortes en el eje longitudinal. Finalmente se realiza una suavización de la forma de las lesiones 3D.

Veamos el proceso de forma más detallada mediante la descripción de las reglas. Previamente presentaremos la notación usada. Así, el superíndice k en ellas se refiere al número del corte, $k+\nu$, con $\nu \in \{-1, 1\}$ se refiere a los dos cortes adyacentes, $N_{i^k, j^{(k+\nu)}}$ representa el solape entre la región i del corte k y la región j de uno de los cortes vecinos.

El conjunto ordenado de reglas se aplica de tal forma que inicialmente todos los cortes se exploran en la búsqueda de la continuidad de la lesión. Previamente, e independientemente de esta continuidad 3D, una región con una posible etiqueta de lesión se etiqueta como nula si su área A_i^k es menor que un umbral A_{TH}.

SI
(1) $f_i = -1$
(2) $A_i^k < A_{TH}$
ENTONCES
(1) $f_i = 0$

Esto se hace con el fin de eliminar volúmenes estrechos etiquetados como lesiones, pero que en realidad corresponden a agujeros de la estructura trabecular.

El conjunto de reglas que evalúan la continuidad es el siguiente. En primer lugar, si una lesión en un corte tiene continuidad simple (no bifurcación), entonces preserva su etiqueta. En otro caso, esa región y sus bifurcaciones reciben etiqueta nula. La regla que evalúa esta última condición es la siguiente:

SI
(1) $f_i^k = -1$
(2) $\exists\ n > 2\ j_n^{(k+\nu)}$ tal que $N_{i^k, j_n^{(k+\nu)}} > 0$
ENTONCES
(1) $f_i = 0$

(2) $f_{j_n}^{(k+\nu)} = 0, \quad \forall j_n \in$ corte $k+\nu$

Si una lesión sólo presenta continuidad en uno de los cortes adyacentes o si al menos en uno de esos cortes se bifurca en más de dos regiones, entonces a la región en el corte actual y a esas regiones en los cortes vecinos (continuidad múltiple) se les asigna etiqueta nula. La regla se formula en los siguientes términos:

SI

(1) $f_i^k = -1$

(2) $\exists$ 2 $j_n^{(k+\nu)}$ tal que $N_{i^k, j_n^{(k+\nu)}} > 0$ con $n = \{1, 2\}$

(3) $\nexists$ $p^{(k+\nu)}$ tal que $N_{p^{(k+\nu)}, j_n^{(k+\nu)}} > 0$ con $n = \{1, 2\}$

ENTONCES

(1) $f_i = 0$

(2) $f_{j_n}^{(k+\nu)} = 0, \quad n = \{1, 2\}$

Si una región etiquetada como lesión, tras la aplicación del método de identificación ya descrito, presenta continuidad en dos cortes pero se bifurca en dos en uno de ellos y no hay región de contacto tal que la unión de esas tres regiones en el corte vecino tenga un área menor que α veces el área de la región en el corte actual, entonces a todas ellas se les asigna etiqueta nula.

SI

(1) $f_i^k = -1$

(2) $\exists$ 2 $j_n^{(k+\nu)}$ tal que $N_{i^k, j_n^{(k+\nu)}} > 0$ con $n = \{1, 2\}$

(3) $\exists$ $p^{(k+\nu)}$ tal que $N_{p^{(k+\nu)}, j_n^{(k+\nu)}} > 0$ con $n = \{1, 2\}$

(4) $A_{p^{(k+\nu)}} + A_{j_1^{(k+\nu)}} + A_{j_2^{(k+\nu)}} > \alpha A_{i^{(k)}}$

ENTONCES

(1) $f_i = 0$

(2) $f_{j_n}^{(k+\nu)} = 0, \quad n = \{1, 2\}$

Por contra, si la región de contacto existe, entonces se unen las tres regiones en un sola.

SI

(1) $f_i^k = -1$

(2) $\exists$ 2 $j_n^{(k+\nu)}$ tal que $N_{i^k, j_n^{(k+\nu)}} > 0$ con $n = \{1, 2\}$

(3) $\exists\ p^{(k+\nu)}$ tal que $N_{p^{(k+\nu)},j_n^{(k+\nu)}} > 0$ con $n = \{1,2\}$

(4) $A_{p^{(k+\nu)}} + A_{j_1^{(k+\nu)}} + A_{j_2^{(k+\nu)}} < \alpha A_{i^{(k)}}$

ENTONCES

(1) UNIÓN: $j_1^{(k+\nu)} = j_1^{(k+\nu)} \cup j_2^{(k+\nu)} \cup p^{(k+\nu)}$

En una segunda iteración, las regiones con etiqueta nula que no tienen ningún vecino con etiqueta de lesión son etiquetadas como hueso.

SI

(1) $f_i^k = 0$

(2) $f_j^{(k+\nu)} \neq -1,\ \forall j$ tal que $N_{i^k,j^{(k+\nu)}} > 0$

ENTONCES

(1) $f_i^k = 1$

Sobre aquellas regiones con etiqueta nula se decidirá en el procesado posterior.

A continuación se hace una agrupación 3D de las regiones con etiqueta de lesión y posteriormente se incorporan los posibles extremos eliminados erróneamente. Para ello las regiones han de tener etiqueta nula, un área menor del doble del área media en los cortes de la región y ser los más próximos entre los posibles candidatos. Así, las regiones 3D resultantes se unirán siempre y cuando sus extremos estén separados por dos o menos cortes, a través de regiones 2D con etiqueta nula, de la misma forma que en el caso anterior.

Para realizar estas operaciones, el sistema comienza por hacer un crecimiento de regiones con etiqueta de lesión en 3D sin más que concatenar regiones solapantes con esta etiqueta. Se crea de esta forma una lista de nodos, representando cada uno de ellos una posible lesión 3D (L_i) y en el que se guarda la siguiente información:

- Números de los cortes extremos afectados por la lesión: z_{min}, z_{max}.
- Centro de masas de la lesión: $\bar{r} = (\bar{x}, \bar{y}, \bar{z})$.
- Área media de las regiones 2D que componen la región: $\bar{A}$.
- Lista de punteros a las regiones del RAG que forman parte de la lesión: $R_1, \ldots, R_r$.

A partir de esta lista se van uniendo las lesiones 3D que tienen los extremos superior de una e inferior de otra muy próximos a través de regiones con etiqueta nula. El procesado que comentamos afecta pues a la lesión completa. Inicialmente se buscan lesiones que estén alineadas a lo largo del eje z y con extremos muy próximos. La regla es la siguiente:

SI

(1) $dist_1(L_i - L_j) = \min\{|z_{min}^{L_i} - z_{max}^{L_j}|, |z_{min}^{L_j} - z_{max}^{L_i}|\} \leq \epsilon_d$

(2) $dist_2(L_i - L_j) = ||\bar{r}^{L_i} - \bar{r}^{L_j}|| < \max\{\frac{\bar{A}^{L_i}}{\pi}, \frac{\bar{A}^{L_j}}{\pi}\}$

(3) $\exists\ s\ j_l^{(k+l)}$ tal que $f_l^{(k+l)} = 0$ y $N_{j_l^{(k+l)}, j_{l+1}^{(k+l+1)}} > 0,\ \forall\ l < s-1$
con $s = dist_1(L_i - L_j)$ y $k = \min\{z_{max}^{L_i}, z_{max}^{L_j}\}$

ENTONCES

(1) $L_{ij} = \text{UNIÓN}(L_i, L_j)$

El resultado es la unión de ambas zonas de lesión. El origen de esta separación está en un mal etiquetado de las regiones intermedias que pueden dividir una lesión.

El proceso continúa con la eliminación de las lesiones 3D que no abarquen un mínimo número de cortes y que no estén conectadas con el fondo. Éstas regiones serán consideradas como hueso de baja densidad y se les asignará etiqueta de hueso.

SI

(1) $z_{max}^{L_i} - z_{min}^{L_i} < \epsilon_z$

(2) $\text{ADYACENCIA}(R_k^{L_i}, \text{fondo}) = 0, \quad \forall k \in L_i$

ENTONCES

(1) $f_{R_k^{L_i}} = 1, \quad \forall k \in L_i$

Las restantes se etiquetarán como lesión y se suavizarán incorporando vecinos en 2D que aumenten la regularidad de su forma. Esta regularización 3D se aplicará a los cortes que tengan un tamaño al menos un 20% menor que alguno de sus vecinos. Esta incorporación se producirá para regiones que el presegmentador detecta como transición entre hueso y lesión, cuando en realidad son lesión. Si la región tiene asociada una etiqueta de hueso (1), entonces ha de ocurrir que el área de la elipse de dispersión de la unión sea menor o igual que la de la región por si sola, o que, si la etiqueta no es 1, que el área de la elipse de dispersión de la unión sea como mucho 1.5 veces la de la de la región por si sola. Este proceso se lleva a cabo de forma iterativa sobre el volumen completo hasta que no haya más incorporaciones. Las reglas son las siguientes:

SI

(1) $\epsilon_{A,1} \leq \frac{\text{ÁREA}(R_i^{k+\nu,L_i})}{\text{ÁREA}(R_j^{k,L_i})} < \epsilon_{A,2} \quad \nu \in \{-1,1\}$

(2) $\text{ADYACENCIA}(R_j^{k,L_i}, R) > 0$

(3) $\frac{\text{COMPACTACIÓN}(R_j^{k,L_i} \cup R)}{\text{COMPACTACIÓN}(R_i^{k,L_i})} > \epsilon_c$

ENTONCES

(1) $R_j^{k,L_i} = R_j^{k,L_i} \cup R$

SI

(1) $\epsilon_{A,1} \leq \frac{\text{ÁREA}(R_i^{k+\nu,L_i})}{\text{ÁREA}(R_j^{k,L_i})} < \epsilon_{A,2} \quad \nu \in \{-1,1\}$

(2) $\text{ADYACENCIA}(R_j^{k,L_i}, R) > 0$

(3) $\text{SOLAPE}(R_i^{k+\nu,L_i}, R) > \epsilon_s$

[4] $\frac{\text{ÁREA}(R_j^{k,L_i} \cup R)}{\text{ÁREA}(R_i^{k+\nu,L_i})} < \epsilon_{A,2}$

ENTONCES

(1) $R_j^{k,L_i} = R_j^{k,L_i} \cup R$

donde $\text{COMPACTACIÓN} = \frac{4\pi \text{ÁREA}}{(\text{perímetro})^2}$, es una medida de suavidad de la forma de la región. Esta función es igual a 1 para formas circulares y decrece hasta cero cuando la irregularidad aumenta. La función ADYACENCIA devuelve la longitud del perímetro compartido entre dos regiones. La función SOLAPE calcula el tamaño de la proyección de la intersección de las dos regiones en el mismo plano XY.

Finalmente, se calcula una medida de suavidad de las lesiones 3D y se le asignará etiqueta de hueso siempre que no se supere un valor mínimo de suavidad. La regla es la siguiente:

SI

(1) $\#ND = \#\{m | \ \frac{\text{ÁREA}(R_m^{k+1,L_i})}{\text{ÁREA}(R_j^{k,L_i})} \notin [\epsilon_{A,3}, \epsilon_{A,4}]\} > \epsilon_{nd}$

ENTONCES

(1) $f_{R_j^{k,L_i}} = 1, \quad \forall j,k \in L_i$

donde ND es el conjunto de regiones 2D que presentan un valor alto de disparidad en cuanto a área con cualquiera de los vecinos encontrados en los cortes adyacentes y

pertenecientes a la misma zona de lesión 3D. El operador # devuelve la cardinalidad del conjunto. Cuando la cardinalidad del conjunto supere un valor se considerará que la región 3D no es lo suficientemente suave.

Aún quedan por considerar los casos de bifurcación. Cuando se obtenga alguna lesión con bifurcaciones, se conservará únicamente el camino más largo, desechando los demás, puesto que es el más probable de entre los que han superado todas las restricciones previas. En general, el resto de bifurcaciones son mucho más cortas.

Como resultado de la aplicación de estas reglas se alcanza la identificación de las lesiones 3D. Durante el proceso se eliminan las falsas clasificaciones de lesiones obtenidas en el proceso inicial de etiquetado. Además, se realiza la unión de partes no conectadas de la misma región.

4.4 Resultados

Con objeto de ilustrar el proceso seguido, en la figura 4.2.a se representa un corte intermedio en el que aparece un defecto central. Esta imagen corresponde al paciente que había sufrido la fractura de la meseta tibial y cuyos contornos apilados mostramos en la figura 3.53. En la imagen que mostramos ahora se puede observar como existen áreas óseas que aparentemente presentan una densidad menor que la correspondiente a la lesión. Por este motivo aparecen como lesiones tras el proceso de etiquetado, (figura 4.2.e). En la secuencia de cortes mostrada en las figuras 4.2.(d-f) se observa que las verdaderas regiones de lesión preservan una forma y un tamaño similar en los tres cortes, cosa que no sucede con las regiones incorrectamente etiquetadas. Con el procesado 3D basado en la continuidad de tamaño y forma, longitud mínima y conexión con el exterior se consigue un etiquetado correcto, tal y como se ilustra en la figura 4.2.c. La figura 4.3 muestra cortes a distintas alturas de esta lesión correctamente identificada.

La figura 3.53 mostraba ya los contornos apilados de las secuencia de imágenes de este paciente. En ella se incluían los contornos internos de esta lesión, que supone una pérdida ósea. Este conjunto de contornos puede constituir la entrada a un paquete de software convencional que permita reconstruir y visualizar la superficie de la estructura anatómica analizada. Así, en la figura 4.4 se muestra una visualización 3D de esta tibia para la que se usaron los paquetes software NUAGES y GEOMVIEW. En ella se observa el defecto

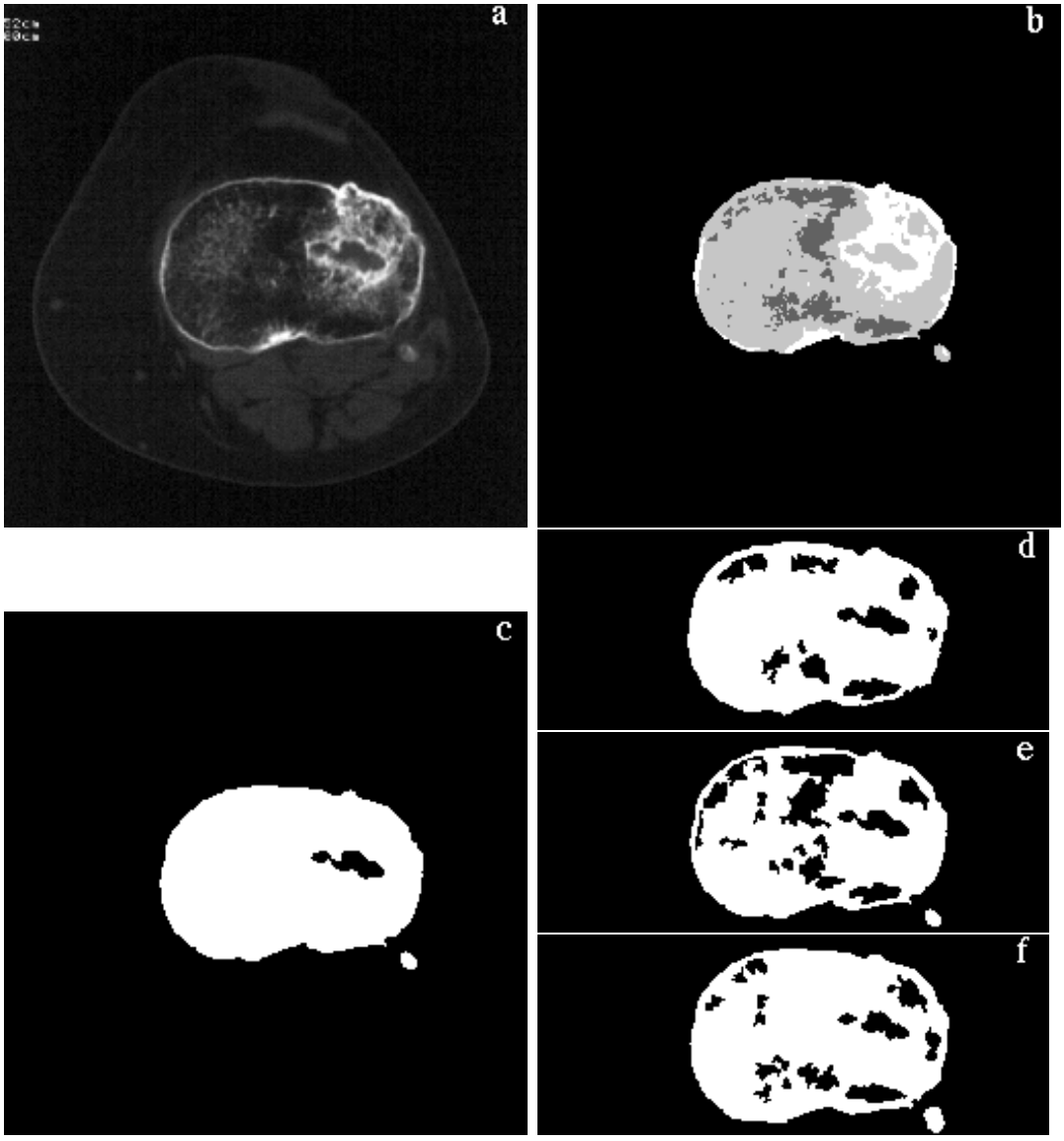


Figura 4.2: (a) Imagen original, (b) presegmentación, (c) etiquetado final, (d-f) etiquetado intermedio de slices contiguos.

