

Segmentación No Paramétrica de Imágenes  
Cerebrales de Resonancia Magnética

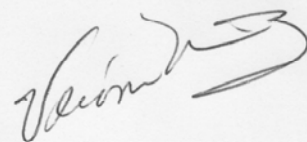
M. en C. Juan Ramón Jiménez Alaniz

Tesis presentada para obtener el grado de  
Doctor en Ciencias (Ingeniería Biomédica)



Laboratorio de Investigación en NeuroImagenología  
Departamento de Ingeniería Eléctrica  
Universidad Autónoma Metropolitana  
Iztapalapa

Dirigida bajo la supervisión de:  
Dra. Verónica Medina Bañuelos



Junio de 2006

**Segmentación No Paramétrica de Imágenes  
Cerebrales de Resonancia Magnética**

**M. en C. Juan Ramón Jiménez Alaniz**

Tesis presentada para obtener el grado de  
Doctor en Ciencias (Ingeniería Biomédica)



**Casa abierta al tiempo**

Laboratorio de Investigación en NeuroImagenología  
Departamento de Ingeniería Eléctrica  
Universidad Autónoma Metropolitana  
Iztapalapa

Dirigida bajo la supervisión de:  
**Dra. Verónica Medina Bañuelos**

Junio de 2006

## Agradecimientos

Una de las cosas que me había propuesto en la vida, fue estudiar una carrera universitaria y lo hice. Doy gracias a Dios por haberme permitido llegar mucho más allá de lo que me había propuesto o imaginado y hacer un doctorado.

Quiero agradecer a mi esposa Briss su apoyo y empuje para que continuara estudiando; a mis hijos Kaori y Juanito por no haberme reclamado el tiempo que les robé para dedicarlo al estudio, muchas gracias. Sin el amor, comprensión y apoyo de mi familia no podría haberlo hecho.

También quiero expresar mi agradecimiento a mi asesora por guiarme y acompañarme durante toda la investigación doctoral. Fue de gran valor su orientación durante el curso de la investigación, la procuración de los medios y recursos necesarios para el proyecto, así como el ambiente propicio para la discusión de dudas y aportación de ideas.

Aprecio mucho también el enorme apoyo y ayuda que en todo tiempo recibí de Oscar Yáñez, gracias por tus palabras de animo cuando el camino a seguir era incierto y difícil.

Agradezco a Raquel Valdés por escucharme y darme su punto de vista cuando necesitaba expresar alguna idea y recibir una opinión al respecto. Muchas gracias a todos mis compañeros del Laboratorio de Neuroimagenología por propiciar un foro de discusión de nuestros proyectos de investigación.

Le doy las gracias al Consejo Nacional de Ciencia y Tecnología (CONACyT), por el apoyo económico que me otorgó, al aceptarme como becario (registro 162482) de su programa de becas para estudios de posgrado. Este apoyo es de gran ayuda y valor para formar recursos humanos especializados que contribuyen en el desarrollo científico y tecnológico de nuestro país.

¿De dónde, pues, procede la sabiduría y dónde se encuentra el lugar de la inteligencia? Y dijo al hombre: "El temor del Señor es la sabiduría, y el apartarse del mal, la inteligencia."

Job

# Índice general

<b>Resumen</b>	<b>4</b>
<b>Introducción</b>	<b>6</b>
<b>1. Antecedentes en segmentación de imágenes cerebrales</b>	<b>10</b>
1.1. Métodos y aplicaciones . . . . .	13
1.1.1. Extracción de características . . . . .	14
1.1.2. Segmentación de imagen de una sola escala de gris . . . . .	15
1.1.3. Segmentación multiespectral . . . . .	15
1.1.3.1. Métodos de segmentación supervisados . . . . .	16
1.1.3.2. Segmentación no supervisada . . . . .	17
1.1.4. Otros métodos . . . . .	18
1.1.5. Clasificación . . . . .	20
1.1.6. Aplicaciones . . . . .	20
1.2. Corrimiento de media . . . . .	24
1.3. Planteamiento del problema y principales contribuciones . . . . .	31
<b>2. Estimación por corrimiento de media</b>	<b>33</b>
2.1. Estimación de densidad . . . . .	33
2.1.1. Métodos univariados . . . . .	34
2.1.2. Medidas de discrepancia . . . . .	37
2.1.3. Método de kernel multivariado . . . . .	40
2.1.4. Convergencia para el kernel de Epanechnikov . . . . .	43
2.1.5. Agrupamiento por corrimiento de media . . . . .	45
2.2. Información espacial . . . . .	46
2.2.1. Dominio espacial-rango . . . . .	46
2.2.2. Filtrado . . . . .	47
2.2.3. Segmentación por corrimiento de media . . . . .	48

2.3. Corrimiento de media ponderado . . . . .	49
<b>3. Mapas de confianza de bordes</b>	<b>50</b>
3.1. Medida de confianza . . . . .	50
3.2. Estimación del gradiente . . . . .	52
3.3. Mapa de confianza . . . . .	55
<b>4. Fusión de regiones</b>	<b>56</b>
4.1. Contracción del grafo . . . . .	56
4.2. Algoritmo encontrar y unir . . . . .	57
4.3. Podado . . . . .	60
<b>5. Incorporación de información a priori</b>	<b>61</b>
5.1. Normalización espacial . . . . .	61
5.1.1. Normalización afín . . . . .	62
5.1.2. Normalización espacial no lineal . . . . .	65
5.2. Clasificación . . . . .	66
5.2.1. Probabilidad de pertenencia . . . . .	67
5.2.2. Etiquetado de las regiones . . . . .	69
5.3. Resumen . . . . .	70
<b>6. Resultados y discusión</b>	<b>74</b>
6.1. Datos sintéticos . . . . .	74
6.1.1. Mapa de confianza . . . . .	76
6.1.2. Filtrado . . . . .	77
6.1.3. Análisis de adyacencia . . . . .	79
6.2. Cerebro normal simulado . . . . .	83
6.2.1. Parámetros del proceso . . . . .	83
6.2.2. Filtrado del volumen simulado . . . . .	99
6.2.3. Análisis de adyacencias de región . . . . .	100
6.2.4. Registro de datos y mapas de probabilidad . . . . .	101
6.2.5. Clasificación de regiones por probabilidad de pertenencia . . . . .	103
6.3. Datos reales . . . . .	110
6.3.1. Filtrado por CM . . . . .	110
6.3.2. Análisis de adyacencias . . . . .	110
6.3.3. Registro de datos y mapas de probabilidad . . . . .	112
6.3.4. Clasificación de regiones . . . . .	114

6.4. Segmentación volumétrica . . . . .	118
<b>7. Conclusiones</b>	<b>123</b>
<b>A. Trabajos publicados derivados de la investigación doctoral</b>	<b>127</b>
<b>B. Mapa de confianza</b>	<b>128</b>
<b>C. Filtrado por corrimiento de media</b>	<b>131</b>
<b>D. Análisis de adyacencia de regiones</b>	<b>133</b>
D.1. Etiquetar imagen filtrada . . . . .	133
D.2. Fusión de regiones por contracción de un grafo de adyacencias . . . . .	134
D.3. Podado de la imagen fusionada . . . . .	138
<b>E. Clasificación de regiones</b>	<b>142</b>
E.1. Determinar probabilidad de pertenencia de cada región . . . . .	142
E.2. Etiquetado final de las regiones . . . . .	143
<b>F. Índice de Dice</b>	<b>146</b>
<b>Bibliografía</b>	<b>148</b>

# Resumen

La segmentación de las estructuras cerebrales, a partir de imágenes obtenidas por Resonancia Magnética, se aplica en el estudio de muchas enfermedades asociadas con el cerebro, así como en su diagnóstico, tratamiento, planeación quirúrgica y mapeo funcional de la actividad cerebral. La segmentación se lleva a cabo empleando una técnica no paramétrica de estimación de densidad, basada en el algoritmo de corrimiento de media aplicado en el dominio conjunto de espacio-rango. La implementación del corrimiento de media está guiada totalmente por los datos, por lo que no es necesario usar parámetros de inicialización o conocer el número de clases a segmentar para calcular el corrimiento de media. La calidad de las fronteras entre regiones se mejora integrando un mapa de confianza de bordes, que representa la confianza de estar verdaderamente ante la presencia de una frontera entre regiones adyacentes, así el corrimiento de media se mejora ponderando cada dato por una función de su confianza de borde. De esta manera, las técnicas de segmentación por región y detección de bordes se combinan para la extracción de bordes débiles pero significativos. El procedimiento estima las modas locales de la función de densidad de probabilidad para definir los centros de clase en el espacio característico, dando como resultado una imagen filtrada con preservación de bordes, conteniendo regiones de intensidad homogénea. Después del filtrado de la imagen, se construye un grafo de adyacencias que se analiza para fusionar las regiones adyacentes que están separadas por un borde de magnitud pequeña, calculada a partir del mapa de confianza, y su diferencia en modas de intensidad es menor de un valor predeterminado. Posteriormente, la imagen fusionada se somete a un proceso de podado, con el propósito de unir las regiones que tienen pocos píxeles, con la región adyacente que tiene un número suficiente de píxeles y una diferencia mínima en intensidad.

Para clasificar las regiones de la imagen a un tipo de tejido cerebral, se realiza una normalización espacial entre los datos de la imagen y los mapas de probabilidad estándar existentes para materia gris, materia blanca y líquido cefalorraquídeo, de modo que para cada estructura cerebral se aplica un criterio de probabilidad máxima para determinar la región semilla correspondiente. Finalmente, se clasifican las regiones asignándoles la

etiqueta de la semilla con la que tengan una diferencia en modas de intensidad mínima.