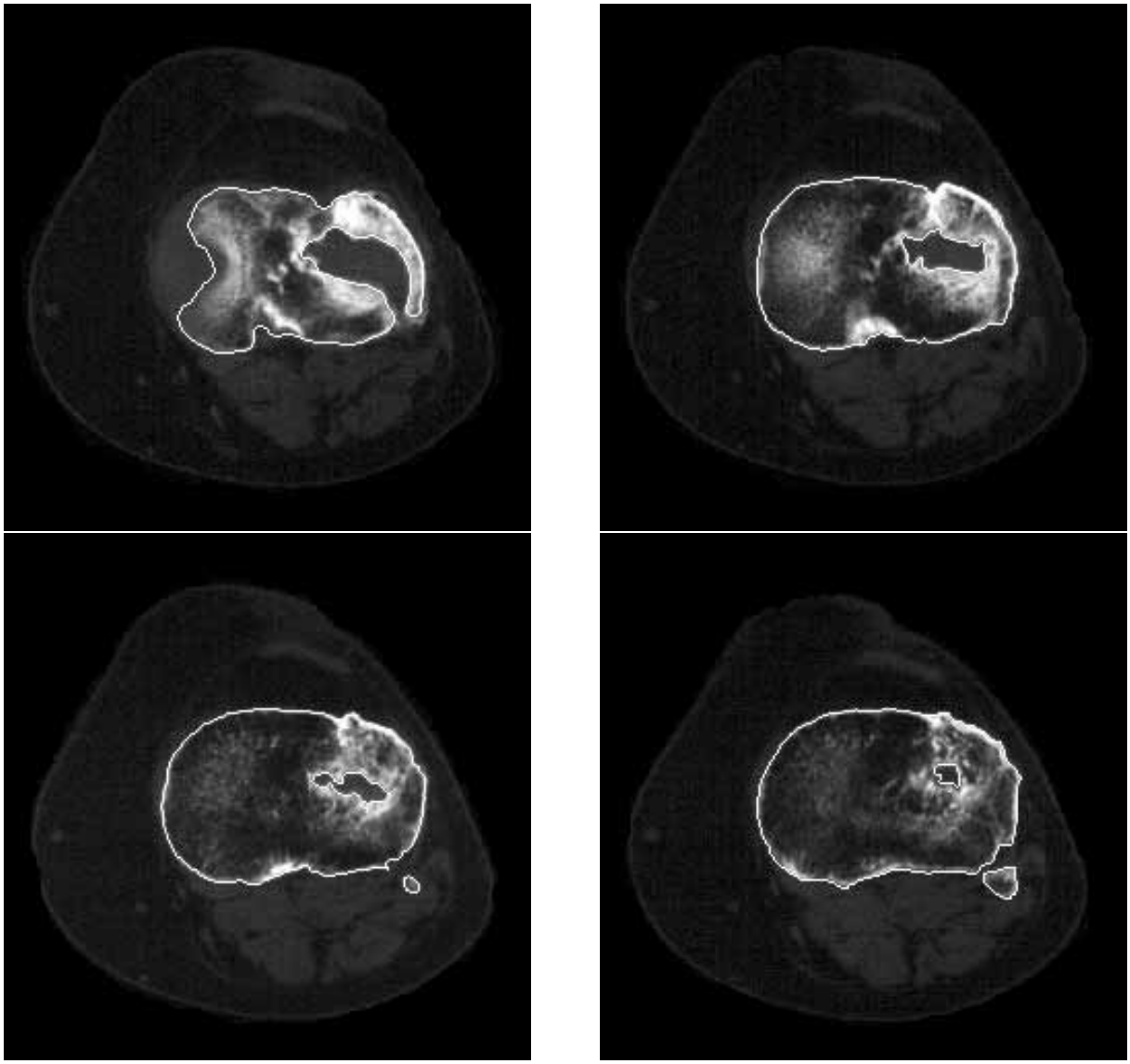


Figura 4.3: Diferentes cortes de la lesión mostrada en la figura 4.4.

central provocado por la fractura.

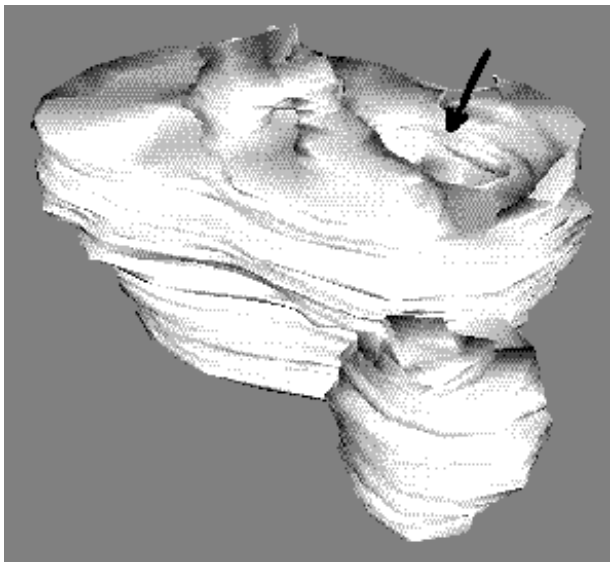


Figura 4.4: Visualización 3D de las superficies de tibia y peroné para los contornos de la figura 3.53. La flecha indica la localización de la lesión.

Otro ejemplo de la identificación de una lesión en diferentes cortes se observa en la figura 4.5. Este caso se corresponde con el del paciente que presentaba una fractura lateral interna de la rodilla izquierda y cuyo conjunto de contornos apilados se mostraba en la figura 3.54. En la figura 4.6 se muestran dos vistas 3D de la misma tibia. En la figura 4.6.a se muestra la la fractura lateral y en la figura 4.4.b muestra otra vista de la reconstrucción de la tibia del mismo paciente, que muestra otra parte de la fractura en la meseta tibial y que se manifiesta únicamente en los últimos cortes de la secuencia. Con el objetivo de apreciar mejor esta lesión es por lo que esta visualización se hace a mayor escala.

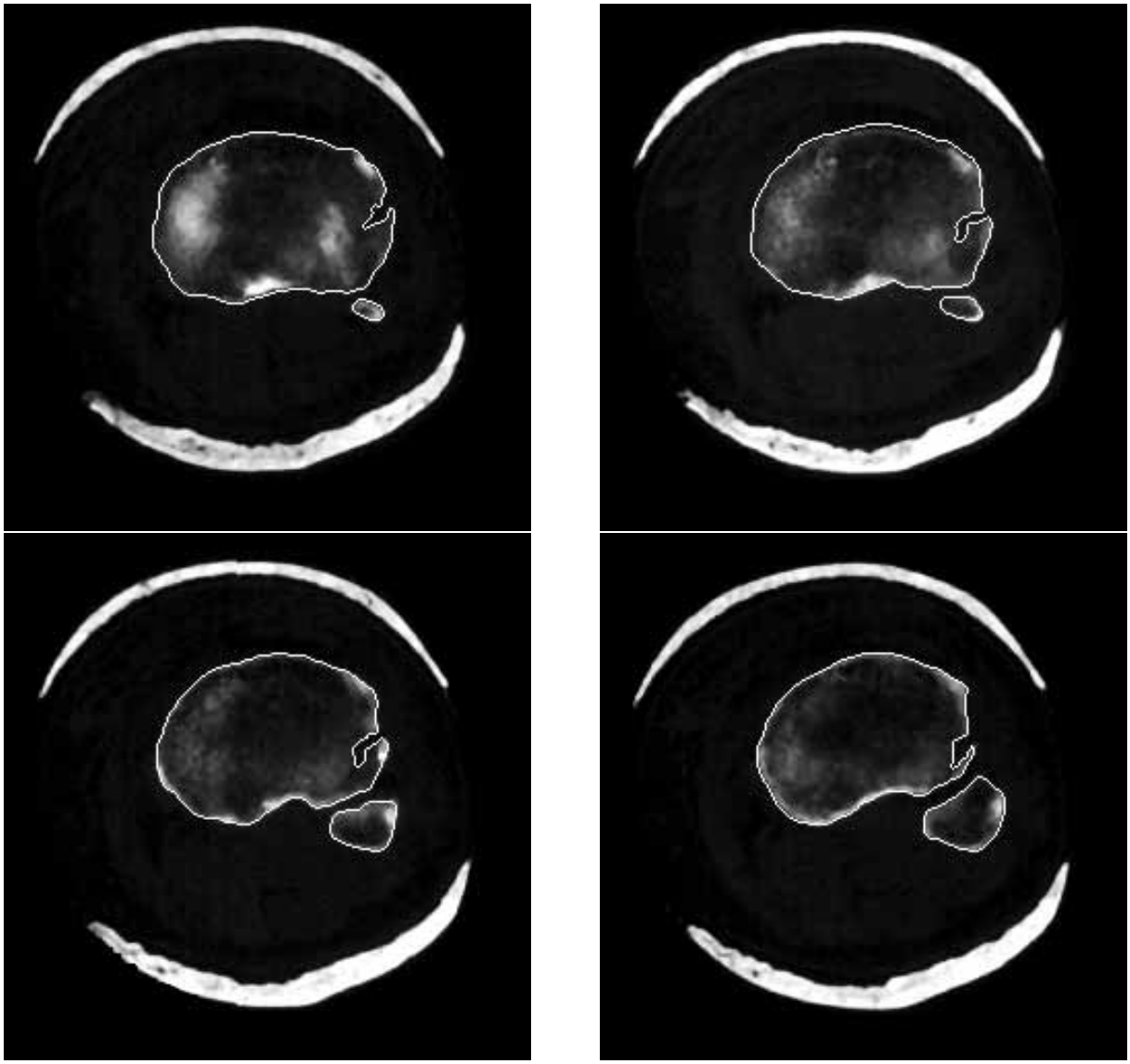


Figura 4.5: Diferentes cortes de la lesión mostrada en la figura 4.6.

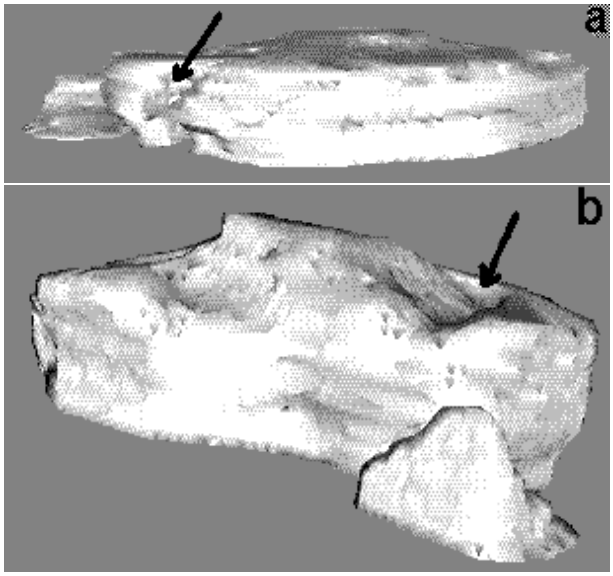


Figura 4.6: Visualizaciones de la reconstrucciones 3D de tibia cuyos contornos se mostraron en la figura 3.54. Las flechas indican las localizaciones de las lesiones.

Capítulo 5

Superficies

5.1 Introducción

Aún cuando en el capítulo anterior hemos usado un software convencional para generar y visualizar las superficies a partir de la secuencia de contornos extraida, todavía queda por generar un modelo sólido de superficie que sea fácilmente trasladable a un software de elementos finitos con el que llevar a cabo el análisis tensional. Para el análisis mediante elementos finitos será preciso llevar cuenta de la numeración de los vértices y de sus conexiones para, por una parte, poder trasladar la definición de las triangulaciones a la sintaxis del software concreto y, por otra parte, controlar los parámetros de mallado que emplee el software de elementos finitos. Por estos motivos será preciso desarrollar un sistema de construcción de superficies propio.

Así pues, en este capítulo abordaremos el problema de la representación de la superficie de un objeto sólido a partir de un conjunto de secciones planas paralelas. Estas secciones transversales o cortes son la base para la interpolación de la superficie de órganos obtenidos mediante exploraciones CT o MR. La superficie interpolada será la que se use como descripción del modelo geométrico de entrada para el análisis tensional mediante elementos finitos. Otra posible aplicación podría ser la definición de las conexiones entre puntos de una malla de superficie para la creación de un modelo de superficies deformables.

El problema de la reconstrucción de superficies consiste en dado un conjunto de datos discretos $\{x^i \in \mathcal{R}^3\}_{i=1}^N$, ajustar una superficie a dicho conjunto. Se pueden distinguir en-

tre dos tipos de conjuntos de datos; aquellos cuyos puntos están espaciados regularmente y aquellos que no satisfacen ninguna condición particular de espaciado o densidad. Estos últimos son los llamados datos dispersos. Nuestro caso se corresponde con el primero de ellos. Nosotros partiremos de un conjunto de puntos en cada plano XY que define los contornos de los objetos de interés en ese plano. A continuación efectuaremos una aproximación poligonal de cada uno de esos contornos seleccionando un conjunto de vértices, o puntos de control, de entre los puntos que definen el contorno cerrado.

El problema de reconstruir superficies se presenta en diferentes campos de la ciencia y se puede plantear de formas diversas [Jain y col., 1995]. Una de ellas consiste en suponer que los datos tienen su origen en una función de dos variables definida sobre un dominio W de $\mathcal{R}^2$. En este caso los datos se representan en la forma (x_1^i, x_2^i, f^i) donde (x_1^i, x_2^i) son coordenadas de un punto en W y f^i su correspondiente valor. El problema es encontrar una función de dos variables f tal que $f(x_1^i, x_2^i) = f^i$ para todo i. Sin embargo, el problema funcional no cubre todas las aplicaciones. Así, en el modelado de estructuras anatómicas no se puede esperar el encontrar siempre funciones simples que interpolen la superficie 3D a través de todos sus datos. En este caso, y otros similares, se particiona la superficie en trozos más pequeños, cada uno descrito por una función separada, formando así una representación a trozos.

La aproximación más simple de la superficie es la poliédrica, obtenida mediante triangulaciones entre los puntos de la superficie. Ésta es la que nosotros adoptaremos, ya que es la más simple que proporciona una organización de los datos de superficie y que admite cualquier software de elementos finitos. Además, es la base para representaciones de orden superior. Finalmente, dentro de este capítulo exploraremos la posibilidad de abordar la segmentación directamente en 3D mediante superficies deformables. Para la definición de las superficies deformables emplearemos el mismo algoritmo de triangulación entre vértices.

5.2 Representación de la superficie

Para la representación geométrica de la superficie partiremos de los conjuntos de puntos (puntos muestra) que definen contornos cerrados en cada corte. El término reconstrucción de superficie significa estimación de la función continua para la superficie a partir

de un conjunto de puntos clave extraidos de entre todos los puntos muestra disponibles inicialmente. En la representación por triangulación de la superficie se emplea una función de interpolación bilineal a trozos de tal forma que la posición de la superficie en las zonas no definidas a priori (fuera de los puntos clave) se hace a partir de los tres vértices de cada triángulo. El problema está, por tanto, en la definición de los triángulos para un recubrimiento completo de la superficie.

Existen varios métodos para la selección de puntos clave de entre los puntos muestra [Kang y col., 1993]. Cuando el número de pixels entre dos secciones adyacentes es el mismo, una de las alternativas consiste considerar todos los puntos muestra como puntos clave y realizar el emparejamiento uno-a-uno de todos ellos. Sin embargo, esto sólo es aplicable a geometrías simples. Cuando el número de puntos en cada corte varía, no se puede generalizar sobre la numeración de los elementos en la reconstrucción de la forma y el modelado, por lo que la superficie del objeto ha de determinarse mediante recubrimiento triangular. Una forma de abordar el problema sería chequear la dirección de los puntos, de tal forma que aquellos que presenten un cambio en la dirección se seleccionan como puntos de control o clave de la forma.

Una segunda alternativa consiste en la selección de puntos clave de entre los puntos que definen cada contorno (puntos muestra) a intervalos equiangulares. Este método divide el contorno completo de una sección transversal en n segmentos que sustentan ángulos idénticos con respecto al centroide del contorno cerrado. Sin embargo, este método da malos recubrimientos cuando la diferencia de pixels es mayor del 15% [Kang y col., 1993].

La tercera alternativa es la selección por igual longitud, la cual es aplicable a cualquier forma irregular de sección. Ésta consiste en seleccionar puntos equidistantes en el parámetro de la curva que define el contorno. Para seleccionar puntos por este método es preciso escoger de forma consistente un punto de referencia para cada sección. Una posibilidad es hacerlo con respecto al centroide de cada sección. Así, tomando un segmento desde $\vec{X} = (x, y, z)$ en la misma dirección para todos los cortes, la intersección de este segmento y el contorno sería el punto de referencia, en el que comienza la selección de puntos clave equidistantes. La longitud total del contorno se divide en n segmentos de la misma longitud. Los extremos de los segmentos definen los puntos de control. La numeración de los puntos permitirá establecer el emparejamiento con los de los cortes vecinos. Esta última solución es la que nosotros adoptaremos. Este esquema hay que adaptarlo para el punto en que se produzca una bifurcación, donde habrá que buscar la correspondencia entre un

contorno de un corte con dos o más del corte siguiente.

El algoritmo que nosotros proponemos para la reconstrucción de la superficie consta de los siguientes pasos:

- Adquisición de los puntos de cada uno de los contornos de los cortes de entrada con una orientación consistente (mismo sentido de recorrido: horario o antihorario).
- Discretización de cada contorno en una secuencia cíclica de vértices, equidistantes en el parámetro de la curva.
- Triangulación. Si el número de contornos es idéntico en los dos cortes, se emparejan contornos según proximidad de centro de masas y se procede con la triangulación. Si no es así, nos encontramos ante un punto de bifurcación en el que la correspondencia de contornos no es como antes uno–a–uno, sino uno–a–varios. Esta situación se resolverá mediante la generación de contornos intermedios previa a la triangulación.

Indicaremos en primer lugar la forma de resolver el problema de la bifurcación para finalmente ver el modo de efectuar la triangulación de la superficie.

5.2.1 Bifurcación

Cuando se trata de determinar la superficie delimitada por dos contornos situados en los planos $z = z_1$ y $z = z_2$, la solución no es única, por lo que se buscará una que sea intuitivamente *buena* [Baqueret y Sharir, 1996]. Las situaciones que nosotros tratamos se caracterizan por un alto grado de solape entre las regiones encerradas por los contornos, por lo que, una vez establecido el criterio más adecuado para la determinación de la superficie (maximizar volumen, minimizar superficie, minimizar longitud de lado o maximizar ángulo mínimo), la *conexión* de los contornos es relativamente simple. Así, hemos adoptado como criterio el de minimizar la longitud de lado y maximizar el ángulo mínimo, por ser más adecuado para análisis mediante elementos finitos. Sin embargo, cuando se trata de crear la superficie que está limitada por más de un contorno dentro del mismo plano, el problema se complica. La estrategia que nosotros adoptamos consiste en la creación de contornos intermedios para el manejo de la bifurcación. Este idea ya fue propuesta por Kankanhalli y col. [1993] para el modelado tridimensional de la

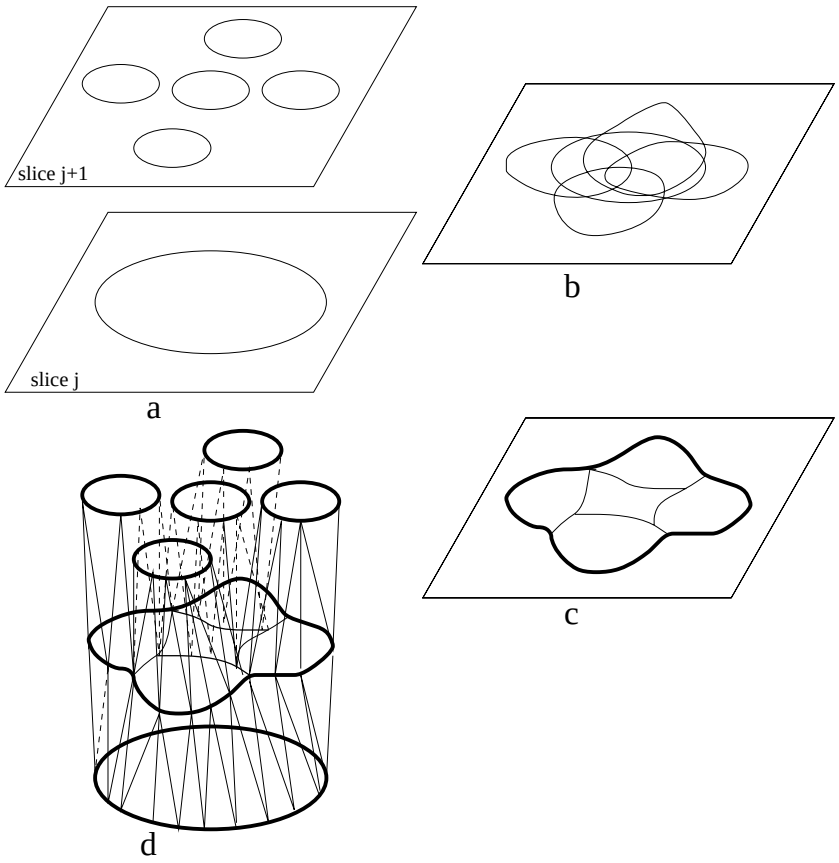


Figura 5.1: Bifurcación: (a) cortes contiguos y sus contornos, (b) contornos intermedios entre los del corte j+1 y el del corte j, (c) división del contorno unión, (d) triangulación con los contornos intermedios.

articulación de rodilla. El algoritmo comienza por etiquetar todos los contornos presentes en los cortes, almacenando tamaño y coordenadas de cada uno de sus puntos. Una vez hecho ésto, se hace un procesado de cada par de cortes consecutivos. Si existe un único contorno por estructura en cada uno de los cortes, entonces se procede con la triangulación de la superficie entre contornos uno–a-uno en la forma que se indicará más adelante. En el caso en el que el emparejamiento de contornos sea uno–a–varios, se procederá con la creación de contornos intermedios, de tal forma que nos permitan establecer de nuevo una triangulación de la superficie entre contornos uno–a–uno.

Para la creación de los contornos intermedios a los cuales se conectarán los contornos

del punto de bifurcación se considerarán los diferentes pasos que se ilustran en la figura 5.1. En primer lugar se genera el contorno intermedio entre cada una de los contornos de la bifurcación y el correspondiente contorno del corte previo. A continuación, se crea la máscara correspondiente a cada uno de estos contornos y se eliminan los solapes entre las máscaras. Finalmente, se determinan los contorno para cada una de estas máscaras. A continuación describiremos de forma detallada cada uno de los pasos del proceso tomando como referencia la situación de correspondencia uno-a-varios ilustrada en la figura 5.1.a. En el proceso de creación de contornos intermedios se realizará inicialmente una aproximación poligonal del contorno, de tal forma que la distancia entre cada dos vértices sea constante en el parámetro de la curva. La razón de adoptar este criterio es que en la mayoría de pares de contornos se va a tener una alta similaridad en forma y tamaño, por lo que eligiendo un vértice de comienzo de forma consistente, el algoritmo de triangulación de la superficie no va a tener una excesiva influencia en la forma final. Por otra parte, permitirá manejar bien los casos de cambios bruscos de forma en partes del contorno. Para determinar el vértice de comienzo de la aproximación poligonal haremos uso del centro de masas del área delimitada por el contorno, definido como

$$\overline{x} = \frac{\int x dA}{A}, \quad \overline{y} = \frac{\int y dA}{A} \tag{5.1}$$

A continuación buscaremos la intersección del segmento con origen en el centro de masas y paralelo al eje y con la parte más externa del contorno en esa dirección y en el sentido de los y's positivos, (x_k, y_k). El punto (x_k, y_k) será el primer vértice de la aproximación poligonal, y los demás se escogerán equidistantes dentro de la curva de contorno,

$$V = \{(vx_i, vy_i) \in C | (vx_i, vy_i) = (x_{(k+i*\delta)\%l_c}, y_{(k+i*\delta)\%l_c}), 0 \leq i < n\} \tag{5.2}$$

donde n es número de vértices de la aproximación poligonal, δ la distancia entre vértices a lo largo de la curva, l_c el tamaño del contorno, (x_j, y_j) el punto del contorno C para el que el parámetro de la curva toma el valor j, y $\%$ es el operador módulo.

Sean $C_{j+1,l}, \quad l \geq 2$, los contornos del corte $j+1$ que se proyectan sobre un mismo contorno C_j del corte anterior. El siguiente paso consiste en la obtención de formas intermedias entre el contorno C_j y cada uno de los contornos $C_{j+1,l}$. Para realizar esta tarea

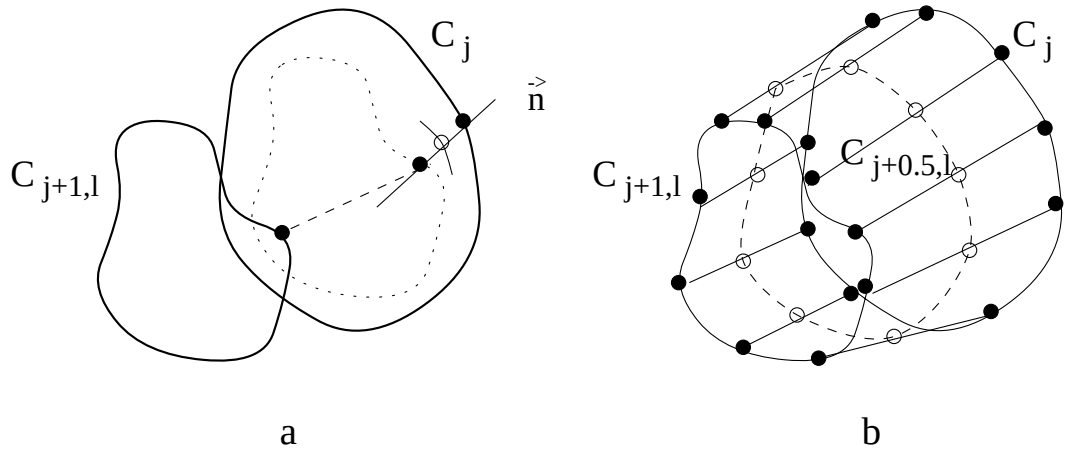


Figura 5.2: Generación del contorno intermedio ($C_{j+0.5,l}$) a dos contornos de la misma estructura en cortes diferentes: el único contorno de la estructura en el corte j (C_j) y uno de los contornos ($C_{j+1,l}$) en el que se bifurca la estructura en el corte $j+1$. (a) Proyección del contorno $C_{j+1,l}$ en el plano del corte j y búsqueda de correspondencias. (b) Contorno intermedio final en línea discontinua.

buscamos el punto medio del segmento entre cada vértice de la aproximación poligonal del contorno $C_{j+1,l}$ proyectado en el corte j y el punto más próximo del contorno C_j. Este punto medio se busca en la dirección normal al contorno $C_{j+1,l}$ proyectado en el plano j, tal y como se muestra en la figura 5.2.a. Para determinar el segmento de búsqueda lo que se hace es determinar la normal al contorno proyectado y localizar el punto del mismo contorno más alejado situado hacia el interior del mismo. Definiremos un *centro de masas local* (cm_l) como el punto medio del segmento que une ambos puntos del contorno proyectado. A partir del cm_l, y en el sentido de la normal hacia el exterior del contorno, buscamos la localización del primer pixel de contorno C_j. Repitiendo esta operación para cada vértice de la aproximación poligonal del contorno $C_{j+1,l}$ obtenemos un conjunto de vértices que representarán la aproximación poligonal del contorno intermedio buscado. La figura 5.2.b muestra el contorno intermedio final ($C_{j+0.5,l}$) localizado entre dos contornos correspondientes de cortes diferentes (C_j y $C_{j+1,l}$). Los puntos indican los vértices de las aproximaciones poligonales. La figura 5.1.b muestra el resultado de la generación de los contornos intermedios para el caso de la figura 5.1.a. Los contornos intermedios generados se solaparán, por lo que será necesario un proceso posterior que elimine dichos solapes.

La coordenada z de dichos vértices será la media de la de los dos cortes, que suponemos, sin pérdida de generalidad, normales al eje $\vec{Z}$.

Una vez creadas las aproximaciones poligonales de los contornos intermedios, el siguiente paso consiste en determinar las máscaras de las regiones encerradas por cada una de ellas. Este cálculo se realiza para facilitar la determinación del contorno unión de todos los contornos intermedios, y la intersección entre cada par de ellos.

El algoritmo utilizado para la obtención de la máscara a partir de un contorno poligonal es aplicable a polígonos cóncavos y convexos. El algoritmo opera mediante el cálculo de segmentos entre extremos derecho e izquierdo del polígono. Es preciso determinar qué pixels, dentro de estos segmentos, están dentro del polígono. Para hacer ésto procedemos de la forma que se indica a continuación. Se explora la imagen con el contorno de izquierda a derecha y de arriba abajo. Al principio de cada fila se pone un indicador de polaridad a valor 0, y en cada intersección con el contorno se hace una inversión de polaridad ($polaridad = (polaridad + 1)\%2$). Los pixels accedidos con polaridad 1 corresponden al interior del polígono, mientras que los otros no. Aquí sin embargo se presentan los dos problemas que se reflejan en la figura 5.3: tratamiento del caso especial de vértices en los extremos de partes cóncavas y convexas, y tratamiento del caso especial en el que los vértices definen una línea horizontal. Para resolver estos dos problemas se hace inicialmente un etiquetado de los puntos del contorno poligonal según el tipo de vértice: 0 si no es vértice, 1 si es máximo aislado en Y, 2 si es mínimo aislado en Y, 3 si es extremo inicial de segmento horizontal, y 4 si es extremo final de segmento horizontal, tal como muestra la figura 5.4.

En la exploración de la imagen, cada vez que se encuentra un punto del contorno que no sea punto intermedio en un trozo de contorno horizontal, se cambia la polaridad. Adicionalmente, si el tipo de vértice asignado es 1 ó 2, se vuelve a invertir la polaridad y se sigue examinado la fila. Si el tipo de vértice es 3, se busca el final del trozo horizontal del contorno y se mira si alguno de los vecinos 8–adyacentes a los pixels de este segmento en las filas anterior y siguiente son puntos de contorno; si hay vecinos en las dos filas, entonces se vuelve a invertir la polaridad. Si el tipo de vértice del pixel actual es 0 ó 4 y la polaridad es 1, se anota como primer extremo de un segmento interior; si la polaridad es 0, indica que estamos en el segundo extremo y se etiquetan como interiores todos los pixels del segmento. El tratamiento que se hace para el tipo de vértice 3 consiste en buscar el extremo del segmento horizontal (tipo de vértice igual a 4). Una vez encontrado, en el

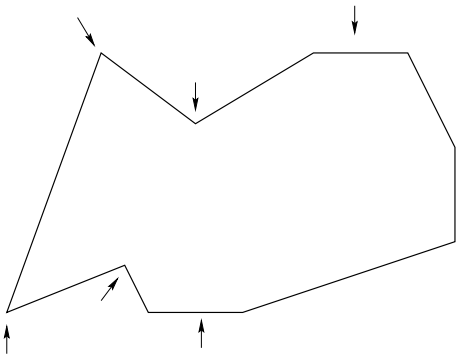


Figura 5.3: Puntos conflictivos para la determinación de máscaras.

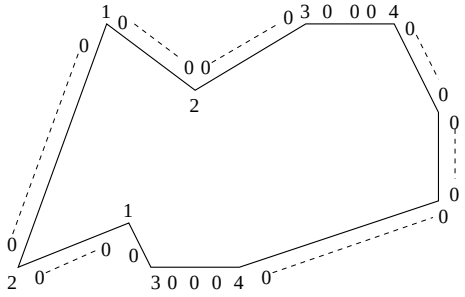


Figura 5.4: Ejemplo de etiquetado de vértices.

caso de que sus extremos sean mínimos o máximos locales del contorno, se incluye como parte de la máscara y se empieza el examen del resto de la fila con el valor de polaridad a 0, de igual forma que al inicio de la misma. Si no lo son, significará que los pixels a la derecha del extremo con tipo de vértice 4 también serán interiores, por lo que aún no se ha alcanzado el segundo extremo del segmento.