El método se aplicó tanto a imágenes sintéticas como reales, en 2 y 3 dimensiones espaciales, manteniendo todos los parámetros constantes durante todo el proceso para cada tipo de datos. La combinación de la segmentación por región y detección de bordes probó ser una técnica robusta, ya que las clases se detectaron adecuadamente y de manera automática, sin importar el nivel de ruido e inhomogeneidad. Comparando los resultados con las segmentaciones de referencia, los índices de similitud promedio obtenidos para los datos sintéticos fueron de 0.9 a 0.99 y de 0.59 a 0.99 para datos reales, considerando solamente la materia gris, la materia blanca y fondo.



# Introducción

La imagenología por Resonancia Magnética (IRM) se ha convertido en una de las modalidades de imágenes médicas más importantes. Las ventajas de la IRM sobre otras modalidades de imágenes de diagnóstico son su alta resolución espacial, su excelente discriminación de tejidos suaves [62], y el hecho de ser virtualmente no invasiva, significa que es particularmente buena para estudios de órganos como el cerebro [35]. La IRM proporciona un medio altamente efectivo para observar la anatomía cerebral. El análisis morfométrico proporciona mediciones cuantitativas de localización, volumen, forma y homogeneidad de componentes de las estructuras cerebrales. Este tipo de análisis, en conjunción con observaciones neurosicológicas, neurológicas, psiquiátricas y acopladas con neuroimágenes funcionales se puede usar para responder interrogantes acerca de la estructura y función cerebral, tanto de sujetos sanos como enfermos. La segmentación de imágenes del cerebro es de interés en el estudio de muchos desórdenes, tales como la esclerosis múltiple, esquizofrenia, epilepsia, mal de Parkinson, enfermedad de Alzheimer, atrofia cerebral o asimetría, etc. Otras aplicaciones de la segmentación incluyen: simulación y planeación de cirugías, medición de volumen de tumor, mapeo funcional, registro de imágenes, identificación de áreas anatómicas de interés para diagnóstico, y correlación mejorada de áreas anatómicas de interés con métricas funcionales localizadas [16].

En la IRM, los tejidos están caracterizados por intensidades de voxel resultado de señales de Resonancia Magnética (RM) diferenciadas espacialmente. Por lo tanto, la meta de la segmentación es dividir una imagen en regiones significativas, donde los voxels de cada región deben ser homogéneos de acuerdo a una propiedad específica. Dado que los voxels de un tejido están conectados espacialmente, el problema de la segmentación es encontrar las propiedades o estadísticas del tejido y etiquetar los voxels.

Es común segmentar imágenes de IRM utilizando el reconocimiento de patrones. Estas técnicas han sido exitosas particularmente para imágenes de cerebro [16]. En el reconocimiento de patrones estadístico, usualmente se define un modelo paramétrico suponiendo formas de distribución particular del espacio de características seleccionado.

Las aproximaciones paramétricas están basadas en la suposición de formas específicas de distribuciones de probabilidad para las características del voxel en la imagen. Un modelo paramétrico usado con frecuencia en la segmentación es el de mezcla finita (FM) [9, 40], suponiendo que los datos se distribuyen normalmente [53, 61]; y también son de uso frecuente las técnicas de expectation-maximization (EM) [45, 62], y probabilidad *máxima a posteriori* (MAP) [46], para el cálculo de parámetros.

Los métodos paramétricos sólo son útiles cuando las distribuciones de características para las diferentes clases son bien conocidas, o pueden ser adecuadamente modeladas. Sin embargo, los datos de RM generalmente no soportan las suposiciones base del método paramétrico [9]. Esto es, la distribución de las clases de tejido no es necesariamente una distribución estadística conocida, y esto con frecuencia resulta en estimaciones que producen segmentaciones pobres. Los métodos no paramétricos, no confían en distribuciones predefinidas, sino en las distribuciones actuales de los datos, y no requieren ningún conocimiento o suposición acerca de sus propiedades estadísticas [9, 16]. Los métodos no paramétricos no tienen la restricción de pertenecer a una forma funcional conocida; la motivación de la aproximación no paramétrica es que cuando se tiene una distribución de datos, de la que no se puede discernir una forma funcional conocida, entonces es posible permitir que los datos decidan cuál es la función que mejor los ajusta sin la restricción impuesta por el modelo paramétrico.

En las técnicas estadísticas, se intenta representar los tejidos por un conjunto finito de parámetros. En este caso, la segmentación se obtiene estimando los parámetros del modelo, permitiendo identificar las clases dentro de la imagen. Sin embargo, el número de clases (i.e., el número de tejidos) con frecuencia es un conjunto elevado de valores que pueden causar problemas de estabilidad y convergencia. La estabilidad y convergencia se pueden mejorar usando algoritmos de agrupamiento (clustering) que son no paramétricos. En tales métodos, el problema es encontrar agrupaciones (clusters) de voxeles con intensidades homogéneas, i.e., encontrar los centros de las agrupaciones y valores de membresía de cada uno de los voxeles para los diferentes agrupamientos [53]. El problema de encontrar las agrupaciones homogéneas se resuelve aplicando el algoritmo de corrimiento de media (CM), que usa las modas locales de la función de densidad fundamental para definir los centros de las agrupaciones [33, 13, 19].

La determinación de las agrupaciones homogéneas, se combina de manera simultánea, con una técnica de detección de bordes para mejorar la calidad de la segmentación [15]. La detección de bordes hace uso de una medida de confianza, que se determina a partir de un mapa de bordes, para evaluar si efectivamente se está ante la presencia de un borde, y si en realidad existe una delimitación en las fronteras de la región ho-

mogénea. La medida de confianza también se utiliza para fusionar regiones que tienen bordes débiles entre ellas, a través de un proceso iterativo sobre un grafo de adyacencia de regiones (GAR). La etapa de fusión se termina con un proceso de podado para eliminar las regiones muy pequeñas.

En esta investigación se propone la aplicación del algoritmo de CM y la medida de confianza para segmentar imágenes de cerebro de IRM. La aplicación de esta técnica produce una segmentación en diferentes clases, cuyo número pudiera ser, dependiendo del propósito de la aplicación, mayor al esperado o al deseado. Cuando el proceso obtiene la imagen segmentada en regiones homogéneas, falta aún por realizarse la asociación de clase con los diferentes tipos de tejido. Ya que la clasificación automática puede ser muy complicada cuando no hay interacción con un operador, y se pretende que en la segmentación final solo las estructuras anatómicas principales del cerebro: materia gris, materia blanca y líquido cefalorraquídeo estén contenidas, se propone el uso de un atlas digital del cerebro con los mapas probabilísticos correspondientes a las estructuras de interés, para llevar a cabo la clasificación final de la imagen, utilizando una regla de asociación basada en el cálculo de probabilidades y agrupación de regiones alrededor de su región semilla más cercana.

El principal interés de usar un atlas digital es la disponibilidad de conocimiento *a priori*, que es explícitamente proporcionado por el atlas, a través de la distribución completa de probabilidades de que un voxel pertenezca a una o más estructuras cerebrales. El enfoque sobre atlas probabilísticos ha continuado, debido a que los atlas traen más información *a priori* al proceso de definición de órganos complejos [50], y se pueden usar para encontrar la membresía de clase más probable de los voxeles. Antes de usar la información *a priori* del atlas, se debe realizar el registro entre el atlas y el volumen a ser segmentado para igualar sus espacios. El registro se realiza a través de un proceso de normalización espacial que minimiza las diferencias cuadradas residuales entre la imagen y el atlas. El atlas utilizado es el mismo que se usa en [45, 44, 47].

A continuación se da una breve descripción del contenido de esta tesis. El Capítulo 1 describe los antecedentes de la segmentación de imágenes cerebrales, los problemas que tiene asociados y algunos métodos que se han aplicado. Se justifica el uso del corrimiento de media como parte principal de la metodología de segmentación, culminando con el planteamiento del problema a resolver. En el Capítulo 2 se explican las bases de la estimación de la función de densidad de probabilidad, estableciendo que el estimador de densidad de kernel multivariado es la mejor elección práctica, como método no paramétrico, para localizar los máximos de la función de densidad local, por medio del algoritmo de corrimiento, y utilizarlos como los centros de las agrupaciones en la

imagen segmentada.

En el Capítulo 3 se presentan los mapas de confianza como alternativa para resolver el problema de extraer una dirección de borde confiable. La incorporación de la medida de confianza mejora la estimación por CM de los máximos locales de la función de densidad, permitiendo alcanzar un filtrado espacial de alta calidad conservando los bordes.

El Capítulo 4 muestra cómo se realiza la fusión de las regiones filtradas, a través de un proceso iterativo sobre un grafo de adyacencia de regiones, así como la reducción del grafo utilizando un algoritmo de encontrar y unir. La fusión final de las regiones se refina aplicando un proceso de podado de las regiones muy pequeñas.

El Capítulo 5 trata acerca de la manera de incorporar información *a priori* en la clasificación de las regiones de la imagen a un tipo de tejido cerebral. También se trata la etapa de registro, entre el estudio de IRM del cerebro y un atlas cerebral, requerida para poder emplear el conocimiento *a priori* en una regla de clasificación de las regiones.

En el Capítulo 6 se muestran y se discuten los resultados obtenidos al aplicar la técnica de segmentación propuesta a dos tipos de datos: sintéticos y reales; en el Capítulo 7 se establecen las conclusiones obtenidas de este proyecto de investigación, y finalmente, se incluyen varios apéndices en los que se muestran las publicaciones derivadas de esta investigación y la descripción de las funciones que se emplean en el proceso de segmentación.