Una vez calculadas todas las máscaras, se procede con el cálculo de los contornos de la unión de todas ellas y de las uniones e intersecciones de pares no disjuntos. Este contorno unión y los contornos intersección permitirán dividir el contorno unión en otros varios iguales en número a los del corte $j+1$. La figura 5.5 ilustra el proceso de partición de la máscara unión en tantos contornos adyacentes no solapantes como contornos haya en el corte $j+1$. Se buscan los puntos compartidos por la unión y la intersección y se traza el contorno medio de la intersección que une esos dos puntos. Cada uno de estos

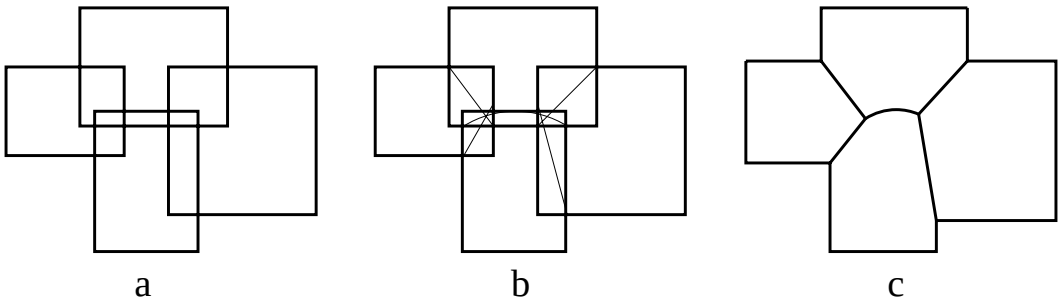


Figura 5.5: Obtención de los contornos intermedios finales: (a) contornos intermedios solapados generados de forma individual entre pares de contornos, (b) eliminación de los solapes mediante división por puntos medios de intersección, (c) contornos cerrados finales para el corte intermedio.

contornos medios se llevan a la imagen del contorno de la unión de todas las máscaras, como se indica en la figura 5.5.b. Se eliminan los trozos de borde que tengan un extremo libre y se obtiene un conjunto de contornos de regiones adyacentes no solapantes, como el de la figura 5.5.c. Como resultado se obtiene un conjunto de contornos cerrados que se utilizarán para hacer las triangulaciones con los contornos $C_{j+1,l}$.

Finalmente, se hará una triangulación entre el contorno C_j y el contorno unión $C_{j+0.5,\cup l}$, y triangulaciones entre cada una de los contornos cerrados en los que se dividió la unión $C_{j+0.5,l}$ con el correspondiente contorno $C_{j+1,l}$. De esta forma se consigue resolver el problema de la triangulación en los puntos de bifurcación. En la figura 5.1.c se muestran los contornos cerrados finales para el corte intermedio. El único contorno del corte previo se conectará al contorno externo de la figura 5.1.d, y cada uno de los contornos del corte siguiente se conectará con el contorno correspondiente del corte intermedio.

5.2.2 Triangulación

El objetivo ahora es derivar una superficie (triangulada) interpolando los vértices de control de cada par de contornos correspondientes entre dos cortes adyacentes. Para realizar esta operación de creación de la superficie procederemos empleando una simplificación del algoritmo de Mencl [1995], que se estructura en los siguientes pasos:

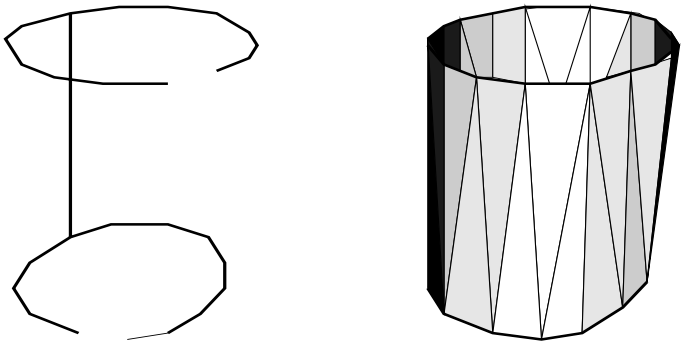


Figura 5.6: El EMST conecta dos anillos con un segmento (a); posteriormente se triangula la superficie entre dichos anillos (b).

1. Cálculo del árbol de mínima extensión euclídea (*euclidean minimum spanning tree* (EMST)) del conjunto de puntos que comprenden los vértices de las dos aproximaciones poligonales de los dos contornos.

2. Extensión del EMST a un grafo de descripción de superficie (SDG) más denso, definiendo contornos sobre la superficie.

3. Rellenar los contornos del SDG con triángulos.

El propósito de los dos primeros pasos es construir un esqueleto global de la superficie. En el paso 3, el esqueleto se extiende a una malla triangular que define la superficie reconstruida. La ventaja del planteamiento es que el EMST se adapta bien a conjuntos de puntos de densidad variable, lo que puede causar dificultades a otros métodos. Sea $P = \{p_1, \ldots, p_n\}$ un conjunto de puntos que describe una superficie. Un EMST es un árbol que conecta todos los puntos de P con segmentos de línea de tal forma que la suma de las longitudes de los segmentos es mínima. Un árbol es un grafo conectado no cíclico.

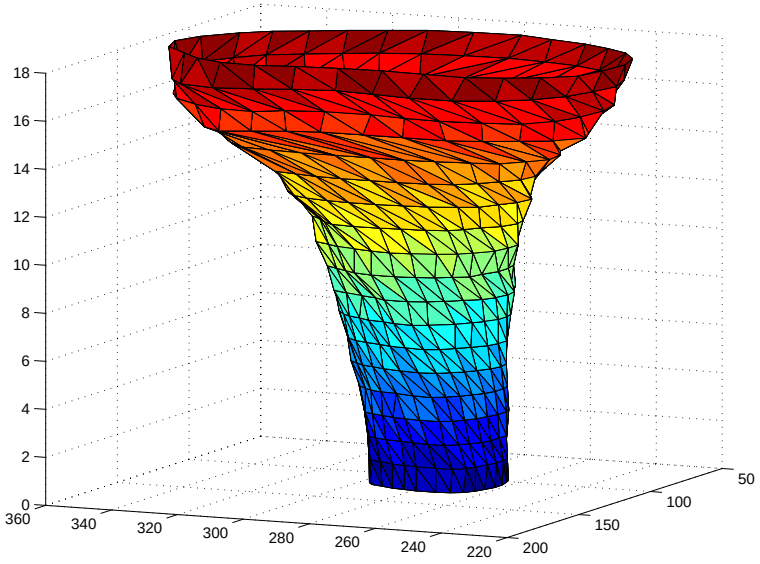
Para simplificar el cálculo del EMST vamos a suponer que la distancia entre contornos es mayor que la existente entre vértices adyacentes del mismo contorno. De esta forma el cálculo se reduce a la determinación de la máxima distancia entre vértices adyacentes en cada uno de los contornos, y la distancia mínima entre cada par de vértices (v_1^k, v_2^j), $v_1^k \in C_1, v_2^j \in C_2$. Como resultado se obtiene un árbol similar al de la figura 5.6. Esta

operación se simplifica si a la hora de extraer los vértices de la aproximación poligonal comenzamos en una dirección fija a partir del centro de masas del contorno, tal y como habíamos dicho antes. Así, la conexión entre los dos contornos se efectuaría a través del primer vértice de ambos contornos. Con la suposición anterior de la distancia entre contornos mayor que la existente entre vértices vecinos del mismo contorno, el SDG se completa con unir los vértices hoja del árbol con el otro perteneciente al mismo contorno. El siguiente paso es recubrir con triángulos el esqueleto dado por el SDG con el fin de obtener una malla triangular como representación de la superficie. La triangulación se construye de forma incremental mediante la adición de segmentos que conectan vértices. El siguiente segmento a añadir se selecciona escogiendo el par de segmentos incidentes, entre todos aquellos que aún no han inducido un triángulo, con el ángulo encerrado más pequeño. El segmento que cierra este par formando un triángulo será añadido al grafo. El proceso se repite mientras existan pares de segmentos incidentes adecuados. Para la adecuada triangulación de la superficie es preciso incluir dos ligaduras en este proceso:

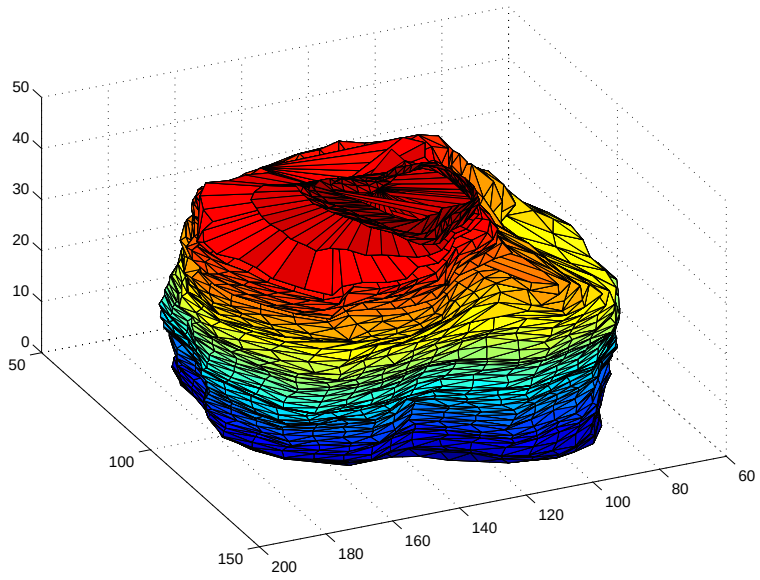
- **Ligadura del triángulo incidente**: en un segmento pueden incidir como máximo dos triángulos.
- **Ligadura de intersección de triángulos**: los triángulos no pueden intersectar.

Estas ligaduras son precisas para garantizar una única superficie cerrada que utilice los contornos poligonales como límites superior e inferior.

En la figura 5.7.a se muestra el resultado de la triangulación de la superficie para la segmentación mostrada a través de un conjunto de contornos apilados en la figura 3.56. Aquí se representa la triangulación de las superficies laterales. Los contornos en el primer y último corte ya definen los límites superior e inferior del hueso. De esta forma, definida la superficie lateral y conocidos estos extremos, ya tenemos perfectamente delimitado el volumen de entrada al software de elementos finitos. La figura 5.7.b contiene otro ejemplo en el que se muestra el resultado de la triangulación de la superficie para el conjunto de contornos apilados en la figura 3.54. Dependiendo del resultado del mallado de volumen que efectuará el software de elementos finitos, puede ser necesario ajustar el número de puntos de control que definen la superficie. Esta modificación puede implicar el aumento de puntos de control para evitar ángulos demasiado agudos, o la disminución para evitar formas demasiado complejas. Estas modificaciones son dependientes del mallador del software utilizado.



a



b

Figura 5.7: Generación de la superficie mediante triangulación para el conjunto de contornos apilados: (a) de la figura 3.56, (b) de la figura 3.54.

Finalmente, en la figura 5.8 se muestra la triangulación obtenida para el punto en que una lesión provocaba la bifurcación de la tibia, cuyos contornos apilados se representaron en la figura 3.54.

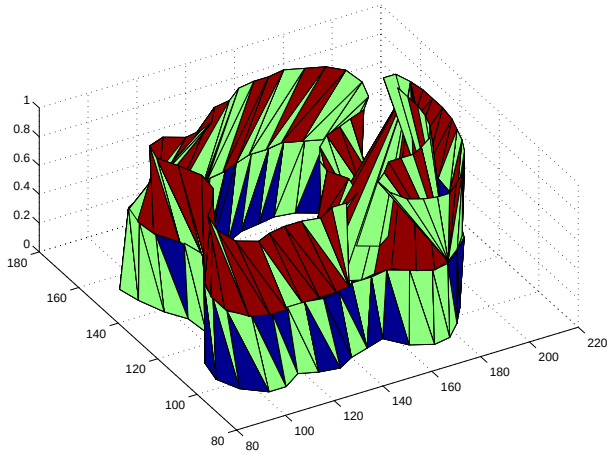


Figura 5.8: Generación de la superficie mediante triangulación en un punto de bifurcación de la tibia visualizada en la figura 4.4.a.

5.3 Superficies deformables

La segmentación de imágenes 3D corte a corte mediante el apilamiento de contornos 2D puede dar lugar a inconsistencias. El uso de verdaderos modelos de superficies deformables 3D puede proporcionar una técnica de segmentación más robusta, que asegure una suavización de forma global y la coherencia de la superficie entre cortes. Los modelos de superficies deformables en 3D fueron utilizados por primera vez en visión computacional por Terzopoulos y col. [1988]. Cohen y Cohen [1993] propusieron el uso de elementos finitos y técnicas basadas en principios físicos para implementar un plano deformable colocado sobre imágenes 3D de cabezas humanas obtenidas mediante Resonancia Magnética. Staib y Duncan [1992a] describieron un modelo de superficie 3D para la puesta en correspondencia de imágenes médicas 3D. El modelo usa una parametrización de Fourier que representa la superficie como una suma ponderada de funciones base sinusoidales. La determinación de la superficie se formula como un problema de optimización mediante una función obje-

tivo de máximo a posteriori. La ventaja de la parametrización de Fourier es que permite que una amplia variedad de superficies suaves sean descritas con un reducido número de parámetros. Metaxas y Terzopoulos [1993] presentaron un esquema basado en principios físicos para la estimación de formas y movimientos no rígidos. Sus modelos dinámicos incorporaron principios mecánicos de cuerpos rígidos y deformables dentro de las primitivas geométricas convencionales. Aplicaron la mecánica Lagrangiana y el método de los elementos finitos para derivar las ecuaciones diferenciales del movimiento que gobiernan la dinámica de los modelos. Nosotros exploramos aquí la posibilidad de extender nuestro modelo propuesto para contornos deformables en 2D a superficies deformables en 3D. Un importante inconveniente de las superficies deformables es la no disposición de un modelo inicial aproximado, por lo que el modelo que aquí se formula es un planteamiento inicial que necesita de importantes mejoras que mitiguen los problemas ocasionados por la mala inicialización.

5.3.1 Modelo de superficies activas 3D

Cuando la imagen 3D es una secuencia de imágenes 2D, la reconstrucción 3D se puede abordar mediante un procesado corte a corte, inicializando el proceso mediante una curva obtenida para el contorno final del corte más alejado del extremo proximal, tal y como hemos comentado en el capítulo 3. De esta forma se propagan los resultados a los cortes vecinos. El principal incoveniente es que no hay otra interacción entre contornos de cortes vecinos en el proceso de deformación, a parte de la inicialización. Si el contorno se perdiese en un corte particular, el método fallaría, mientras que si se usase un verdadero modelo 3D, se abriría la posibilidad de recuperarlo. Mediante el modelo 3D, la curva en un corte se ve afectada también por características de cortes vecinos, lo cual ayuda a reconstruir la superficie completa. El modelo 3D se representa mediante una superficie $\mathcal{S}$, que se define como una proyección v:

$$\begin{aligned} v : \Omega &= [0,1] \times [0,1] \rightarrow \mathcal{R}^3 \\ (s,r) \mapsto v(s,r) &= (v_1(s,r), v_2(s,r), v_3(s,r)) \end{aligned}$$

La energía asociada E se define sobre una clase asociada $\mathcal{A}$ de proyecciones v:

$$E : \mathcal{A} \rightarrow \mathcal{R}$$

$$v \mapsto E(v) \quad = \quad \int_{\Omega} [E_{int}(v(s,r)) + P(v(s,r)]ds\ dr \tag{5.3}$$

donde el primer término es la energía interna asociada con las fuerzas internas actuantes sobre la forma de la superficie. El segundo término representa el potencial de las fuerzas externas derivadas de la imagen 3D. La solución se obtiene miminizando la energía de la ecuación 5.3, dada la estimación inicial $v(0,s,r)$ y las condiciones de contorno.

Energía interna

El aspecto más comúnmente usado para incorporar la energía interna es el de la suavidad. En el esquema de los contornos activos se usaron como medidas de suavidad las primera (estiramiento) y segunda derivadas (flexión). En 3D, la energía de estiramiento puede formularse en base a la distancia al cuadrado promedio de un vértice de la superficie con los vecinos directos a los que está conectado,

$$E_{stretching}(v_i(x,y,z)) = \alpha_i \frac{1}{\#N_i} \sum_{j \in N_i} ||v_i - v_j||^2, \tag{5.4}$$

mientras que la curvatura de la superficie puede utilizarse como medida de flexión de la misma.