# Capítulo 1

## Antecedentes en segmentación de imágenes cerebrales

La segmentación es la separación de estructuras de interés del fondo y de entre ellas mismas, es una función de análisis esencial para la que se han desarrollado muchos algoritmos en el campo del procesamiento de imágenes. En imágenes médicas, la segmentación es importante para la extracción de características, mediciones y despliegue de imágenes. En algunas aplicaciones se busca clasificar los píxeles de la imagen en regiones, tales como huesos, músculos y vasos sanguíneos, mientras que, en otras, en regiones patológicas, como cáncer, deformaciones de tejido y esclerosis múltiple. En estudios de IRM del cerebro el objetivo es dividir la imagen en subregiones de materia gris (MG), materia blanca (MB) y líquido cefalorraquídeo (LCR).

La segmentación de un objeto se logra ya sea identificando todos los píxeles o voxeles que pertenecen al objeto o localizando aquellos que forman su frontera. El primero de estos métodos está basado principalmente en la intensidad de los píxeles, pero otros atributos, como la localización espacial, pueden ser asociados con el píxel para realizar la segmentación, cuyo objetivo principal es dividir una imagen en regiones, también llamadas clases, que son homogéneas con respecto a una o más características. Así, las técnicas de segmentación más comúnmente usadas se pueden clasificar en dos amplias categorías: técnicas de *segmentación por región*, que buscan grupos de píxeles que satisfacen un criterio de homogeneidad dado, y técnicas de *segmentación basadas en borde*, que buscan las fronteras entre regiones con diferentes características.

La delineación manual de MG, MB y LCR en imágenes de IRM por un experto humano es difícil porque requiere un conocimiento anatómico detallado y un examen cuidadoso de grandes cantidades de datos, su realización manual es tediosa, consume mucho tiempo, y por lo tanto es propensa a errores debido a la complejidad de las

estructuras cerebrales; mostrando probablemente una gran variabilidad intra e interobservador. Sin embargo, la ventaja de la técnica manual es que refleja la interpretación del radiólogo, que es la única verdad válida disponible. Por otra parte, una técnica de segmentación automática presenta ventajas si se compara con la técnica manual, las principales son: mejora de la repetibilidad, mejora de la confiabilidad y disminuye la variabilidad. Los resultados serán más reproducibles porque los métodos automatizados producirán los mismos resultados para los mismos datos. La confiabilidad aumenta porque los errores debidos a la fatiga se eliminan. La variabilidad puede disminuirse con las técnicas automáticas mientras se conserva alta la confiabilidad. Por lo anterior, el diseño de métodos de segmentación automáticos y confiables es altamente deseable.

En IRM, los tejidos están caracterizados por las intensidades del voxel, como resultado de los diferentes tiempos de relajación que presentan. Así, la segmentación se puede definir como el encuentro de las fronteras entre regiones con intensidades homogéneas. Dado que los voxeles de un tejido están conectados espacialmente, el problema de la segmentación se puede reformular para encontrar las estadísticas o propiedades del tejido y etiquetar los voxeles de manera apropiada [53].

La segmentación de una imagen no sólo es la decisión de colocar cada voxel dentro de una clase de tejido general, sino que se debe etiquetar cada voxel como perteneciente a una estructura neuroanatómica específica. El problema para realizar esta tarea es que con frecuencia no se dispone de suficiente información dentro de la imagen para segmentar completamente una estructura; y la principal dificultad, que deben enfrentar los métodos de segmentación, es que las diferentes estructuras del cerebro, que corresponden al mismo tipo de tejido, presentan intensidades similares en las imágenes de RM, de esta manera, las intensidades de señal de diferentes estructuras cerebrales, con frecuencia se traslapan como puede apreciarse en la Figura 1.1. El uso de técnicas automáticas para segmentar IRM de cerebro se dificulta, adicionalmente, por la posibilidad de encontrar artefactos en la imagen. Un artefacto es una característica que aparece en una imagen, la cual no está presente en la imagen original; algunas veces es el resultado de una operación impropia del equipo, y otras veces es consecuencia de procesos naturales o propiedades del cuerpo humano. Los artefactos puede ser muy notables o sólo incluir unos cuantos pixeles, pero pueden dar apariencias que se confunden con patologías que pueden ser mal diagnosticadas. Algunos ejemplos de artefactos que pueden ser encontrados en la imagen, son los siguientes:

- **Inhomogeneidad del campo magnético estático ( $B_0$ ).** Todas las imágenes por RM suponen un campo magnético estático homogéneo, ya que un campo no homogéneo originará imágenes distorsionadas. La distorsión puede ser ya sea

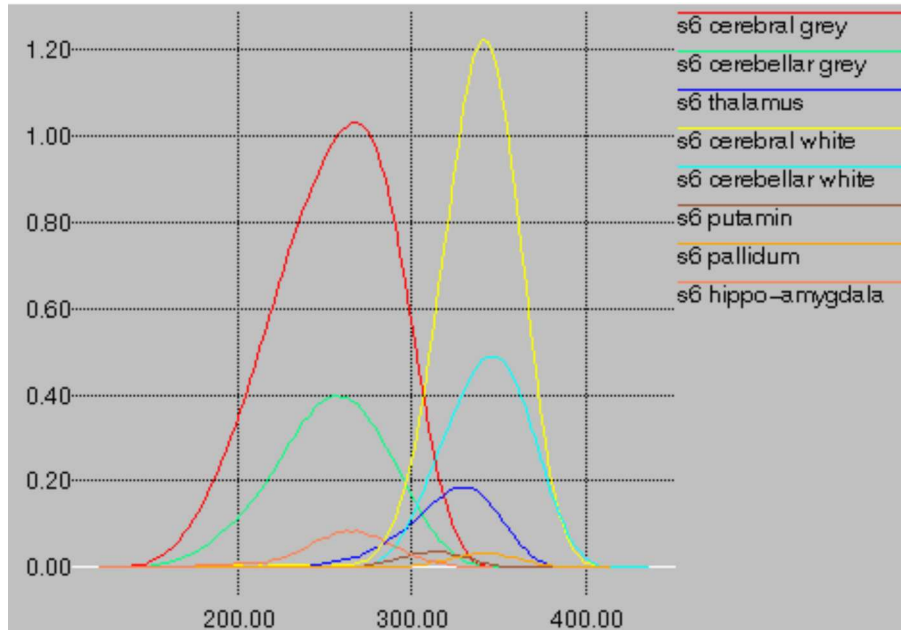


Figura 1.1: Intensidades de diferentes estructuras del cerebro. Las intensidades de las estructuras cerebrales presentan traslape entre ellas. Por ejemplo, el tálamo es una estructura de materia gris central que tiene una intensidad promedio intermedia porque es una mezcla de materia blanca y gris. Tomada de [39].

espacial, en intensidad o ambas. Las distorsiones de intensidad resultan de la homogeneidad del campo en una ubicación que es mayor o menor que en el resto de la imagen. La distorsión espacial resulta de variaciones en la frecuencia de resonancia que afectarán a la relación frecuencia-espacio. En consecuencia se producirá distorsión en la dirección de selección de plano y en el de lectura pero no en la dirección de diferencia de fase si se usa una técnica de codificación en fase para la reconstrucción de imágenes.

- **Gradiente.** Los artefactos que surgen de problemas con el sistema de gradiente algunas veces son muy similares a las inhomogeneidades de  $B_0$ . Un gradiente que no es constante con respecto a la dirección del gradiente distorsionará la imagen. Esto típicamente sólo es posible si una bobina de gradiente se ha dañado. Otros artefactos relacionados al gradiente se deben a corrientes anormales pasando a través de las bobinas de gradiente.
- **Inhomogeneidad RF.** El problema de inhomogeneidad RF es una variación en intensidad a través de la imagen. La causa es ya sea un campo magnético RF ( $B_1$ ) no uniforme o una sensibilidad no uniforme en una sola bobina receptora.

Algunas bobinas RF, tales como las de superficie, naturalmente tienen variaciones y siempre mostrarán este artefacto. La presencia de este artefacto en otras bobinas representa la falla de un elemento en la bobina RF o la presencia de material no ferromagnético en el objeto del que se obtiene la imagen.

- **Volumen parcial.** Este es un artefacto que es causado por el tamaño del voxel de la imagen. Se deriva del hecho de que el valor de intensidad asignado a un voxel es el promedio de las contribuciones de todos los tejidos presentes en el volumen correspondiente, es decir, este efecto se presenta cuando un voxel representa más de una clase de tejido, empañando la distinción de intensidad entre las clases. Por ejemplo, si un voxel pequeño contiene solo señal de agua o grasa, y un voxel grande puede tener una combinación de ambos, el voxel grande posee una intensidad de señal igual al promedio ponderado de la cantidad de agua y grasa presente en el voxel. Otra manifestación de este artefacto es una pérdida de resolución causada por características múltiples presentes en el voxel.

La segmentación basada en intensidad de IRM presenta problemas debido a las inhomogeneidades del campo de la bobina, causando gradientes de intensidad a través de la imagen, lo que origina que se necesite un umbral distinto en partes diferentes del conjunto de datos, las rebanadas pueden ser notablemente más brillantes o más oscuras. Aunque las imágenes de IRM pueden parecer visualmente uniformes, tales inhomogeneidades con frecuencia alteran la segmentación basada en intensidad. En el caso ideal, la diferenciación entre materia blanca y gris debiera ser fácil ya que estos tejidos presentan intensidades distintas. En la práctica, si las inhomogeneidades de intensidad espacial son de suficiente magnitud, causan que las distribuciones de las intensidades de señal asociadas con estas clases de tejido se traslapen significativamente, lo que aunado al traslape de intensidades de diferentes estructuras agrava aun más el problema.

## 1.1. Métodos y aplicaciones

Además de los artefactos inherentes de las imágenes por RM, la aplicación de los métodos de segmentación de IRM enfrentan otro tipo de problemas significativos. Por ejemplo, la selección óptima de características en IRM multiespectral es importante para maximizar la diferenciación de contraste en tejido o la segmentación en el espacio característico. El nivel de supervisión del operador en el proceso de segmentación impactará en la estabilidad de los métodos de segmentación, particularmente en términos de variación inter e intraobservador.



### 1.1.1. Extracción de características

La segmentación de imágenes por RM está basada en conjuntos de características que se pueden extraer de las imágenes, tales como la intensidad del voxel, que a su vez se puede usar para calcular otras características como los bordes y la textura. La selección de buenas características es la clave para una segmentación exitosa.

**Características basadas en intensidad del pixel.** Muchas aproximaciones de segmentación simplemente usan los valores de escala de gris de los píxeles. La entrada de intensidad del pixel puede venir de una sola imagen, de un volumen o de un conjunto de datos multispectral de la misma ubicación anatómica. Si más imágenes están disponibles, se pueden agregar a los vectores de características para formar un espacio característico N-dimensional.

**Características basadas en intensidad de pixel calculada.** El uso de parámetros de RM calculados (derivados del tipo de relajación) se ha propuesto como características para segmentación de imágenes. Sin embargo, su aplicación ha sido limitada. La sensibilidad y especificidad de estos parámetros (absolutos) calculados, por ejemplo para la diferenciación de tejidos normales y patológicos, y la reproducibilidad de estas métricas calculadas, ha planteado problemas adicionales para obtener una segmentación estable.

**Características basadas en textura y bordes.** Debido a la naturaleza no uniforme de las imágenes por RM, un umbral global no puede segmentar datos de IRM pesados en T1, T2 o DP, y se requieren características adicionales. Para obtener estas características adicionales se han usado los métodos de detección de bordes, aunque con poco éxito porque no tienen una localización óptima. Se puede alcanzar una mejora en la detección de borde usando el detector de Canny [11], pero no es muy probable que los métodos basados solo en detección de bordes generen una segmentación confiable.

Otra característica usada en la segmentación de imágenes por RM es la textura. Las características de textura se aplican principalmente para clasificación, y no para delineación de regiones. La textura es una característica estadística necesariamente derivada de un gran número de píxeles; por lo que no es adecuada para clasificación de pixel.

La selección de características se vuelve importante si la dimensionalidad de los datos plantea cargas computacionales. Sin embargo, generalmente no hay suficientes características disponibles para plantear una carga computacional, o aun permitir una segmentación robusta. Los criterios para extracción y selección de características no han sido estudiados extensivamente pero se convertirá en un área importante a medida que el número de características aumente debido a los avances en los métodos de IRM.

### 1.1.2. Segmentación de imagen de una sola escala de gris

Esta técnica usa una sola imagen o escala de gris, representada por un solo tipo de datos, para realizar la segmentación y se puede subdividir en los siguientes métodos.

**Métodos de segmentación basados en umbralización.** La aproximación más intuitiva para la segmentación es la umbralización global, la dificultad de esta aproximación es determinar el valor de los umbrales. Se han propuesto los métodos de umbralización guiados por el conocimiento, donde los umbrales se determinan basándose en una función de bondad describiendo la separación de fondo, hueso y cerebro. El método está limitado, y su aplicación para uso clínico está impedido por la variabilidad de la anatomía y los datos de RM, así como por los artefactos de la imagen.

**Métodos de segmentación basados en borde.** Un borde o frontera en una imagen está definido por el gradiente de intensidad del pixel local. Los esquemas de detección de bordes sufren de la detección incorrecta de bordes debido al ruido, sobre y subsegmentación, y variabilidad en selección de umbral en la imagen de bordes. El rastreo de fronteras requiere que un operador seleccione un pixel en una región a delinear, después el método encuentra un punto en la frontera de la región y sigue la frontera desde este punto. La aplicación exitosa para segmentación global de IRM debe demostrarse, ya que se requiere una buena suposición inicial para las fronteras. Los métodos de rastreo de frontera, probablemente estén restringidos para segmentación de estructuras grandes y bien definidas como el parénquima cerebral.

**Métodos de segmentación de crecimiento de semilla.** Los métodos de crecimiento de región buscan grupos de pixeles con intensidades similares. Este tipo de segmentación requiere que un operador seleccione de manera empírica las semillas y umbrales. Los pixeles alrededor de las semillas se examinan, y se incluyen en la región si están dentro de los umbrales, algunas veces se agrega el requerimiento de que sean suficientemente similares a los pixeles que ya están en la región. Cada pixel agregado se convierte en una nueva semilla cuyos vecinos se inspeccionan para inclusión en la región. Los resultados con crecimiento de semilla, generalmente son dependientes de lo establecido por el operador.

### 1.1.3. Segmentación multiespectral

Si se adquieren diferentes imágenes del mismo objeto usando varias modalidades de imagen, o se coleccionan imágenes a través del tiempo, estas imágenes pueden proporcionar características diferentes del objeto para usarse en la segmentación. Los métodos de segmentación basados en integración de información de varias imágenes se llaman

multiespectrales o multimodales, donde están disponibles múltiples imágenes de RM con diferentes contrastes de escala de gris.

La aproximación más común para la segmentación de IRM multiespectral es el reconocimiento de patrones. Estas técnicas generalmente parecen ser exitosas particularmente para imágenes de cerebro, pero aun hay mucho por hacer en el área de validación.

La diferencia entre métodos supervisados y no supervisados se enfatiza en la literatura debido a la necesidad de mediciones reproducibles. Los métodos supervisados requieren que un usuario introduzca información previa para la segmentación. Esto se puede hacer seleccionando píxeles o regiones de entrenamiento en las imágenes. Los métodos no supervisados son automáticos, en el sentido de que la intervención del usuario puede no ser necesaria para completar el proceso, pero el resultado será independiente del usuario. Los datos de entrada a estos métodos son las características seleccionadas de la imagen. Una segmentación no supervisada define regiones en la imagen sin intervención de un usuario. Sin embargo, estas regiones no tienen un significado anatómico asociado. Por lo tanto, es necesario un paso de clasificación para obtener una imagen etiquetada de salida.

#### 1.1.3.1. Métodos de segmentación supervisados

**Métodos de reconocimiento de patrones.** Muchos métodos de reconocimiento de patrones asumen distribuciones particulares de las características, y son llamados métodos paramétricos. Por ejemplo el método de máxima verosimilitud (ML) asume distribuciones Gaussianas multivariadas [17]. Las medias y matrices de covarianza para cada uno de los tejidos se estiman de un conjunto de entrenamiento suministrado por el usuario. Los píxeles restantes se clasifican después calculando la verosimilitud de cada clase de tejido, y seleccionando el tipo de tejido con la probabilidad más alta. Los métodos paramétricos solamente son útiles cuando las distribuciones de características para las diferentes clases son bien conocidas, lo cual no es necesariamente el caso de IRM.

Los métodos no paramétricos, tal como el de los k-vecinos más cercanos (kNN) no confía en distribuciones predefinidas, sino en la distribución actual de las muestras de entrenamiento. kNN ha dado resultados superiores tanto en términos de exactitud y reproducibilidad comparado con los métodos paramétricos [17, 59].

**Métodos de segmentación de espacio característico manual.** Otros métodos de segmentación supervisada incluyen segmentaciones basadas en espacio característico que utilizan fronteras de decisión definidas por el operador. Aquí, se enfatiza el paso

de selección y extracción de características, y se despliegan estas características en algún tipo de gráfica multidimensional. Las áreas de más alta densidad (agrupaciones, clusters), que son visibles en la gráfica, se delinean manualmente y se asocian con tipos de tejido. También se hace selección manual de umbrales en cada característica para cada tipo de tejido, forzando las agrupaciones hacia cajas rectangulares.

**Variabilidad de los métodos de segmentación.** Todos los métodos supervisados son dependientes del operador, en virtud de lo cual existe una variabilidad inter e intraoperador. La medición de esta variabilidad ha mostrado ser relativamente grande. La reproducibilidad es un indicador importante de la confianza en las segmentaciones, no proporciona una medida de exactitud, ya que simplemente da una medida de fiabilidad ante la variación de operador o datos. Para anular o minimizar los efectos de la variabilidad se sugieren los métodos no supervisados.

### 1.1.3.2. Segmentación no supervisada

Las técnicas no supervisadas, también llamadas agrupamiento (clustering), automáticamente encuentran la estructura de los datos. Un grupo (cluster) es un área conectada en el espacio característico con una densidad alta de puntos. Los métodos no supervisados empleados para la segmentación de IRM incluyen el k-means, y su equivalente difuso, fuzzy c-means (FCM). Se ha encontrado que el agrupamiento no supervisado es prometedor para determinar el volumen de un tumor, pero el resultado depende de la etapa de inicialización requerida por la técnica.