La curvatura de una superficie se define usando el concepto de curvatura de una curva plana. Considérese un plano que corta la superficie en el punto $\vec{p}$ y es normal a la superficie en este punto. La intersección del plano con la superficie es una curva plana que incluye el punto $\vec{p}$. La curvatura normal en ese punto es la curvatura de la curva plana de intersección. La curvatura normal no es única, puesto que depende de la orientación del plano respecto de la normal. Las curvaturas mínima y máxima, κ_1 y κ_2, se llaman curvaturas principales, y las direcciones en el plano tangente correspondientes a estas curvaturas son las direcciones principales. La curvatura Gaussiana (K) es el producto de las curvaturas principales, y la curvatura media (H) es el promedio de las curvaturas principales,

$$K = \kappa_1 \times \kappa_2, \quad H = \frac{\kappa_1 + \kappa_2}{2} \tag{5.5}$$

La curvatura media de una superficie es una medida intrínseca de su flexión. La curvatura

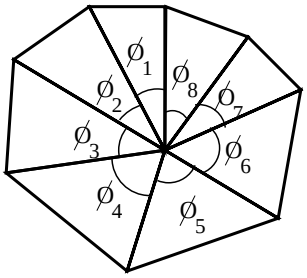


Figura 5.9: Ángulo sólido.

media describe sólo un aspecto de la forma de la superficie. Una medida complementaria es la curvatura Gaussiana, medida a lo largo de la geodésica de la superficie. En una triangulación de la superficie, la curvatura Gaussiana es fácilmente calculable mediante el exceso esférico o ángulo sólido. Esta medida se puede ilustrar con la figura 5.9, donde el punto central corresponde al punto i de la superficie en el que se está midiendo la curvatura, y el resto de vértices de los triángulos son los vértices vecinos de i en el mismo contorno y en los contornos de los cortes adyacentes. Aquí cada uno de los triángulos está, en general, en un plano diferente, por lo que la suma de los ϕ_j no es, necesariamente, igual a 2π. La curvatura Gaussiana en un vértice i de una triangulación se puede estimar como la razón entre exceso esférico y la suma de las áreas de los triángulos:

$$K_i = \frac{\Theta_i}{\text{Sum Area}} \tag{5.6}$$

donde Sum Area es la suma de las áreas de todos los triángulos con un vértice en i, y siendo:

$$\Theta_i = 2\pi - \sum_{j \in N_i} \phi_j, \tag{5.7}$$

el exceso esférico.

En una triangulación sólo se pueden estimar medidas de la curvatura media sobre los triángulos que rodean a cada vértice; éstas se obtienen como la suma de la longitud de cada lado multiplicada por los ángulos del diedro asociado:

$$\int\int_T H dA = \sum_i l_i \Psi_i \tag{5.8}$$

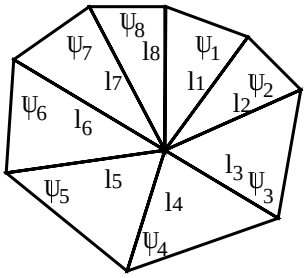


Figura 5.10: Curvatura media.

donde T es la suma de las áreas de los triángulos que rodean al vértice, l_i es la longitud del lado i, y Ψ_i es el ángulo del diedro sobre el lado i, tal y como se muestra en la figura 5.10. Por tanto, el índice de curvatura natural en una triangulación es la curvatura Gaussiana [Delingette, 1994]. Por tanto nosotros tomamos la siguiente expresión como energía de flexión:

$$E_{bending}(v_i) = \beta_i |K_i| \tag{5.9}$$

Cuando se trabaja con verdaderos modelos deformables 3D no es posible, en general, garantizar que el modelo inicial esté próximo a la posición final deseada. Típicamente se parte de posiciones internas (externas) a la estructura a segmentar, y se aplican fuerzas de inflación (deflación) para intentar superar los mínimos locales. Aquí resulta muy ventajoso hacer que los coeficientes de elasticidad α y β varíen en función de la intensidad del pixel sobre el que esté posicionado cada uno de los vértices. Con ello se permite dar mayor flexibilidad a la superficie cuando ésta se encuentra alejada de la superficie del objeto. De esta forma la evolución de la superficie deformable se hace menos dependiente de la forma del modelo inicial. Se permiten así deformaciones provocadas por las fuerzas externas débiles, como las de inflación o deflación, en zonas con alto grado de homogeneidad y valores bajos de curvatura y distancia al centroide. De acuerdo con esto escogemos las siguientes expresiones para los coeficientes de elasticidad:

$$\alpha_i = \alpha_0 + \alpha_1(t) I(v_i(s,r)), \quad \beta_i = \beta_0 + \beta_1(t) I(v_i(s,r)) \tag{5.10}$$

donde t hace referencia al tiempo. El coeficiente se hará mayor a medida que va evolucionando el sistema.

Energía externa

Uno de los problemas principales con los métodos basados en modelos deformables es el de la mala inicialización, lo que puede provocar que la superficie quede atrapada en mínimos locales. Una situación de este tipo se alcanzaría, por ejemplo, cuando intentando delinear la superficie externa del hueso, se posicionase el modelo inicial dentro del hueso cortical, o cuando se situase el modelo inicial en la parte exterior del hueso, pero próximo a la superficie de otra estructura. En un intento de mitigar estos problemas de mala inicialización, nosotros proponemos dos nuevos término de energía externa que añadir a los términos clásicos de gradiente e intensidad. Incorporando estos nuevos términos, la expresión del potencial externo será:

$$\begin{aligned} P(v(s,r)) = & - \gamma\left|\nabla[G(v(s,r)) * I(v(s,r))]\right| - \delta I(v(s,r)) \\ & - \zeta F_1(I(v(s,r)))\overrightarrow{n} - \eta F_2(t)\frac{\partial}{\partial \overrightarrow{n}} I(v(s,r)) \end{aligned} \tag{5.11}$$

donde $F_1(I(v(s,r)))$ es función de la intensidad en el punto $v(s,r)$; $F_2(t)$ es función del tiempo; y $\gamma, \delta, \zeta, \eta$ son constantes positivas. El tercer término corresponde a una fuerza normal hacia el exterior de la superficie, ponderada por la intensidad. El cuarto término provoca que la superficie se desplace hacia la superficie externa para la situación de estructuras con dobles contornos (hueso cortical), o hacia la superficie más próxima para el caso de la proximidad de otra estructura. Aquí $\overrightarrow{n}$ representa la normal a la superficie activa. Este cuarto término es del tipo usado con el modelo de los contornos deformables y su efecto es, por tanto, el de empujar la superficie deformable hacia la parte externa del hueso que rodea. La normal a una superficie se define mediante la expresión:

$$\overrightarrow{n}(s,r) = -\frac{v_s(s,r) \times v_r(s,r)}{||v_s(s,r) \times v_r(s,r)||} \tag{5.12}$$

en ella los subíndices denotan derivada parcial. Dado que definimos la superficie como una triangulación entre un conjunto de vértices, entonces aproximamos la normal por el promedio de las normales a cada uno de los triángulos incidentes en el vértice correspondiente v_i; o dicho de otra forma, como el promedio de los productos vectoriales entre los vectores $P_k(i)$ que unen el vértice actual con cada par de vértices vecinos, adyacentes en

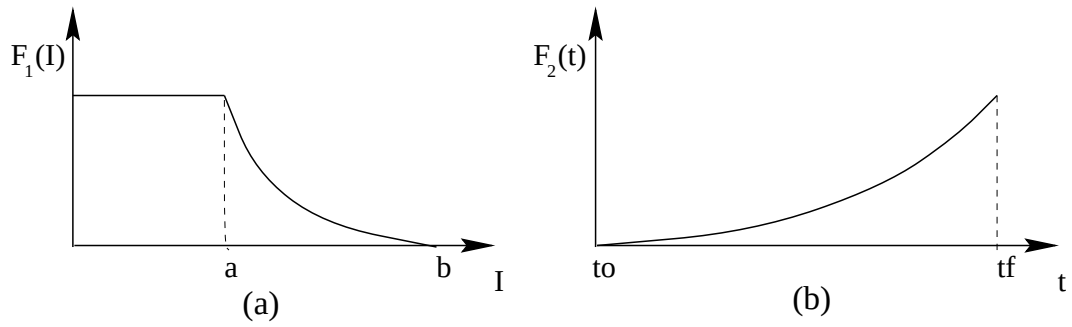


Figura 5.11: (a) Módulo de la función de deflación tomando como parámetro la intensidad; (b) variación del coeficiente (peso) de la función de derivada normal con el tiempo (iteraciones del proceso de minimización).

sentido angular:

$$\vec{n}_i = \frac{\sum_j P_j(i) \times P_{j+1}(i)}{||\sum_j P_j(i) \times P_{j+1}(i)||} \tag{5.13}$$

Para asegurar una buena segmentación 3D utilizamos una fuerza externa normal a la superficie y hacia su exterior que tienda a reducir problemas de inicialización en puntos interiores de la estructura a segmentar. Esta fuerza será función de la intensidad, otorgándole valores altos donde la intensidad es alta, y decayendo hasta cero a medida que la intensidad disminuye y se aproxima a valores propios de los tejidos no óseos. La figura 5.11.a muestra la dependencia del módulo de esta función con la intensidad. La importancia de la fuerza de la derivada normal es mayor, sin embargo, a medida que el modelo se aproxima a la superficie deseada. Esta es una fuerza que pretende mejorar la precisión de la superficie final. Por este motivo esta fuerza se pondera de tal forma que en las primeras etapas su peso sea muy bajo y que se incremente su contribución a medida que el número de iteraciones crece, de la forma que se muestra en la representación funcional de la figura 5.11.b.

Minimización de energía

En la sección 3.6.2 del capítulo 3 ya se introdujo el uso del algoritmo de Temple Simulado como técnica de optimización. Storvik [1994] propuso un método para la detección

Inicializar temperatura T y superficie S;

Repetir

Seleccionar a vértice V_i de S.

Mover V_i a uno de sus voxels vecinos: $S \rightarrow S^{'}$.

Repetir

Estimar la cantidad de cambio en la función de energía E;

Si $\Delta E \leftarrow E(S^{'}) - E(S) < 0$ aceptar la localización actual para V_i;

en otro caso $P = \min\{1, e^{-\Delta E/T}\}$;

si $Randon[0,1) < P$ entonces $S \leftarrow S^{'}$.

Hasta que se alcance el equilibrio.

Reemplazar T por $f(T)$ donde f es una función monótona decreciente.

Hasta $(T \rightarrow 0)$.

Tabla 5.1: Algoritmo de Temple Simulado para minimización de superficies deformables.

de curvas basado en un planteamiento Bayesiano y empleó este algoritmo para encontrar el mínimo global, haciendo que el comportamiento del modelo deformable fuese menos sensible a los malos posicionamientos iniciales. Nosotros también empleamos este algoritmo para la minimización del funcional de energía, adaptado a la situación actual tal y como se muestra en la tabla 5.1. Construimos un esquema iterativo donde los cambios en cada etapa se hacen localmente en la superficie. De esta forma la búsqueda del mínimo es dinámica. Una transición de la etapa $t^{(r)}$ a la etapa $t^{(r+1)}$ se define a través de dos pasos, uno en el que se define la posición a la que se va a desplazar el vértice, y otro que define el tipo de cambio.

El algoritmo de Temple Simulado evitará que la superficie deformable quede atrapada en mínimos locales. Cuando se trabaja con superficies deformables no se partirá, en general, de una superficie próxima a la deseada. Además, las imágenes médicas presentan el problema de baja relación señal/ruido, lo que provoca la existencia de un elevado número de mínimos locales. Estos inconvenientes pueden mitigarse mediante el uso de algoritmos que permitan cierto número de evoluciones aleatorias de la superficie que

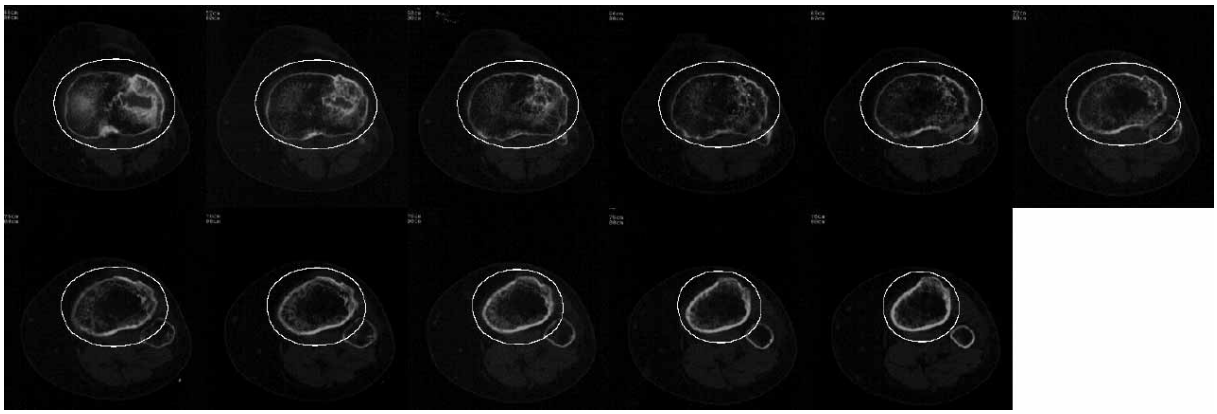


Figura 5.12: Superficie inicial definida a través de una elipse en cada uno de los cortes.

aumenten la energía del sistema. De esta forma, si se trata de un mínimo local poco pronunciado, se superará; sin embargo, si se trata de un mínimo de energía significativo, se volverá a caer en él. El inconveniente relacionado con el algoritmo de temple simulado es su coste computacional elevado. En principio se podría pensar en parametrizar la superficie utilizando modelos deformables paramétricos. Sin embargo, cuando se trata de superficies muy irregulares, como las que se pueden encontrar en imágenes médicas, se necesitaría tal número de parámetros que resultará preferible utilizar la descripción mediante puntos de superficie; esta descripción es capaz de manejar un amplio espectro de formas con un número limitado de puntos vértice.

Si se porporcionase una superficie inicial que fuese totalmente exterior, y no muy distante de la final deseada, no serían precisas muchas iteraciones. Por lo que se hará que la temperatura T del sistema decrezca rápidamente . De esta forma el sistema evolucionará en las proximidades del punto de congelación.

5.3.2 Ejemplo

Con objeto de comprobar la validez del método propuesto de superficies deformables hemos analizado una secuencia de imágenes correspondientes al paciente cuya secuencia final de contornos se mostraba en la figura 3.53. En la figura 5.12 se muestra la posición de la superficie inicial en diferentes cortes. Esta superficie inicial ha sido generada mediante

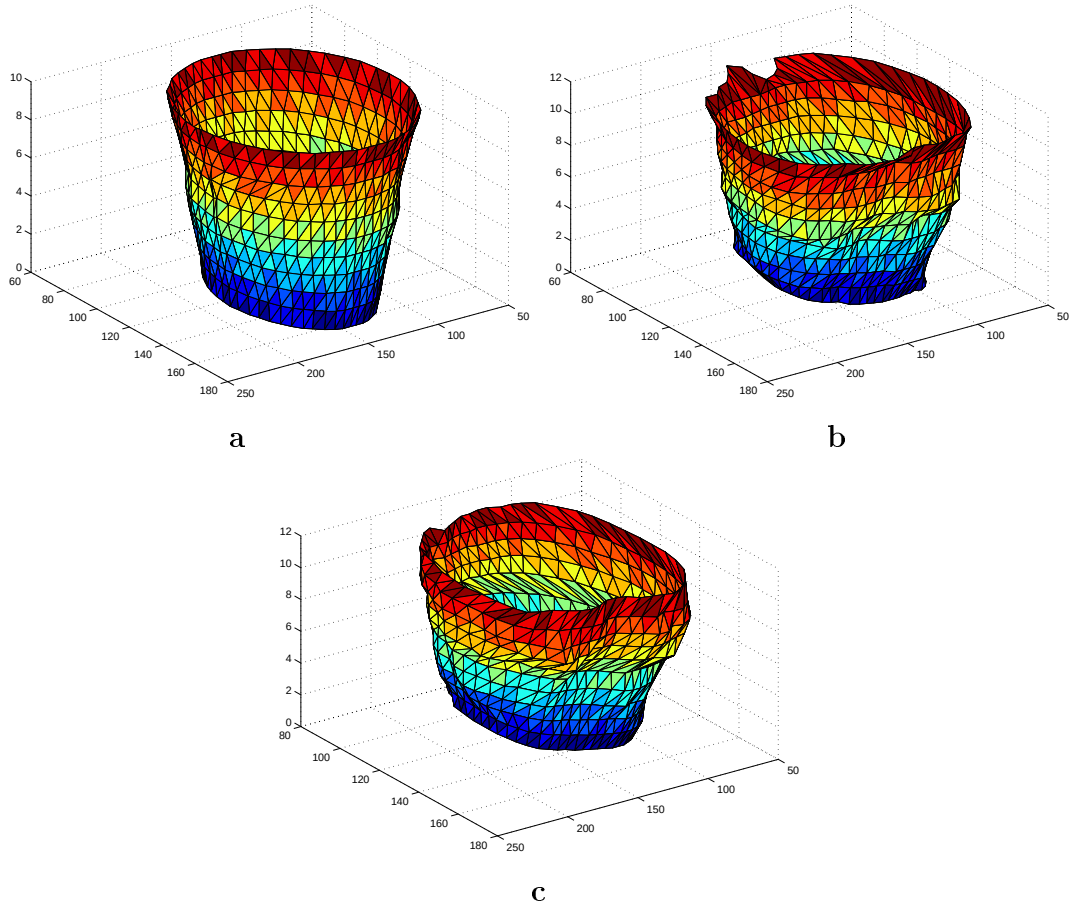


Figura 5.13: Superficies: (a) inicial para el método de superficies deformables, (b) resultado final de la deformación de la superficie, (c) resultado obtenido para la misma secuencia de cortes mediante el método de integración.

una serie de contornos elípticos conectados mediante triangulaciones, y que engloban a la estructura de interés. En la figura 5.13.a se muestra la superficie inicial, en la figura 5.13.b se muestra el resultado de la deformación de la superficie y en la figura 5.13.c aparece la superficie que resulta de la triangulación de los contornos obtenidos mediante el esquema de integración. Los resultados ofrecidos por el método de superficies y el de integración difieren sustancialmente. En la figura 5.14 se muestra el resultado de la segmentación en cada uno de los cortes.