Como ya se mencionó, los métodos supervisados son dependientes del operador, debido a esta razón, se prefieren los métodos no supervisados, ya que desde el punto de vista de la reproducibilidad, son altamente deseables. Sin embargo, los métodos no supervisados no necesariamente llegan a segmentaciones significativas, y con frecuencia requieren tiempos de cálculo grandes. Para contrarrestar estas limitaciones se han desarrollado metodologías que usan, por ejemplo, un algoritmo de reagrupamiento guiado por validación, con el que las particiones de baja calidad se refinan iterativamente hacia otras mejores. Otro desarrollo, que permite una supervisión parcial, se ha implementado como un FCM semisupervisado. Es posible, que en el futuro, se puedan alcanzar resultados clínicamente útiles, con suficiente reproducibilidad y requerimientos reducidos en tiempo de computación y poca habilidad del operador, usando técnicas parcialmente supervisadas [16], en las que la entrada del operador guíe al algoritmo en la detección de estructuras en los datos.

**Adiciones de dominio específico.** Los métodos de segmentación potencialmente se pueden mejorar, cuando el algoritmo de segmentación se complementa con un

conocimiento *a priori* acerca de la imagen. Los métodos se pueden optimizar para el dominio por inserción de heurísticas codificadas implícitamente o por el uso de una base de conocimiento explícito. Con frecuencia el conocimiento usado en segmentación de IRM se adquiere principalmente por el uso de datos de entrenamiento. La característica más importante usada en heurísticas y bases de conocimiento en la literatura sobre segmentación de RM es el nivel de gris en la imagen. También se ha usado el tamaño, posición y forma de las regiones, así como restricciones anatómicas sobre regiones vecinas. Asimismo, se ha reportado el uso de plantillas para describir, por ejemplo, un cerebro normal [3]. Esta plantilla se genera de un conjunto de estudios de IRM normales, transformados geoméricamente para igualar uno a otro. Esta aproximación toma en cuenta un paso adicional de una máscara de probabilidad para varios tejidos normales en el cerebro. Este atlas no está limitado a sistemas de IRM y protocolos específicos, y puede encontrar un uso amplio en procesamiento de IRM.

#### 1.1.4. Otros métodos

La segmentación general y de objetos específicos se ha estudiado con las alternativas del encuentro de fronteras y el crecimiento de región basado en intensidad. En el primer caso, se han considerado aproximaciones que incluyen gradientes de intensidad de imágenes dentro de estructuras coherentes en una variedad de formas. En el crecimiento de región, se han utilizado los datos multiparámetro inherentes en IRM para segmentar estructuras basándose en las propiedades de la señal física. Aquí, el árbol de decisión básico (minimum spanning tree) [65] y las estrategias de agrupamiento estadístico se han empleado para intentar separar la materia gris de la materia blanca, usando imágenes construidas matemáticamente de los tres parámetros de IRM: T1, T2 y Densidad de Protones.

Dentro de la alternativa del encuentro de fronteras o bordes en la segmentación, las estrategias basadas en modelos deformables, para aplicaciones médicas, han desarrollado “snakes” completamente en 3D que se pueden ejecutar en conjuntos de datos de imagen volumétrica. Algunas aproximaciones interesantes han incluido el uso de vectores de gradiente [64, 63], en lugar de sólo la magnitud del gradiente, para guiar la deformación de la superficie/contorno. Básicamente todas las estrategias se han implementado por optimización de funciones objetivo de una forma u otra; las formas más básicas, simplemente intentan encontrar un compromiso entre algún término de energía basado en la imagen y otro relacionado a una energía interna o modelo de forma (típicamente, solo suavidad de puntos adyacentes).

Otra alternativa interesante para contornos deformables basados en función objeti-

vo, plantea la deformación de contorno como una propagación de frente de onda, que puede ser vista como el conjunto de nivel cero de una función evolutiva. Esta evolución está gobernada por las velocidades dependientes de la curvatura de frentes o curvas en movimiento. La función evolutiva después se puede capturar en la forma de una ecuación diferencial, donde un término de velocidad obliga a la propagación a detenerse de acuerdo a información derivada de la imagen (por ejemplo, los bordes) [57]. Esta idea se ha tomado acoplando conjuntos de nivel (level sets) evolucionando para buscar la corteza exterior del cerebro, usando la restricción espacial de que la materia gris cortical es aproximadamente de igual espesor en la mayor parte de la superficie del cerebro.

Los algoritmos que encuentran superficie/frontera con frecuencia continúan siendo sensibles a la variación en los parámetros de adquisición de la imagen y a sus posiciones iniciales; por lo que los esfuerzos se están dirigiendo a integrar estrategias basadas en agrupación de intensidad y basadas en frontera para ganar robustez al ruido, en imagen borrosa e inicialización [12].

La segmentación basada en la intensidad de las imágenes ha progresado en algunas formas interesantes. Quizá de lo más notable es la idea de intentar agrupar valores de intensidad y corregir las distorsiones de los datos por RM simultáneamente. En las imágenes por RM la variación de intensidad debida a los efectos de campo sesgado (bias) puede afectar a los algoritmos de crecimiento de región. Una idea para estimar simultáneamente el campo sesgado y clasificar regiones de intensidad en el cerebro fue propuesta por Wells [62], la cual proporcionó una solución interesante y prometedora para este problema. La estrategia utilizó una aproximación Expectation-Maximization (EM) para ir hacia adelante y atrás, entre la clasificación de una imagen corregida de campo sesgado y la estimación misma del campo sesgado. Los esfuerzos aun continúan, intentando mejorar las estrategias de crecimiento de región basadas en intensidad, para encontrar las inhomogeneidades del campo usando fuzzy c-means y otras aproximaciones.

Mientras que se han hecho algunos esfuerzos para formar conjuntos de datos estandarizados para evaluar y validar algoritmos de segmentación en datos simulados, poco se ha hecho para organizar bases de datos con alguna forma de verdad de tierra (ground truth), expertos humanos u otra manera, que pudiera permitir la comparación cuantitativa y contrastar metodologías. Un esfuerzo inicial en esta dirección es el Internet Brain Segmentation Repository (IBSR) [39], que proporciona los resultados de la segmentación guiada manualmente por el experto, junto con los datos de imágenes de cerebro por RM. Su propósito es fomentar la evaluación y desarrollo de métodos de segmentación [27].

### 1.1.5. Clasificación

La segmentación de imagen se refiere a la descomposición de una imagen en regiones, pero estrictamente hablando no clasifica los tipos de tejido. La clasificación es el paso que trata con el etiquetado de las regiones. En la segmentación supervisada, esto es hecho por un operador cuando etiqueta los datos de entrenamiento, donde los píxeles se asocian con tejidos anatómicos. Los métodos no supervisados regresan una segmentación con diferentes clases, pero la asociación de clases con tipos de tejido debe aun ser realizada. Lo más común es que esta asociación sea hecha por un operador, o implícitamente por un médico. Mucho del conocimiento usado para etiquetar está basado en el significado de intensidades medias de región; por ejemplo, la región más oscura corresponde al fondo, y una región que es brillante en T2 pero oscura en T1 corresponde a LCR.

El etiquetado automático puede resultar ser muy complicado cuando no hay interacción con un operador o radiólogo. Sin embargo, en la práctica clínica, la intervención de un operador para etiquetar manualmente las regiones en el volumen segmentado, para permitir el cálculo de sus medidas geométricas, no parece presentar problemas.

### 1.1.6. Aplicaciones

Existe una cantidad enorme de propuestas enfocadas a resolver el problema de la segmentación de imágenes en general. A pesar de los logros alcanzados, no se tiene un algoritmo que pueda segmentar robustamente una variedad de estructuras relevantes en imágenes médicas sobre diversos conjuntos de datos. Por lo anterior, la investigación sobre métodos de segmentación se ha seguido realizando, y a continuación se presentan algunas aplicaciones específicas para la segmentación de imágenes cerebrales por RM.

Xialon et al. [66], proponen un planteamiento nuevo de propagación de superficies acopladas, vía los métodos de conjunto de nivel para la segmentación y medición de corteza cerebral. Logran una segmentación robusta y automática de la corteza, por la evolución simultánea de dos superficies empotradas, cada una determinada por su propia información basada en la imagen, y simultáneamente, obtienen una representación de las superficies corticales exterior e interior a partir de las cuales se puede calcular el área de la superficie.

En este planteamiento, en lugar de usar las características del gradiente, diseñan un operador local que hace uso de la información del nivel de gris, dando una medida de verosimilitud de que un voxel esté sobre la frontera entre dos tejidos. La verosimilitud se calcula suponiendo que los tejidos se obtienen independientemente de distribuciones Gaussianas. El operador local se aplica después de corregir los efectos de inhomogenei-

dad de la imagen.

Hacen uso de los métodos de conjunto de nivel para segmentar la corteza, inicializando con esferas concéntricas. Consideran dos ecuaciones de evolución, una para cada una de las superficies de frontera cortical interior (MB/MG) y exterior (LCR/MG); expresadas como funciones de la dirección normal a la superficie, información derivada de la imagen, y la distancia entre las dos superficies. Para mejorar la eficiencia computacional, incorporan en su algoritmo un método de banda angosta, modificando el método de conjunto de nivel tal que sólo se actualizan los puntos cercanos a los frentes de propagación actuales.

Zhang et al. [67] mencionan que los modelos de mezcla finita son muy usados en segmentación, sin embargo, tienen una limitación, no toman en cuenta la información espacial porque consideran a todos los datos, muestras independientes de una población. Para tratar este problema han desarrollado un modelo de campo aleatorio de Markov oculto (HMRF). La importancia del modelo deriva de la teoría MRF, en la que la información espacial en una imagen se codifica a través de restricciones contextuales de pixeles vecinos. Por la imposición de tales restricciones, se espera que los pixeles vecinos tengan las mismas etiquetas de clase (en imágenes constantes por tramos) o intensidades similares (en imágenes continuas por tramos).

El algoritmo empieza con un paso de estimación inicial para obtener la clasificación y parámetros del tejido iniciales; seguido de un proceso EM de tres pasos que actualiza las etiquetas de clase, los parámetros del tejido y sesgo del campo iterativamente. En las iteraciones se usa una aproximación MRF-MAP para estimar las etiquetas de clase, MAP se aplica para estimar el sesgo del campo, y los parámetros del tejido se estiman por ML. La salida del algoritmo son los tejidos segmentados, el sesgo del campo estimado y la imagen sin sesgo de campo.

Para el modelo HMRF es esencial un paso de estimación de parámetros, el cual se deriva del algoritmo EM. Se muestra que con la incorporación HMRF-EM se puede alcanzar una segmentación robusta y exacta. HMRF-EM requiere alguna manera de corrección del artefacto de sesgo del campo, debido a que el algoritmo de Wells [62] es problemático para imágenes con clases que no siguen la distribución Gaussiana, en su lugar usan el algoritmo EM modificado (MEM) [36], que unifica todas las clases que no siguen la distribución dentro de una clase separada.

Realizaron experimentos en dos y tres dimensiones. En datos simulados, MEM da resultados inferiores, comparado con HMRF-EM. También se concluye que el algoritmo HMRF-EM tiene ventaja sobre FM-EM cuando el ruido de la imagen no es despreciable. Aunque aplicaron el algoritmo a los datos del IBSR, no hicieron comparaciones



cuantitativas por considerar que la calidad de la segmentación manual era pobre.

Sabiendo que en los modelos deformables, la evolución del contorno es muy sensible a la posición inicial, y que no puede arreglárselas con protuberancias significativas y cambios topológicos, en [8] se propone como solución alterna, minimizar la energía considerando la evolución del contorno como un problema de propagación de frente, usando el formalismo de conjunto de nivel.

Las técnicas iterativas de conjunto de nivel son apropiadas para segmentar formas complejas, y su resultado es menos dependiente de la inicialización que en otros métodos iterativos. Sin embargo, la parametrización es una limitante para usos prácticos. Se han propuesto varios modelos de evolución, pero incluyen muchos parámetros: tamaño del paso, parámetros de ponderación, término de propagación adicional, etc. En este trabajo, aplican la técnica de conjunto de nivel para segmentar estructuras anatómicas del cerebro, realizando un registro denso antes de la segmentación. Usan un método de registro automático para inicializar la superficie, como alternativa a la inicialización manual. La técnica de registro [38], sigue un esquema jerárquico robusto que se puede enfocar sobre las áreas a segmentar. El algoritmo fue probado con datos simulados del BrainWeb para segmentar el cerebro completo, así como en datos reales de 18 volúmenes de RM.

En [47] Marroquin et al., indican que los métodos de segmentación necesitan incluir simultáneamente la estimación de modelos de intensidad variable para cada clase, y el conocimiento previo acerca de la ubicación y coherencia espacial de las regiones correspondientes. El diseño de estos métodos es probabilístico, y las suposiciones de coherencia espacial se pueden incorporar usando una aproximación bayesiana, por medio de los modelos de campo aleatorio de Markov (MRF); y las restricciones anatómicas de localización de las clases se pueden especificar como probabilidades previas, obtenidas de mapas probabilísticos de tejido cerebral.

Debido a las limitaciones que tienen las aproximaciones bayesianas, en este trabajo se presenta un esquema bayesiano totalmente automático que supera estas dificultades, incorporando las siguientes características: En lugar de un sólo campo de sesgo multiplicativo que afecta todas las intensidades del tejido, se proponen modelos paramétricos suavizados para la intensidad de cada clase. El uso de un atlas de cerebro y un método de registro robusto para encontrar la transformación no rígida que mapea el atlas hacia los datos a segmentar. Adicionalmente, proponen una variante del algoritmo EM que permite encontrar segmentaciones óptimas.

Procesaron dos conjuntos de datos: simulados, obtenidos del BrainWeb, y reales de la base de datos del IBSR. Obtienen resultados para los datos reales, superiores a los

que están reportados en la base de datos.

En [42] se presenta un algoritmo que garantiza la minimización de posibles errores, proporcionando así, una clasificación más exacta basada en la calidad de la señal de entrada. La propuesta usa la distribución relativa en 3D de las intensidades del pixel para crear un modelo probabilístico, el cual se construye usando las propiedades de las matrices de co-ocurrencia en todas las direcciones del vecindario para cada pixel.

Antes de segmentar, remueven los artefactos producidos por el ruido, usando un filtro de difusión anisotrópica. La clasificación de las imágenes se hace a través del uso de un modelo estadístico, basado en el conocimiento de las orientaciones del vecindario, clasificando el pixel actual en función de su vecindario.

La mejor clasificación se alcanza determinando la estimación de la probabilidad total mínima de clasificación errónea para cada pixel, la cual a su vez, está basada en el mínimo costo esperado de metodología de clasificación, que intenta separar valores dentro de poblaciones diferentes basándose en un análisis probabilístico y estadístico de la población. Al modelo estadístico propuesto lo llaman la Probabilidad Total Mínima Multivariada de Modelo de Clasificación Errónea, MMTPM.

Las componentes consideradas por MMTPM son las medias de intensidad, sus varianzas y probabilidades previas. El vecindario usado en la decisión de clasificación se restringe con un criterio de distancia Euclidiana, y para cada voxel se evalúa una función de decisión para cada clase posible, asignándolo a la clase resultante mayor. Después de segmentar, el etiquetado fue llevado a cabo encontrando el vecino más cercano a la intensidad media del segmento. El algoritmo fue probado con la base de datos del IBSR.

Se han presentado, a manera de ejemplo, algunas aplicaciones para segmentación de IRM cerebral, no siendo las únicas revisadas en la bibliografía, no se encontró reporte alguno sobre la aplicación del algoritmo de corrimiento de media (mean shift), CM, a imágenes médicas. En este trabajo de investigación, el algoritmo de CM tiene el papel más relevante, y la motivación para usarlo es que a pesar de sus excelentes cualidades, parece que no se conoce en la literatura estadística. Aunque el libro [56, p. 132] discute el algoritmo del CM, las ventajas de emplear un procedimiento de corrimiento de media en la estimación de densidad fueron redescubiertas en [14]. El procedimiento del CM es una herramienta muy versátil para el análisis del espacio característico y puede proporcionar soluciones confiables en la tarea de segmentación de imágenes. A continuación, se presentan los antecedentes del algoritmo de CM.

## 1.2. Corrimiento de media

La segmentación por espacio característico manual puede producir resultados incorrectos, debido a la delineación manual y asociación con tipos de tejido hechas por el usuario. Para evitar este tipo de errores la segmentación debe ser manejada por la aplicación, es decir, soportada por información con un alto nivel de independencia. Para esto se requiere de una representación bastante confiable de los datos de entrada, y que el proceso de extracción de características esté controlado por muy pocos parámetros que correspondan a medidas intuitivas del dominio de la entrada. Lo anterior se puede lograr por el análisis de imágenes basado en el espacio característico, pero sin la intervención del usuario.

Un espacio característico es un mapeo de la entrada obtenido por el procesamiento de los datos en subconjuntos pequeños a la vez. Para cada subconjunto se obtiene una representación paramétrica de la característica de interés, y el resultado está mapeado hacia un punto en el espacio multidimensional del parámetro. Después de procesar la entrada, las características significativas corresponden a las regiones más densas en el espacio característico, es decir, a grupos (clusters). El objetivo del análisis de la imagen es la delineación de estos grupos.

Las ventajas y desventajas del espacio característico surgen de la naturaleza global de la representación derivada de la entrada. Por un lado, se reúne toda la evidencia para la presencia de una característica significativa, proporcionando una tolerancia excelente al nivel de ruido, evitando que el ruido pueda dar decisiones locales que no son confiables. Por otra parte, las características con menor soporte en el espacio característico pueden no detectarse a pesar de ser sobresalientes para la tarea a ejecutarse. Esta desventaja, sin embargo, se puede evitar ya sea aumentando el espacio característico con parámetros (espaciales) adicionales desde el dominio de entrada, o por posprocesamiento robusto del dominio de entrada guiado por los resultados del análisis del espacio característico.

Mientras que existe una gran cantidad de técnicas de agrupamiento publicadas, muchas de ellas no son adecuadas para analizar los espacios característicos derivados de datos reales. Los métodos que confían en un conocimiento *a priori* del número de grupos presentes (incluyendo aquellos que usan optimización de un criterio global para encontrar el número), así como los métodos que implícitamente asumen la misma forma para todos los grupos en el espacio, no son capaces de manejar la complejidad de un espacio característico real.

Los espacios característicos estructurados arbitrariamente sólo se pueden analizar por métodos no paramétricos, debido a que estos métodos no tienen suposiciones incrus-

tadas. Los métodos de agrupamiento no paramétricos se pueden clasificar en dos clases: agrupamiento jerárquico y estimación de densidad. Las técnicas jerárquicas agregan o dividen los datos basándose en alguna medida de proximidad; tienden a ser computacionalmente costosas y no es directa la definición de un criterio de paro significativo para la fusión o división de los datos.

En el enfoque de agrupamiento no paramétrico basado en la estimación de densidad, el espacio característico de una imagen se puede considerar como la función de densidad de probabilidad ( $fdp$ ) empírica del parámetro representado. Las regiones densas en el espacio característico corresponden al máximo local de la  $fdp$ , esto es, a las modas de la función desconocida. Una vez que se determina la ubicación de una moda, el grupo asociado con ella se delinea basándose en la estructura local del espacio característico.