La segmentación obtenida mediante superficies deformables dista de la deseada, pero

supone un primer paso en el que ya no se parte de un modelo aproximado, intentando suplir esta carencia de información con una mayor ligadura en la continuidad 3D. Por otra parte, en este esquema aún no se ha incluido información de homogeneidad (regiones), que va a resultar indispensable para la obtención de un sistema robusto. La forma en la que se ha de combinar información de homogeneidad y gradientes mediante la formulación de potenciales es uno de los objetivos de estudio más inmediatos.

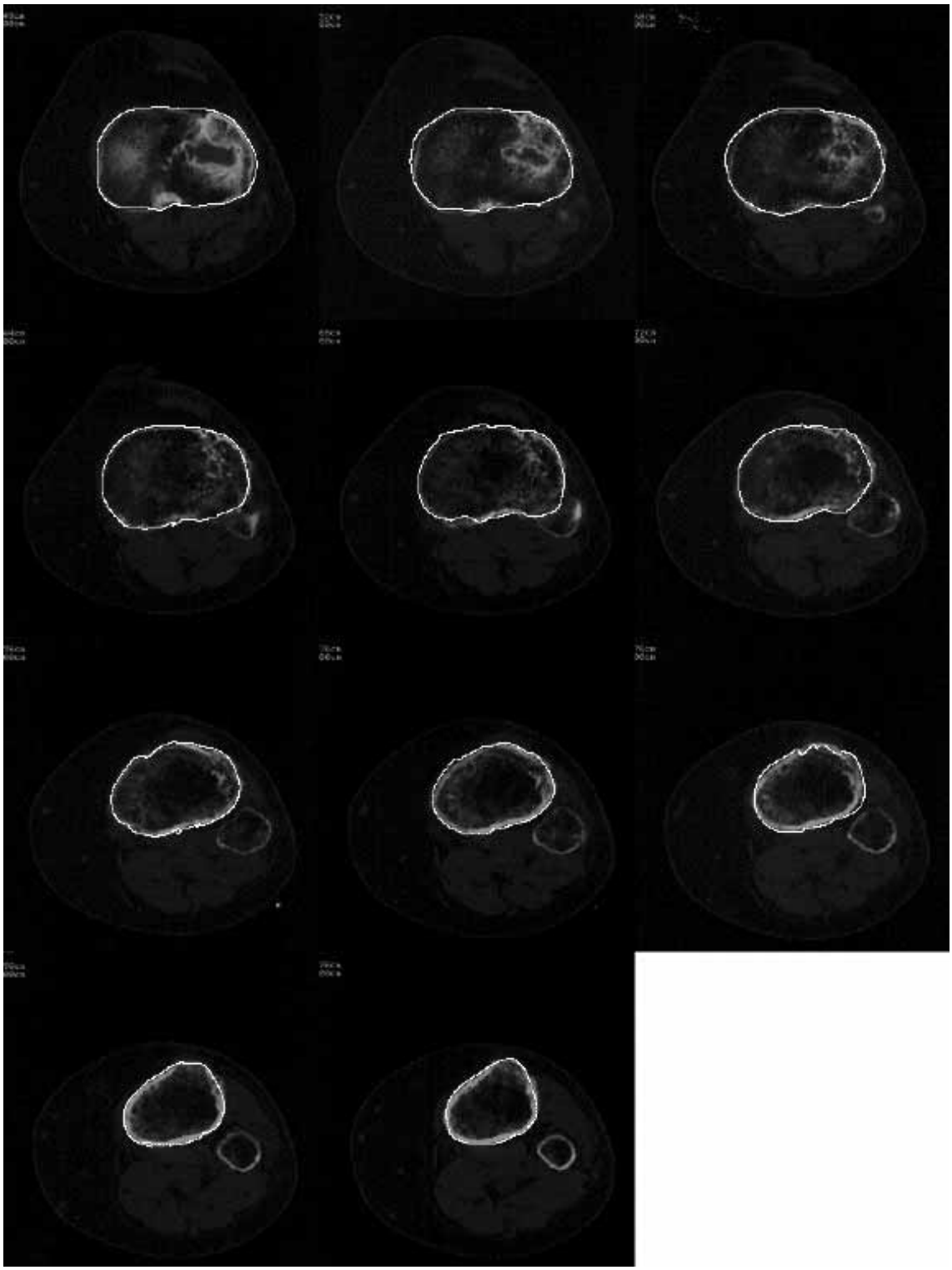


Figura 5.14: Posición de la superficie final en cada uno de los cortes.

Conclusiones y principales aportaciones

En esta memoria se ha abordado el problema de la extracción automática de la geometría 3D de la parte proximal de la tibia a partir de una secuencia de imágenes de cortes transversales paralelos obtenidos mediante Tomografía Computerizada. Esta tarea comprende tres etapas principales: delimitación del contorno exterior de la tibia corte a corte, detección de lesiones y obtención de la superficie 3D que define el modelo geométrico.

La primera de las tareas se ha abordado mediante un sistema basado en conocimiento para segmentación mediante unión y división de regiones, un sistema de segmentación basado en modelos deformables y un bloque de integración de las segmentaciones obtenidas por los dos sistemas anteriores.

El sistema de segmentación por unión y división consta de un bloque de bajo nivel, que genera una presegmentación inicial proporcionada por un clasificador neuronal y organiza esta información en un RAG, y un bloque de alto nivel que, a partir del RAG, completa la segmentación mediante la incorporación de conocimiento sobre el dominio a través de un conjunto de reglas para unión y división de regiones. El sistema basado en modelos deformables implementa una estrategia de segmentación basada en contornos, habiendo definido nuevos términos de energía adecuados al problema concreto. Estos dos sistemas presentan una serie de ventajas e inconvenientes que conjugamos en un bloque de integración, con el fin de obtener una segmentación robusta y fiable. El sistema basado en regiones ofrece mejor comportamiento en cuanto al reconocimiento de los objetos; es decir, en la determinación de la localización de los objetos y en la distinción con respecto de otras entidades. Sin embargo, el sistema basado en contornos activos permite obtener una definición más precisa de la localización espacial de los contornos del objeto, una vez

que se conoce su localización aproximada. La estrategia de integración explota estas dos características con el fin de obtener la mejor segmentación posible.

Una vez identificadas las estructuras de interés en cada una de las imágenes que constituyen la secuencia, se procede con la identificación de posibles lesiones. Esta tarea se aborda empleando un esquema basado en el formalismo de los Campos Aleatorios de Markov para el etiquetado de regiones y un sistema basado en reglas que refina su identificación inicial. El esquema basado en el modelo de los Campos Aleatorios de Markov se aplica corte a corte en una base 2D y, posteriormente, el sistema basado en reglas, aplicado en una base 3D, evalúa la consistencia de la clasificación de lesión efectuada anteriormente, efectuando a su vez una regularización de forma de las lesiones teniendo en cuenta información espacial en las 3 dimensiones.

Finalizados los procesos anteriores, hemos abordado la creación del modelo 3D a partir del apilamiento de los contornos 2D mediante la generación de una superficie envolvente conectando puntos de contorno por triangulación. A partir de este momento se dispone de un modelo geométrico apto para ser transferido a un sistema de análisis mediante Elementos Finitos.

El sistema ha sido implementado en su totalidad utilizando lenguaje C, a excepción de los procesos de visualización, para los que se ha empleado software libre: NUAGES y GEOMVIEW.

El comportamiento del sistema se ha evaluado utilizando diversas secuencias de imágenes de Tomografía Computerizada proporcionadas por el Servicio de Ciruxía Ortopédica del Complejo Hospitalario Universitario de Santiago de Compostela, así como secuencias extraidas de la base de datos de imágenes médicas del Mallinckrodt Institue of Radiology. Tres son los factores de mayor importancia para la evaluación de cualquier método de segmentación: velocidad, repetitibilidad y precisión. Desafortunadamente, no se puede establecer la segmentación ideal cuando se trabaja con imágenes adquiridas a partir de sujetos vivos. Por ello, la valoración de la precisión ha de hacerse por inspección visual. De ésta se deduce que los contornos obtenidos mediante nuestro sistema se corresponden con aquellos percibidos por un operador humano. Por otra parte, si definimos la velocidad de procesado (segmentación, detección de lesiones y generación de la superficie) como el número de cortes de la misma secuencia procesados por minuto, los valores obtenidos están en el rango de 2 a 3 minutos por corte. Estos tiempos mejoran los de cualquier

procesado semiautomático. Además, el proceso es totalmente repetible, porque en él no interviene el operador humano. Consideramos, por tanto, que el sistema ofrece ventajas definitivas respecto del procesado semiautomático, comúnmente usado en tareas complejas, al proporcionar resultados precisos, ser repetible y ahorrar cantidades de tiempo significativas.

Aún cuando este sistema proporciona ya una herramienta para la generación de un modelo geométrico 3D, para completar el sistema de análisis mediante elementos finitos es preciso añadir más módulos que se encarguen de realizar la interfaze entre nuestra descripción del modelo sólido y la admitida por el software concreto de elementos finitos, así como abordar el desarrollo de herramientas para el modelado de prótesis a partir de un conjunto de especificaciones y de su colocación sobre el modelo de tibia construido. A estas tareas dedicaremos nuestro esfuerzo en un futuro próximo.

Por otra parte, dentro de los propios procesos abordados en esta memoria, avanzaremos también hacia el desarrollo de un sistema de reconstrucción verdaderamente 3D con el que se intentará sacar todo el partido posible de la información espacial. Este sistema se orientará hacia superficies deformables en el que habrá que incluir información de regiones de una forma eficiente, de tal modo que se evite la división del sistema de segmentación en tres bloques diferentes. El trabajo con superficies deformables presenta los problema de determinación del modelo inicial aproximado y el ya mencionado de incorporación de información de regiones, tareas de compleja solución sobre todo para un sistema automático. A su vez la identificación de lesiones se intentará abordar, también, como un sistema completamente 3D.

Bibliografía

[Ahmed y col., 1990] A. M. Ahmed, M. Tissakht, S. C. Shrivastava, and K. Chan. Dynamics stress response of the implant/cement interface: an axisymetric analysis of a knee tibial component. *J. Orthop. Res.*, 8(3):435–447, 1990.

[Allota y col., 1995] B. Allota, P. Dario, M. Marcacci, and S. T. Larse. Experiments towards the design of machtronic tools for orthopaedic surgery. In *Proc. of the Ninth World Congress on the Theory of Machines and Mechanisms*, volume 3, pages 2156–2160, Milano, 1995.

[Amini y col., 1990] A. A. Amini, T. E. Weymouth, and R. C. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Trans. Patt. Anal. Machine. Intell.*, 12(9):855–867, september 1990.

[Ayache y col., 1990] N. Ayache, J. D. Boissonnant, L. Cohen, B. Geiger, J. Levy-Vehel, O. Monga, and P. Sander. Steps toward the automatic interpretation of 3D images. In K. H. Höhne, editor, *3D imaging in medicine*, volume F60 of *NATO ASI Series*. Springer-Verlag, Berlin Heidelberg, 1990.

[Ayache, 1991] N. Ayache. *Artificial vision for mobile robots: stereo vision and multisensory perception*. MIT Press, 1991.

[Bae y col., 1992] K. T. Bae, M. L. Giger, C. T. Chen, and C. E. Kahn,Jr. Automatic segmentation of liver structures in CT images. *Med. Phys.*, 20(1):71–78, Jan/Feb 1992.

[Baker, 1990] H. H. Baker. Surface modeling with medical imagery. In K. H. Höne, editor, *3D Imaging in Medicine*, volume F60 of *NATO ASI*, pages 277–288. Springer Verlag, 1990.

[Ballard y Brown, 1982] D. H. Ballard and C. M. Brown. *Computer Vision.* Prentice–Hall, New Jersey, 1982.

[Banerjee y Majundar, 1996] S. Banerjee and D. D. Majumdar. A 2D shape metric and its implementation in biomedical imaging. *Pattern Recognition Letters*, 17:141–147, 1996.

[Bardinet y col., 1996] E. Bardinet, L.D. Cohen, and N. Ayache. Analyzing the deformation of the left ventricle of the heart with a parametric deformable model. Technical Report 2797, INRIA, Sophia Antipolis, France, 1996.

[Barequet y Sharir, 1996] G. Barequet and M. Sharir. Picewise–linear interpolation between polygonal slices. *Computing Vision and Image Understanding*, 63(2):251–272, 1996.

[Barillot, 1993] C. Barillot. Surface and volume rendering techniques to display 3-D data. *IEEE Engineering in Medicine and biology*, pages 111–119, March 1993.

[Besl y McKay, 1992] P. J. Besl and N.-N. McKay. A method for regristration of 3-D shapes. *IEEE Trans. Patt. Anal. Machine. Intell.*, 14:239–256, 1992.

[Bezdek, 1981] J. C. Bezdek. *Pattern recognition with fuzzy objective function algorithms.* Plenum Press, New York, 1981.

[Bezdek y col., 1993] J. C. Bezdek, L. O. Hall, and L. P. Clarke. Review of MR image segmentation techniques using pattern recognition. *Med. Phys.*, 20(4):1033–1048, 1993.

[Blake y Yuille, 1993] A. Blake and A. Yuille, editors. *Active vision*, Artificial Intelligence. MIT Press, 1993.

[Boissonant, 1988] J. D. Boissonnat. Shape reconstruction from planar cross sections. *Comput. Vision Graphics Image Process.*, 44:1–29, 1988.

[Bomans y col., 1990] M. Bomans, K. Höhne, U. Tiede, and M. Riemer. 3-D segmentation of the head for 3-D display. *IEEE Transactions on Medical Imaging*, 9(2):177–183, June 1990.

[Breau y col., 1991] C. Breau, A. Shirazi-Adl, and J. de Guise. Reconstruction of a human ligamentous lumbar spine using CT images–A three-dimensional finite element mesh generation. *Annals of Biomedical Engineering*, 19:291–302, 1991.

[Brooks y col., 1979] R. A. Brooks, R. Greiner, and T. O. Binford. The ACRONYM model-based vision system. In *Proc. IJCAI-6*, pages 105–113, Tokio, Japan, August 1979.

[Brooks, 1983] R. A. Brooks. Model-based three-dimensional interpretations of two-dimensional images. *IEEE Trans. Patt. Anal. Machine. Intell.*, 5(2):140–150, March 1983.

[Brunie y col., 1991] L. Brunie, P. Cinquin, and S. Miguet. Three–dimensional image of the wrist. In H. U. Lenke, M. L. Rhodes, C. C. Jaffe, and R. Felix, editors, *Proceedings of the International Symposium CAR'91*, pages 549–555. Spring-Verlag, 1991.