En este trabajo el planteamiento para la detección de la moda y agrupamiento está basado en el procedimiento del corrimiento de media propuesto por Fukunaga y Hostetler en 1975 [33].

En muchos problemas de reconocimiento de patrones, se dispone de muy poca información acerca de la  $fdp$  verdadera y de su forma. Debido a esta falta de conocimiento se debe confiar en las técnicas no paramétricas para obtener las estimaciones del gradiente de densidad. Fukunaga y Hostetler derivan un kernel general de estimación del gradiente multivariado usando las estimaciones de kernel de Parzen.

Una vez estimado el gradiente de densidad de probabilidad como el gradiente de la estimación de densidad, basándose en una función de kernel diferenciable, aplican la función del kernel gaussiano para entender el proceso involucrado en la estimación del gradiente, llegando a la interpretación intuitiva de que la estimación del gradiente de densidad es esencialmente una medida ponderada del corrimiento de la media de las observaciones alrededor del punto de análisis.

La interpretación intuitiva la demuestran sustituyendo el kernel de Epanechnikov en la expresión de la estimación del gradiente de densidad, derivando la expresión de lo que llamaron el corrimiento de media muestral [33]:

$$M_h(X) \equiv \frac{1}{k} \sum_{X_i \in S_h(X)} (X_i - X)$$

de las observaciones en la región  $S_h(X)$  alrededor de  $X$ . El tamaño de la región  $S_h(X)$  está en función del radio de ventana  $h$ , y  $k$  es el número de observaciones  $X_i$  que caen dentro de  $S_h(X)$ . Si el gradiente en  $X$  o pendiente es cero, correspondiendo a una densidad uniforme sobre la región  $S_h(X)$ , el promedio del corrimiento de media será cero debido a la simetría de las observaciones cercanas a  $X$ . Sin embargo, con un gradiente

de densidad diferente de cero apuntando en la dirección de más rápido incremento de la  $f_{dp}$ , en el promedio más observaciones deben caer a lo largo de su dirección que en cualquier otra dentro de  $S_h(X)$ . Así, el corrimiento de media promedio debe apuntar en esa dirección y tener una longitud proporcional a la magnitud del gradiente.

La estimación del gradiente la aplican en [33] como un método de agrupamiento, donde asignan cada observación a la moda más cercana a lo largo de la dirección del gradiente en los puntos de observación. Los datos experimentales que usaron fueron tres distribuciones normales bivariadas con medias diferentes y matriz de covarianza igual. Aplicando recursivamente el algoritmo de corrimiento de media, los datos convergieron a tres grupos. También muestran que el algoritmo se puede usar como filtro para contraer los datos sobre su superficie intrínseca, reteniendo las propiedades dominantes de la estructura de los datos.

Olvidado durante 20 años, el algoritmo del corrimiento de media nuevamente es abordado por Cheng [13] en 1995. Cheng generaliza el algoritmo del corrimiento de media, y muestra que el algoritmo de agrupamiento k-means es un caso límite del corrimiento de media. También establece que el corrimiento de media es equivalente al ascenso por gradiente en la estimación de densidad. El corrimiento de media es el mejor paso de ascenso, con un tamaño de paso variable, que está dado por la magnitud del gradiente.

Con respecto a la convergencia del algoritmo, Cheng muestra que los datos convergen a una sola posición, y que en el caso extremo, cuando el kernel se trunca al grado de que sólo cubre un punto de los datos, no ocurre la fusión entre los datos. Sin embargo, cuando el tamaño del kernel es adecuado, los datos pueden tener trayectorias que fusionan hacia números variables de centros de grupo, y así, cuando dos puntos de los datos convergen a la misma posición, se considera que pertenecen al mismo grupo.

La filosofía del artículo de Cheng es que el agrupamiento se puede ver como el resultado de algún proceso natural como el corrimiento de media. Entonces, el resultado o propósito del agrupamiento por corrimiento de media es usar un máximo local como una caracterización de los datos; por lo que el corrimiento de media se considera como una aproximación a la densidad de los datos.

En los dos trabajos anteriores se plantea el procedimiento del corrimiento de media y su generalización, pero es en la Universidad de Rutgers en el Robust Image Understanding Laboratory, donde Peter Meer, con interés en aplicar la estadística a la visión, inicia la investigación del corrimiento de media aplicado a la segmentación de imágenes. A grandes rasgos, el desarrollo de la línea de investigación sobre este tema se presenta a continuación.

Los primeros en aplicar el corrimiento de media para la segmentación de una imagen de color son Comaniciu y Meer[19], quienes presentan un algoritmo no supervisado de agrupamiento no paramétrico basado en el procedimiento del corrimiento de media (CM). El procedimiento del CM es una técnica ascendente de pendiente muy marcada y adaptable que calcula el vector de CM para cada punto de los datos, que desplaza el kernel por esa cantidad, y repite los cálculos hasta alcanzar una moda. El algoritmo de agrupamiento empieza seleccionando aleatoriamente un conjunto muestral de los datos, en seguida, se aplica el procedimiento del CM al conjunto muestral, definiendo un conjunto de  $m$  centros de grupo candidatos definidos por los puntos de convergencia. Después se perturban los centros de clase candidatos y se aplica nuevamente el CM, y a continuación se derivan los centros de grupo desde los centros candidatos; siguiendo una validación de los centros de grupo por medio de un umbral. Finalmente, se delinean los grupos asociando cada punto de la muestra con un centro de grupo. Para asignar los puntos de los datos se emplea una técnica del  $k$ -vecino más cercano, es decir, cada dato pertenece al grupo definido por la mayoría de sus  $k$ -puntos de la muestra más cercanos.

El procedimiento se aplicó a la segmentación de imágenes de color de una persona, un avión y una casa, delineando las regiones homogéneas de las mismas, esto es los grupos de diferente color.

En [20], Comaniciu y Meer muestran que la aplicación del CM, en el dominio combinado espacial-rango, permite obtener una segmentación de imagen y filtrado con preservación de bordes de alta calidad. Los métodos desarrollados son muy simples basándose en la misma idea de desplazar iterativamente una ventana de tamaño fijo, hacia el promedio de los datos dentro de ella.

La salida del filtro de CM para un pixel de la imagen, está definida como la información de rango (nivel de gris, color, o información espectral) implicada por el punto de convergencia. Para la segmentación, los puntos de convergencia suficientemente cerca en el dominio conjunto se fusionan para obtener las regiones homogéneas de la imagen.

El filtrado con CM se aplicó a la muy usada imagen *cameraman* en tonos de gris. La región del campo de pasto se suavizó casi completamente, mientras que se conservaron los detalles. Para el proceso de segmentación se usaron una imagen de un edificio en tonos de gris, una imagen de un lago, y de una mano ambas en color. En todos los casos se obtuvieron contornos de alta calidad permitiendo la delineación de regiones homogéneas. Los temas de filtrado y segmentación de imágenes son abordados de manera más amplia y formal en [21].

El algoritmo de CM también se ha aplicado al rastreo de objetos. En [22] se presenta una propuesta para rastreo (tracking) en tiempo real de objetos no rígidos, basada en

características visuales como color y/o textura, cuya distribución estadística caracteriza el objeto de interés. El análisis de movimiento es una aplicación natural del CM. Las iteraciones del CM se emplean para encontrar el blanco candidato que es el más similar a un blanco modelo preestablecido, con la similitud expresada por una métrica basada en el coeficiente de Bhattacharyya. En la tarea de encontrar la ubicación del blanco, lo que se hace es encontrar la ubicación discreta cuya densidad asociada es la más similar a la densidad del blanco.

El método se aplicó a la tarea de rastrear un jugador de fútbol marcado con una región elipsoidal. El rastreador basado en CM demostró ser robusto a oclusión parcial, desorden, distractores y movimiento de la cámara. Ya que no se supuso un modelo de movimiento, el rastreador se adaptó bien al carácter no estacionario de los movimientos del jugador, el cual alterna abruptamente entre acciones lentas y rápidas. El método también se aplicó a rastrear personas en un andén del metro, demostrando la capacidad de adaptarse a cambios de escala.

Una de las limitaciones del algoritmo de CM es que involucra la especificación de un parámetro de escala. Mientras que los resultados obtenidos parecen satisfactorios, cuando las características locales del espacio característico difieren significativamente a través de los datos, es difícil encontrar un ancho de banda global óptimo para el procedimiento del CM. En [23] se adapta localmente el ancho de banda, y se estudia una alternativa para la selección de escala manejada por los datos. A continuación se presenta un resumen de la alternativa de selección de ancho de banda variable hecha en [23].

El término *ancho de banda fijo* se refiere al hecho de que el radio (ancho de banda) del kernel de la estimación de la función de densidad, se mantiene constante a través de todos los elementos del espacio  $d$ -dimensional. Como resultado de esto, el procedimiento de ancho de banda fijo, estima la densidad en cada punto tomando el promedio de kernels escalados idénticamente y centrados en cada uno de los puntos de los datos. El ancho de banda  $h$  se puede variar de dos maneras. Una es seleccionando un ancho de banda diferente  $h = h(\mathbf{x})$  para cada punto  $\mathbf{x}$  de la estimación, definiendo el estimador como :

$$\hat{f}_1(\mathbf{x}) = \frac{1}{nh(\mathbf{x})^d} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h(\mathbf{x})}\right)$$

llamado el estimador de densidad *globo* (balloon); en este caso, la estimación de la *fdp* en  $\mathbf{x}$  es el promedio de kernels escalados idénticamente y centrados en cada punto de los datos.