[Cabello y col., 1993] D. Cabello, M. G. Penedo, S. Barro, J. M. Pardo, and J. Heras. CT image segmentation by self–organazing learning. In J. Mira, J. Cabestany, and A. Prieto, editors, *New trends in neural computation*, Lecture Notes in Computer Science, volume 686, pages 651–656. Spring-Verlag, 1993.

[Canny, 1986] J. Canny. A computational approach to edge detection. *IEEE Trans. Patt. Anal. Machine. Intell.*, PAMI-8(6):679–698, November 1986.

[Chakraborty y col., 1994] A. Chakraborty, L. H. Staib, and J. S. Duncan. Deformable boundary finding influenced by region homogeneity. In *Proc. Computer Vision and Pattern Recognition*, pages 624–627, 1994.

[Chakraborty y col., 1996] A. Chakraborty, L. H. Staib, and J. S. Duncan. Deformable boundary finding in medical images by integrating gradient and region information. *IEEE Transactions on Medical Imaging*, 15(6):859–870, 1996.

[Chen y Huang, 1988] H. H. Chen and T. S. Huang. A survey of construction and manipulation of octrees. *Computer Vision, Graphics, and Image Processing*, 43:409–431, 1988.

[Chen y col., 1992] S. Y. Chen, K. R. Hoffman, C. T. Chen, and J. D. Carroll. Modelling human heart based on cardiac tomography. *SPIE Imaging Technoligies and Applications*, 1778:14–18, 1992.

[Christiansen y Sederberg, 1978] H.-N. Christiansen and T. W. Sederberg. Conversion of complex contour lines definitions into polygonal element mosaics. *Comput. Graphics*, 13:187–192, 1978.

[Cho y Meer, 1997] K. Cho and P. Meer. Image segmentation from consensus information. *Computer Vision and Image Understanding*, 68(1):72–89, 1997.

[Chow y Kaneko, 1972] C. K. Chow and T. Kaneko. Boundary detection of radiographic images by a thresholding method. In S. Wanatabe, editor, *Frontiers of Pattern Recognition*, pages 61–82, New York, 1972. Academic Press.

[Chu y Aggarwal, 1993] C.-C. Chu and J. K. Aggarwal. The integration of image segmentation maps using region and edge information. *IEEE Trans. Patt. Anal. Machine. Intell.*, 15(12):1241–1252, December 1993.

[Clift, 1992] S. E. Clift. Finite–element analysis in cartilage biomechanics. *J. Biomed. Eng.*, 14:217–220, May 1992.

[Cohen y Cohen, 1993] L. D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Trans. Patt. Anal. Machine. Intell.*, 15(11):1131–1147, November 1993.

[Coleman y Andrews, 1979] G. B. Coleman and H. C. Andrews. Image segmentation by clustering. In *Proc. IEEE*, volume 67, pages 773–785, May 1979.

[Cootes y col., 1992] T. F. Cootes, C. J. Taylor, J. Cooper, and D. H. Graham. Training models of shape form sets of examples. In D. C. Hogg and R. Boyle, editors, *Proceedings of the British Machine Vision Conference*, pages 9–18, London, 1992. Springer–Verlag.

[Crevier y Lepage, 1997] D. Crevier and R. Lepage. Knowledge-based image understanding systems: a survey. *Computer Vision and Image Understanding*, 67(2):161–185, 1997.

[Dario y col., 1996] P. Dario, C. Paggetti, T. Ciucci, D. Bertelli, B. Allota, M. Marcacci, M. Fadda, S. Martelli, and D. Caramella. A system for computer–asisted articular surgery. *Journal of Computer Aided Surgery*, pages 48–49, 1996.

[Declerck y col., 1995] J. Declerck, G. Subsol, J. P. Thirion, and N. Ayache. Automatic retrieval of anatomical structures in 3D medical images. Technical Report 2485, INRIA, Sophia–Antipolis Cedex (France), 1995.

[Delingette, 1994] H. Delingette. Simplex meshes: a general representation for 3D shape reconstruction. Technical Report 2214, INRIA, Sophia–Antipolis Cedex (France), 1994.

[Dhawan y Juvvadi, 1990] A. P. Dhawan and S. Juvvadi. Knowledge-based analysis and understanding of medical images. *Computer Methods and Programs in Biomedicine*, 33:221–239, 1990.

[Dhawan y Arata, 1991] A. P. Dhawan and L. Arata. Knowledge-based 3D analysis from 2D medical images. *IEEE Engineering in Medicine and Biology*, 10(4):30–37, December 1991.

[Dhawan y Arata, 1993] A. P. Dhawan and L. Arata. Segmentation of medical images through competitive learning. *Computer Methods and Programs in Biomedicine*, 40:203–215, 1993.

[Duda y Hart, 1973] R. O. Duda and P. E. Hurt. *Pattern Clasification and Scene Analysis*. Wiley-Interscience, New York, 1973.

[Efford, 1994] N. D. Efford. Knowledge–based segmentation and feature analysis of hand–wrist radiographs. Technical Report 94.31, School of computer studies. University of Leeds, 1994.

[Ehricke, 1991] H. H. Ehricke. Automated 3D MR image analysis with a ruled based segmentation system, 1991.

[Farin, 1993] G. Farin. *Curves and surfaces for CAGD*. Academic Press, 1993.

[Friedland y Rosenfeld, 1992] N. S. Friedland and A. Rosenfeld. Compact object recognition using energy–function–based optimization. *IEEE Trans. Patt. Anal. Machine. Intell.*, 14(7):770–777, 1992.

[Frost, 1987] R. A. Frost. *Introduction to Knowledge Base Systems.* Collins, London, 1987.

[Fuchs y col., 1977] H. Fuchs, Z. M. Kedem, and S. P. Uselton. Optimal surface reconstruction from planar contours. *Commun. ACM*, 20:693–702, 1977.

[Garbay, 1995] C. Garbay. Knowledge adquisition and representation. In J. D. Bronzino, editor, *The Biomedical Engineering Handbook*, pages 2731–2745. CRC Press, IEEE Press, 1995.

[Geiger, 1995] D. Geiger, A. Gupta, L. A. Costa, and J. Vlontzos. Dynamic programming for detecting, tracking, and matching deformable contours. *IEEE Trans. Patt. Anal. Machine. Intell.*, 17(3):294–302, 1995.

[Geman y Geman, 1984] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images. *IEEE Trans. Patt. Anal. Machine. Intell.*, 6(6):721–741, 1984.

[Goldwasser y col., 1986] S. M. Goldwasser, D. Reynolds, R. A. Talton, and E. Walsh. Real-time display and manipulation od 3D CT, PET, and NMR data. In *SPIE proc.*, pages 139–149, 1986.

[González y Wintz, 1987] R. C. González and P. Wintz. *Digital image processing.* Addison–Wesley, 1987.

[Goshtasby y col., 1992] A. Goshtasby, D. A. Turner, and L. V. Ackerman. Matching of tomographics slices for interpolation. *IEEE Transactions on Medical Imaging*, 11(4):507–516, December 1992.

[Grimson, 1990] W. E. L. Grimson. *Object recognition by computer: the role of geometric constraints.* MIT Press, 1990.

[Haddon y Boyce, 1990] J. F. Haddon and J. F. Boyce. Image segmentation by unifying region and boundary information. *IEEE Trans. Patt. Anal. Machine. Intell.*, 12(10):929–948, October 1990.

[Haralick y Shapiro, 1992] R. M. Haralick and L. G. Shapiro. *Computer and Robot Vision*, volume I–II. Addison-Wesley, 1992.

[Hart y col., 1992] R. T. Hart, Z. M. Oden, S. W. Parrish, and D. B. Burr. Computational methods for the biomechanics studies. *The international journal of supercomputer applications*, 6(2):164–174, Summer 1992.

[Hecquard y Acharya, 1991] Jean Hecquard and Raj Acharya. Connected component labeling with linear octree. *Pattern Recognition*, 24(6):515–531, 1991.

[Herman y col., 1992] G. T. Herman, J. Zheng, and C. A. Bucholtz. Shape-based interpolation. *IEEE Comput. Graph. Applic.*, 12(3):69–79, 92.

[Herman, 1995] G. T. Herman. Image reconstruction from projections. *Real–Time Imaging*, 1:3–18, 1995.

[Hong y Rosenfeld, 1984] T. H. Hong and A. Rosenfeld. Compact region extraction using weighted pixel linking in a pyramid. *IEEE Trans. Patt. Anal. Machine. Intell.*, 6(2):222–229, 1984.

[Hoover y col., 1996] A. Hoover, G. Jean-Baptiste, X. Jiang, P.J. Flyn, H. Bunke, D.B. Goldgof, K. Bowyer, D.W. Eggert, A. Fitzgibbon, and R.B. Fisher. An experimental comparison of range image segmentation algorithms. *IEEE Trans. Patt. Anal. Machine. Intell.*, 18(7):673–689, 1996.

[Huertas y Medioni, 1986] A. Huertas and G. Medioni. Detection of intensity changes with subpixel accuracy using laplacian-gaussian masks. *IEEE Trans. Patt. Anal. Machine. Intell.*, PAMI-8(5):651–664, Septmber 1986.

[Huiskes y Chao, 1983] R. Huiskes and E. Y. S. Chao. A survey of finite element analysis in orthopedic biomechanics: the first decade. *J. Biomechanics*, 16(6):385–409, 1983.

[Hull y col., 1990] M. E. C. Hull, J. H. Frazer, and R. J. Millar. Octree-based modelling of computed-tomography images. *IEEE Proceedings*, 137, pt. I(3):118–122, June 1990.

[Ibaroudene y col., 1990] V. Ibaroudene, D. Demjanenko and R. S. Acharya. Adjacency algoritms for lineat octree nodes. *Image and Vision Computing*, 8(2):115–123, May 1990.

[Jackins y Tanimoto, 1980] C. L. Jackins and S. L. Tanimoto. Oct-Trees and their use in representing three dimensional objects. *Computer Graphics And Image Processing*, 14:249–270, 1980.

[Jain y col., 1995] R. Jain, R. Kasturi, and B. G. Schunck. *Machine Vision*. McGraw–Hill, New York, 1995.

[Kang y col., 1993] Y. K. Kang, H.C. Park, Y. Youm, I. K. Lee, M. H. Ahn, and J.C. Ihn. Three dimensional shape reconstruction and finite element analysis of femur before and after the cementless type of total hip replacement. *J. Biomed. Eng.*, 15:497–504, November 1993.

[Kankanhalli y col., 1993] M. S. Kankanhalli, C. P. Yu, and R. C. Krueger. Volume modelling for orthopedic surgery. *IFIP Transactions B (Applications in Technology)*, B-9:303–315, 1993.

[Kass y col., 1988] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision*, 1:321–331, 1988.

[Keppel, 1975] E. Keppel. Approximating complex surfaces by triangulation of contour lines. *IBM J. Res. Devel.*, 19:2–11, 1975.

[Keyak y col., 1990] J. H. Keyak, J. M. Meagher, H. B. Skinner, and C. D. Mot,Jr. Automated three-dimensional finite element modelling of bone: a new method. *J. Biomed. Eng.*, 12:389–397, September 1990.

[Keyak y Skinner, 1992] J. H. Keyak and H. B. Skinner. Three-dimensional finite element modelling of bone: effects of element size. *J. Biomed. Eng.*, 14:483–489, November 1992.

[Keyak y col., 1993] J. H. Keyak, M. G. Fourkas, J. M. Meagher, and H. B. Skinner. Validation of an authomated method of three-dimensional finite element modelling of bone. *J. Biomed. Eng.*, 15:505–509, 1993.

[Kim y Yang, 1996] I. Y. Kim and H. S. Yang. An interpretation scheme for image segmentation and labeling based on Markov random field model. *IEEE Trans. Patt. Anal. Machine. Intell.*, 18(1):69–73, 1996.

[Kirkpatrick y col., 1983] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[Kittler y Illingworth, 1985] J. Kittler and J. Illingworth. Threshold selection based on a simple image statistics. *Computer Vision, Graphics, and Image Processing*, 30, 1985.

[Klaus y Horn, 1997] B. Klaus and P. Horn. *Robot vision*. MIT Press, 11 edition, 1997.

[Kohonen, 1982] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cibernetics*, 43:59–62, 1982.

[Kübler y Gerig, 1990] O. Kübler and G. Gerig. Segmentation and analysis of multidimensional data-sets in medicine. In K. H. Höhne, editor, *3D imaging in medicine*, volume F60 of *NATO ASI Series*. Spring-Verlag, Berlin Heidelberg, 1990.

[Lai y Chin, 1994] K. F. Lai and R. T. Chin. On regularization, formulation and initialization of the active contour models (snakes). In *Proc. 1st Asian Conference on Computer Vision*, pages 542–545, 1994.

[Leahy y col., 1989] R. Leahy, T. Herbert, and R. Lee. Application of Markov Random Fields in medical imaging. *Information Processing in Medical Imaging*, pages 1–14, 1989.

[Lei y Sewchand, 1992] T. Lei and W. Sewchand. Statistical approach to X-Ray CT imaging and its applications in image analysis -Part II: A new stochastic model-based image segmentation technique for X-Ray CT image. *IEEE Transactions on Medical Imaging*, 11(1):62–29, March 1992.

[Levine y Nazif, 1985] Martin D. Levine and Ahmed M. Nazif. Dynamic measurement of computer generated image segmentation. *IEEE Trans. Patt. Anal. Machine. Intell.*, PAMI-7(2):155–164, March 1985.

[Levoy, 1988] M. Levoy. Display of surfaces from volume data. *IEEE Comput. Graphics and Appl.*, pages 29–37, 1988.

[Li, 1995] S. Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer–Verlag, Tokio, 1995.

[Liang, 1993] Zhengrong Liang. Tissue clasification and segmentation of MR images. *IEEE Engineering in Medicine and Biology*, pages 81–85, March 1993.

[Lin y col., 1989] W. C. Lin, S. Y. Chen, and C. T. Chen. A new surface interpolation technique for reconstruction 3-D objects from serial cross-sections. *Comput. Vision, Graph Image Processing*, 48:124–143, 1989.

[Loncaric y col., 1995] S. Loncaric, A. P. Dhawan, J. Broderick, and T. Brott. 3-D image analysis of intra-cerebral brain hemorrhage from digitized CT films. *Computer Methods and Programs in Biomedicine*, 46:207–216, 1995.

[Lorensen y Cline, 1987] W. E. Lorensen and H. E. Cline. Marching cubes: a high resolution 3D surface construction algorithm. *ACM Computer Graphics*, 21(4):163–169, july 1987.

[Lu y Jain, 1989] Y. Lu and R. C. Jain. Behaviour of edges in scale space. *IEEE Trans. Patt. Anal. Machine. Intell.*, 11(4):337–356, April 1989.

[Lu y Jain, 1992] Y. Lu and R. C. Jain. Reasoning about edges in scale space. *IEEE Trans. Patt. Anal. Machine. Intell.*, 14(4):450–468, April 1992.

[Ma y Overton, 1991] X. Q. Ma and T. R. Overton. Automated image analysis for bone density measurements using computed tomography. *IEEE Transactions on Medical Imaging*, 10(4):611–615, December 1991.

[Malandin y col., 1995] G. Malandain, S. Fernadez-Vidal, and J.M. Rocchisani. Physically based rigid registration of 3-D free-form objects: application to medical imaging. Technical Report 2453, INRIA, Sophia Antipolis, France, 1995.

[Marr y Hildreth, 1980] D. Marr and E. Hildreth. Theory of edge detection. *Proc. R. Soc. Lond. B*, 207:187–217, 1980.

[McInerney y Terzopoulos, 1995] T. McInerney and D. Terzopoulos. Topologically adaptable snakes. In *Proceedings ICCV,95*, pages 840–845, Berlin, Germany, 1995.

[McInerney y Terzopoulos, 1996] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medial Image Analysis*, 1(2):91–108, 1996.

[Meagher, 1982] D. Meagher. Geometric modelling using octree encoding. *Compt. Graphics Image Process.*, 19:129–147, 1982.

[Meagher, 1987] D. J. Meagher. The manipulation, analysis and display of 3-D medical objects using octree encoding techniques. *Innov. Tech. Biol. Med.*, 8(n. special 1):22–36, 1987.

[Mencl, 1995] R. Mencl. A graph–theoretic approach to surface reconstruction. Technical Report 568/1995, Universität Dormund, 1995.

[Metaxas y Terzopoulos, 1993] D. Metaxas and D. Terzopoulos. Shape and nonrigid motion estimation through physics–based synthesis. *IEEE Trans. Patt. Anal. Machine. Intell.*, 15(6):580–591, 1993.

[Metaxas y Kakadiaris, 1996] D. Metaxas and A. Kakadiaris. Elastically adaptive deformable models. In *Proceedings of the Fourth Europena Conference on Computer Vision*, volume II, pages 550–559, 1996.

[Michael y Nelson, 1989] David J. Michael and Alan C. Nelson. HANDX: A model-based system for automatic segmentation of bones from digital hand radiographs. *IEEE Transactions on Medical Imaging*, 8(1):64–69, March 1989.

[Modestino y Zhang, 1992] J. W. Modestino and J. Zhang. A Markov random field model–based approach to image interpretation. *IEEE Trans. Patt. Anal. Machine. Intell.*, 14(6):606–615, 1992.

[Müller y Rüegsegger, 1995] R. Müller and P. Rüegsegger. Three-dimensional finite element modelling of non-invasively assessed trabecular bone structures. *Med. Eng. Phys.*, 17(2):126–133, 1995.

[Münch y Rüegsegger, 1990] B. Münch and P. Rüegsegger. Quantitative evaluation of the human knee in 3D. *SPIE Close-range photogrametry meets machine vision*, 1395(pt 1):580–586, 1990.

[Murphy y col., 1986] S. B. Murphy, P. K. Kiejewski, S. R. Simon, H. P. Chandler, P. P. Griffin, D. T. Reilly, B. L. Peneberg, and M. M. Landy. Computer-aided simulation, analysis, and design in orthopedic surgery. *Orthopedic Clinics of North America*, 17(4):637–649, October 1986.

[Nanning y Jianquin, 1992] Zheng Nanning and Liu Jianquin. Visual knowledge representation and intelligent image segmentation. *Journal of computer science and technology*, 7(3):219–225, July 1992.

[Narayan y col., 1994] S.Ñarayan, D. Sensharma, E. M. Santori, A. Lee, a. A. Sabherwal, and A. W. Toga. Animated visualization of a high resolution color three dimensional digital computer model of the whole human head. *Year Book of Medical Informatics*, pages 482–501, 1994.

[Nazif y Levine, 1984] Ahmed M. Nazif and Martin D. Levine. Low-level image segmentation: an expert system. *IEEE Trans. Patt. Anal. Machine. Intell.*, PAMI-6(5):555–577, 1984.

[Ney y Fisman, 1991] D. R. Ney and E. K. Fisman. Edition tools for 3D medical imaging. *IEEE Computer Graphics & Applications*, pages 63–71, November 1991.

[Ohta, 1985] Y. Ohta. *Knowledge–based interpretation of outdoor natural color scenes.* Pitman Publishing, Boston, 1985.

[Otsu, 1979] N. Otsu. A threshold selection method for gray–level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.

[Pal y Pal, 1993] N. R. Pal and S. K. Pal. A review on image segmentation techniques. *Pattern Recognition*, 26:1277–1294, 1993.

[Panda y Rosenfeld, 1978] D. P. Panda and A. Rosenfeld. Image segmentation by pixel clasification in (gray level, edge value) space. *IEEE Transctions on Computers*, C–27:875–879, 1978.

[Pankanti y Jain, 1995] S. Pankanti and A. K. Jain. Integrating vision modules: stereo, shading, grouping, and line labeling. *IEEE Trans. Patt. Anal. Machine. Intell.*, 17(8):831–842, 1995.

[Pao, 1989] Y.-H. Pao. *Adaptive Pattern Recognition and Neural Networks.* Addison–Wesley, 1989.

[Pardo y col., 1995a] J. M Pardo, D. Cabello, M. J. Carreira, M. G. Penedo, and J. Heras. Knowledge-based CT image analysis: automatic 3D shape reconstruction

of bones. In *Proc. of the Second Asian Conference on Computer Vision*, volume I, pages 504–508, 1995.

[Pardo y col., 1995b] J. M Pardo, D. Cabello, and J. Heras. Automatic 3D shape reconstruction of bones. In *Basic & Applied Biomedical Engineering Building Blocks for Health Care*, 17th Annual International Conference of the IEEE Engineering in Medicine and Biology Society and 21st Canadian Medical and Biological Engineering Conference, pages 387–388, 1995.

[Pardo y col., 1997a] J. M Pardo, D. Cabello, M.J. Carreira, and J. Heras. Integrating region and edge information in CT image segmentation. In A. Sanfeliu, J.J. Villanueva, and J. Vitria, editors, *Pattern Recognition and Image Analysis*, volume 1, pages 7–12. Centre de Visió per Computador / Universitat Autónoma de Barcelona, 1997.

[Pardo y col., 1997b] J. M Pardo, D. Cabello, and J. Heras. A Markov Random Field model for bony tissue classification. In A. Del Bimbo, editor, *Image Analysis and Processing*, Lecture Notes in Computer Science, volume 1311, pages 364–371. Springer-Verlag, 1997.

[Pardo y col., 1998a] J. M Pardo, D. Cabello, and J. Heras. A snake for model-based segmentation of biomedical images. *Pattern Recognition Letters*, 1998 (aceptado).

[Pardo y col., 1998b] J. M Pardo, D. Cabello, J. Heras, and J.Couceiro. A Markov Random Field model for bony tissue classification. *Computerized Medical Imaging and Graphics*, 1998 (aceptado).

[Parrot y col., 1993] R. W. Parrott, M. R. Stytz, P. Amburn, and D. Robinson. Towards statistically optimal interpolation for 3D medical imaging. *IEEE Engineering in Medicine and Bioligy*, pages 49–59, September 1993.

[Pavlidis, 1977] T. Pavlidis. *Structural pattern recognition*. Springer–Verlag, Berlin, 1977.

[Pavlidis y Liow, 1990] Theo Pavlidis and Yuh-Tay Liow. Integration region growing and edge detection. *IEEE Trans. Patt. Anal. Machine. Intell.*, 12(3):225–233, March 1990.

[Pepino y col., 1993] E. Pepino, M. Cesarelli, F. Di Salle, A. Mosca, and M. Bracale. Preliminary notes on the conversion of three-dimensional images of bones into CAD structures for the pre-operative planning of intertrochantheric osteotomies. *Medical and Biological Engineering and Computing*, 31:529–534, 1993.

[Pratt, 1991] W. K. Pratt. *Digital Image Processing.* Wiley, New York, 1991.

[Press y col., 1992] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C.* Cambridge University Press, 1992.

[Pung y col., 1994] T. Pun, G. Gerig, and O. Ratib. Image analisys and computer vision in medicine. *Computerized Medical Imaging and Graphics*, 18(2):85–96, March/April 1994.

[Radeva y Serrat, 1993] P. Radeva and J. Serrat. Rubber snake: implementation on signed distance potential. In *Vision Conferences SWISS'93*, pages 187–194, september 1993.

[Radeva y col., 1995] P. Radeva, J. Serrat, and E. Marti. A snake for model–based segmentation. In *Proceedings of International Conference on Computer Vision*, 1995.

[Radeva y col., 1997] P. Radeva, A. A. Amini, and J. Huang. Deformable B–solids and implicit snakes for 3D localization and tracking of SPAMM MRI data. *Computer Vision and Image Understanding*, 66(2):163–178, 1997.

[Rakotomanana y col., 1992] R.L. Rakotomanana, P.F. Leyvraz, A. Curnier, J.H. Heegaard, and P.J. Rubin. A finite element model for evaluation of tibial prosthesis-bone interface in total knee replacement. *J. Biomechanics*, 25(12):1413–1424, 1992.

[Rao y Jain, 1988] A. R. Rao and R. jain. Knowledge representation and control in computer vision systems. *IEEE Expert*, 3(3):64–79, 1988.

[Raya, 1990] S. P. Raya. Low-level segmentation od 3-D magnetic resonance brain images—a ruled-based system. *IEEE Transactions on Medical Imaging*, 9(3):327–337, september 1990.

[Richardt y col., 1996] J. Richardt, F. Karl, C. Müller, and R. Klette. The fuzzy local-global duality in detecting pictorial patterns. *Pattern Recognition Letters*, 17:187–195, 1996.

[Robb, 1995] R. A. Robb. *Three-dimensional biomedical imaging*. VCH, 1995.

[Rosenfeld y Kak, 1982] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*. Academic Press, New York, 1982.

[Roux y Caotrieux, 1990] Christian Roux and Jean-Louis Coatrieux. Some trends in 3D medical imaging. In L. Torres and M.A. Masgrau, E. andLagunas, editors, *Signal Processing V: Theories and Applications*. Elsevier Science Publishers B.V., 1990.

[Samet, 1990a] H. Samet. *Applications of spatial data structures*. Addison–Wesley, 1990.

[Samet, 1990b] H. Samet. *The design and anlysis of spatial data structures*. Addison–Wesley, 1990.

[Sandor y Leahy, 1997] S. Sandor and R. Leahy. Surface-based labeling of cortical anatomy using a deformable atlas. *IEEE Transactions on Medical Imaging*, 16(1):41–54, 1997.

[Schereppers y Sauren, 1990] G. Schereppers and A. Sauren. A finite element model for the tibio-femoral contact area in the knee joint. *Finite Element News*, 6:29–34, December 1990.

[Seitz y Rüegsegger, 1983] P. Seitz and P. Rüegsegger. Fast contour detection for high precision quantitative CT. *IEEE Transactions on Medical Imaging*, MI-2(3):136–141, September 1983.

[Sjöberg y Bergholm, 1988] F. Sjöberg and F. Bergholm. Extraction of diffuse edges by edge focusing. *Pattern Recognition Letters*, 7:181–190, March 1988.

[Sloan y Painter, 1988] K. R. Sloan and J. Painter. Pessimal guesses may be optimal: a conterintuitive search result. *IEEE Trans. Patt. Anal. Machine. Intell.*, 10:949–955, 1988.

[Sonka y col., 1996] M. Sonka, W. Park, and E. A. Hoffman. Rule-based detection of intrathoracic airway trees. *IEEE Transactions on Medical Imaging*, 15(3):314–326, June 1996.

[Staib y Duncan, 1992] L. H. Staib and J. S. Duncan. Boundary finding with parametrically deformable models. *IEEE Trans. Patt. Anal. Machine. Intell.*, 14(11):1061–1075, 1992.

[Staib y Duncan, 1992a] L. H. Staib and J. S. Duncan. Deformable Fourier models for surface finding in 3–d images. In *Proc. SPIE Visual. Biomed. Comput.*, volume SPIE–1808, pages 90–104, 1992.

[Stansfield, 1986] S. A. Stansfield. ANGY: a rule-based expert system for automatic segmentation of coronary vessels from digital substracted angiogramas. *IEEE Trans. Patt. Anal. Machine. Intell.*, 8(2):188–199, March 1986.

[Steele y col., 1994] J. R. Steele, A. Basu, and A. Job. A three–dimensional representation of an athletic female knee joint using magnetic resonance imaging. *Med. Eng. Phys.*, 16:363–369, September 1994.

[Storvik, 1994] G. Storvik. A bayesian approach to dynamic contours through stochastic sampling and simulated annealing. *IEEE Trans. Patt. Anal. Machine. Intell.*, 16(10):976–986, 1994.

[Strasters y Gerbrands, 1991] K. C. Strasters and J. J. Gerbrands. Three-dimensional image segmentation using a split, merge and group approach. *Pattern Recognition Letters*, 12:307–325, May 1991.

[Sutherland, 1986] C. J. Sutherland. Practical aplication of computer-generated three-dimensional reconstruction in orthopedic surgery. *Orthopedic Clinics of North America*, 17(4):651–656, October 1986.

[Tagare y col., 1995] H. D. Tagare, F. M. Vos, C. C. Jaffe, and J. S. Duncan. Arrangement: a spatial relation between parts for evaluating similarity of tomographic section. *IEEE Trans. Patt. Anal. Machine. Intell.*, 17(9):880–893, 1995.

[Tagare, 1997] H. D. Tagare. Deformable 2-D template matching using orthogonal curves. *IEEE Trans. Patt. Anal. Machine. Intell.*, 16(1):108–117, 1997.

[Taratorin y Sideman, 1993] Alexander M. Taratorin and Samuel Sideman. Constrained detection of left ventricular boundaries from cine CT images of human hearts. *IEEE Transactions on Medical Imaging*, 12:521–533, September 1993.

[Tek y Kimia, 1997] H. Tek and B.B. Kimia. Volumetric segmentation of medical images by three-dimensional bubbles. *Computer Vision and Image Understanding*, 65(2):246–258, 1997.

[Tensi y col., 1989] H. M. Tensi, H. Gese, and R. Ascherl. Non-linear three-dimensional finite element analysis of a cementless hip endoprosthesis. In *Part H:Journal of Engineering in Medicine*, volume 203, pages 215–222, 1989.

[Terzopoulos, 1987] D. Terzopoulos. On matching deformable models to images: direct and iterative solutions. In *Topical Meeting on Machine Vision*, volume 12 of *Techical Digest*, pages 160–167. Optical Society of America, 1987.

[Terzopoulos y col., 1988] D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: recovering 3D shape and nonrigid motion. *Artificial Intelligence*, 36:91–123, 1988.

[Torre y Poggio, 1986] V. Torre and T. A. Poggio. On edge detection. *IEEE Trans. Patt. Anal. Machine. Intell.*, PAMI-8(2):147–163, March 1986.

[Udupa y Herman, 1991] J. K. Udupa, G. T. Herman. *3D Imaging in Medicine.* CRC Press, 1991.

[Udupa y col., 1991] J. K. Udupa, H. M. Hung, and K. S. Chuang. Surface and volume rendering in 3D imaging: a comparison. *Journal of Digital Imaging*, 4(3):159–168, 1991.

[Vilariño y col, 1997] D.L. Vilariño, D. Cabello, A. Mosquera, and J.M. Pardo. Application of a multilayer discrete-time CNN to deformable models. In J. Mira, R. Moreno-Díaz, and J. Cabestany, editors, *Biological and Artificial Computation: From Neuroscience to Technology*, volume 1240, pages 1193–1202. Springer-Verlag, Berlin Heidelberg, 1997.

[Vilariño y col., 1998a] D. L. Vilariño, V. M. Brea, D. Cabello, and J. M. Pardo. Active contours with cellular neural networks: a new approach. In *Proceedings of the IASTED International Conference Signal Processing and Communications*, pages 427–430, 1988.

[Vilariño y col., 1998b] D. L. Vilaiño, D. Cabello, M. Balsi, and V. M. Brea. Image segmentation based on active contours using dicrete time cellular neural networks. In *5th IEEE International Workshop on Cellular Neural Networks and their Applications*, 1988 (Aceptado).

[Vilariño y col., 1998c] D. L. Vilariño, V. M. Brea, D Cabello, and J. M Pardo. Discrete-time CNN for image segmentation by active contours. *Pattern Recognition Letters*, 1998(aceptado).

[Wang y Aggawal, 1986] Y. F. Wang and J. K. Aggawal. Surface reconstruction and representation of 3-D scenes. *Pattern Recognition*, 19:197–207, 1986.

[Weska y col., 1974] J. S. Weska, R.-N. Nagel, and A. Rosenfeld. A threshold selection technique. *IEEE Transactions on Computers*, C–23:1322–1326, 1974.

[Wood, 1986] S. L. Wood. Surface definition of 3D objects from CT images. In *Proc. 8th Ann. Conf. IEEE Eng. Med. Biol. Soc.*, volume 2, pages 1118–1121, 1986.

[Wu y Leahy, 1993] Zhenyu Wu and Richard Leahy. An optimal graph theoretic approach to data clustering: theory and its applications to image segmentation. *IEEE Trans. Patt. Anal. Machine. Intell.*, 15(11):1101–1113, November 1993.

[Yau y Srihari, 1983] M. Yau and S.-N. Srihari. A hierarchical data structure for multi-dimensional digital images. *Communications of ACM*, 26(7):504–515, July 1983.

[Youla y Webb, 1982] D. C. Youla and H. Webb. Image restoration by the method of convex projections: Part 1–theory. *IEEE Transactions on Medical Imaging*, 1:81–94, 1982.

[Zhang, 1997] Y. J. Zhang. Evaluation and comparison of different segmentation algorithms. *Pattern Recognition Letters*, 18:963–974, 1997.

[Zhu y col., 1991] D. Zhu, R. Conners, and P. Araman. 3-D CT image segmentation by volume growing. *SPIE, Visual Communications and Image Processing'91: Image Processing*, 1606(pt 1):685–692, 1991.

[Zhu y Yan, 1997] Y. Zhu and H. Yan. Computerized tumor boundary detection using a Hopfield neural network. *IEEE Transactions on Medical Imaging*, 16(1):55–67, 1997.

[Zienkiewicz y Taylor, 1989] O. C. Zienkiewicz and R. L. Taylor. *The finite element method.* McGraw–Hill, New York, N Y, 1989.