La otra manera de variar el radio, es seleccionando anchos de banda diferentes  $h = h(\mathbf{x}_i)$

para cada punto de los datos  $\mathbf{x}_i$  obteniendo el estimador de densidad de *punto muestral*:

$$\hat{f}_2(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h(\mathbf{x}_i)^d} K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h(\mathbf{x}_i)}\right)$$

por lo que la estimación de la *fdp* en  $\mathbf{x}$  es el promedio de kernels escalados diferentemente y centrados en cada punto de los datos.

A pesar de que el estimador globo es más intuitivo, su mejora de rendimiento comparado con el de ancho de banda fijo es insignificante, cuando el ancho de banda  $h(\mathbf{x})$  se elige como una función del  $k$ -ésimo vecino más cercano, además de que usualmente no integra a uno. El estimador de punto muestral es por sí mismo una densidad, no negativo e integra a uno, y su mejor propiedad es que reduce el sesgo cuando  $h(\mathbf{X}_i)$  se elige como:

$$h(\mathbf{X}_i) = h_0 \left( \frac{\lambda}{f(\mathbf{X}_i)} \right)^{1/2}$$

donde  $h_0$  representa un ancho de banda fijo y  $\lambda$  es una constante de proporcionalidad. Debido a que  $f(\mathbf{X}_i)$  se desconoce,  $h_0$  y una función inicial (llamada piloto) se encuentran con la regla plug-in. La estimación final depende de la constante de proporcionalidad  $\lambda$ , la cual puede calcularse como la media geométrica de la función piloto. El estimador muestral casi todas las veces es mejor que el estimador de ancho de banda fijo. Cuando se usa el estimador de punto muestral, asocia a cada uno de los datos un kernel escalado diferentemente, y es el paso básico del procedimiento iterativo del *CM de ancho de banda variable*.

El principal problema del CM de ancho de banda variable es que requiere un ancho de banda inicial  $h_0$ , el cual es difícil de determinar y muchos de los aspectos prácticos no se han resuelto. Debido a lo anterior, se propone seleccionar el ancho de banda inicial de otra manera. La idea es imponer una estructura local sobre los datos, asumiendo que localmente la densidad fundamental es normal esféricamente, con media  $\mu$  desconocida y matriz de covarianza  $\Sigma = \sigma^2 \mathbf{I}$ . La tarea de encontrar lo dos parámetros anteriores parece difícil, sin embargo, en el artículo se prueba un teorema que establece que  $h_0 \equiv \sigma$  conduciendo a una regla de selección de escala certera y simple.

Con el nuevo procedimiento segmentan un cuadro de video en manchas (blobs) representativas detectadas en los dominios espacial y de color. El algoritmo mostró estabilidad en segmentar una secuencia de color obtenida por una panorámica de la cámara.

La selección del ancho de banda nuevamente es abordada en [18], sigue la misma línea



presentada anteriormente, pero ahora la matriz de covarianza es la más estable a través de las escalas. Cuando la distribución de datos fundamental es normal, el análisis del ancho de banda no tiene influencia en el cálculo de  $(\mu, \Sigma)$ . Cuando la estructura fundamental se desvía de la normalidad, el ancho de banda afecta la estimación. Por lo tanto, se prueba la estabilidad de  $(\mu, \Sigma)$  contra la variación del análisis del ancho de banda. La prueba de estabilidad para las distribuciones normales  $p_j$  asociadas a los anchos de banda involucra el cálculo de la disimilitud total entre  $p_j$  y sus vecinos. La disimilitud se mide usando una versión especializada de la divergencia de Jensen-Shannon. El par  $(\mu, \Sigma)$  más estable es el que minimiza la divergencia entre distribuciones vecinas. El autor muestra resultados al aplicar el método a la segmentación de imágenes de color.

Todas las aplicaciones del CM hasta ahora, tales como segmentación de imagen, reconocimiento de objetos y rastreo fueron en espacios de baja dimensión. En [34] se presenta una implementación del CM basada en características de alta dimensión. El cuello de botella en dimensiones altas es la necesidad de un algoritmo rápido para realizar consultas de vecindario cuando se calculan las iteraciones del CM. Para enfrentar esta situación se usa el algoritmo del vecino más cercano basado en LSH (locality-sensitive hashing) y adaptado para manejar datos complejos. Así, dada la ubicación actual del kernel en el proceso iterativo, una consulta basada en LSH recupera el conjunto aproximado de vecinos necesarios para calcular la siguiente ubicación. La resolución del análisis de datos está controlada por el usuario, y dado que aplican el método de CM adaptable (esto es, de ancho de banda variable), el usuario determina el número de vecinos usados en el procedimiento de densidad piloto. Los anchos de banda asociados con los puntos de los datos se obtienen ejecutando consultas de vecindario. Una vez determinados los anchos de banda, el procedimiento adaptable corre aproximadamente con el costo del CM de ancho de banda fijo.

En [34] también se muestra la superioridad del CM adaptable, comparado con el de ancho de banda fijo, en dimensiones altas. Además, se muestra que el uso de la estructura de datos LSH en el procedimiento asegura un aumento significativo de velocidad.

En la clasificación de textura, una textura está caracterizada por *textons*, que son centros de grupo en un espacio característico derivados de la entrada. El espacio característico se construye de la salida de un banco de filtros aplicado a cada pixel, y las propuestas difieren en el banco de filtros empleado.

Para encontrar los textons, usualmente se usa el algoritmo k-means, que como se sabe tiene las limitaciones de que la forma de los grupos se restringe a ser esférica y su número tiene que ser dado antes del procesamiento. En este planteamiento se sustituye el k-means por el CM adaptable, así los textons en lugar de estar basados en medias

ahora están basados en modas, y el número de ellas se determina automáticamente. La experimentación se llevó a cabo sobre la base de datos Brodatz [10] que contiene 112 texturas con diferentes grados de dificultad. Los mejores resultados se obtuvieron con un banco de filtros llamado LM (combinación de 48 filtros isotrópicos y anisotrópicos) y aplicando el CM adaptable.

### 1.3. Planteamiento del problema y principales contribuciones

El problema de la segmentación de imágenes ha sido abordado desde diferentes perspectivas y se ha propuesto una gran cantidad de metodologías para delinear las estructuras de una imagen por RM del cerebro. Sin embargo, el problema no se ha resuelto totalmente, puesto que no se tiene una técnica de segmentación robusta, capaz de segmentar una diversidad de estructuras relevantes en imágenes médicas sobre un conjunto de datos amplio.

Entender los espacios de características de una imagen, derivados de los datos reales, se dificulta porque muchas veces tienen una estructura compleja y no se dispone de información a priori para guiarnos en el análisis. Las características significativas que se necesitan recuperar, corresponden a agrupaciones en el espacio. El número de agrupaciones, su forma y reglas de asignación se deben distinguir solamente de los datos dados.

El espacio de características se puede considerar como una muestra de una función de densidad de probabilidad desconocida. La estimación por kernel es una buena elección práctica para estimar esa función desconocida; el algoritmo de corrimiento de media ha mostrado buenos resultados en la segmentación de imágenes no médicas, por lo que en este trabajo de investigación se utilizó para la estimación de la función de densidad de IRM cerebral, usando las modas de la función de densidad para delinear las estructuras neuroanatómicas que conforman el cerebro.

El objetivo de este trabajo de investigación fue: Proponer una metodología de segmentación de imágenes por resonancia magnética cerebrales aplicando una técnica de estimación robusta y no paramétrica.

Las hipótesis de trabajo que se consideraron fueron las siguientes:

La segmentación de imágenes por RM, realizada con un método no paramétrico, supera a las producidas por los métodos paramétricos.

El corrimiento de media es menos sensible al problema de las inhomogeneidades del

campo de la antena que se presenta en la segmentación de imágenes por RM.

## Principales contribuciones de la tesis

La metodología de segmentación propuesta está basada en el algoritmo de corrimiento de media, dando como resultado una imagen segmentada con una preservación adecuada de bordes entre regiones homogéneas. Las aportaciones de la metodología de segmentación, aplicada a imágenes cerebrales de RM, son las siguientes:

- La metodología es guiada únicamente por los datos, lo que permite extraer las características de la imagen sin suponer una distribución particular de los datos y sin requerir información previa o de inicialización.
- La metodología es robusta aun en la presencia de inhomogeneidades, sin requerir de un proceso previo de corrección de inhomogeneidad.
- La técnica usa un mapa de confianza de bordes para conservar las fronteras entre tejido cerebral, incluso conservando los bordes débiles originados por los efectos del volumen parcial.
- Aunque el procedimiento de CM se ha usado con anterioridad en la segmentación de imágenes, esta es la primer propuesta en aplicar y valorar el uso del CM en la segmentación de imágenes cerebrales de RM.
- La fusión de regiones homogéneas, producidas por el procedimiento de CM, se realiza a través de la reducción de un grafo de adyacencias, aplicando criterios de magnitud del borde y distancia en modas entre regiones.
- Se propone una metodología de segmentación automática y completa, que entrega como resultado, una imagen clasificada en regiones anatómicamente significativas, a partir únicamente de los datos de entrada.
- La técnica de segmentación no sólo es aplicable a la característica de intensidad, sino que se puede extender de manera directa al uso de más componentes del espacio característico de la imagen.

# Capítulo 2

## Estimación por corrimiento de media

La técnica de segmentación propuesta es un método no paramétrico basado en la estimación de una función de densidad de probabilidad. En este capítulo se presenta el concepto de estimación de densidad, la estimación por kernel, la medida de discrepancia del estimador, y el ancho de ventana y kernel ideal. Además, se describe el corrimiento de media como la estimación del gradiente de densidad normalizado, así como la convergencia del procedimiento, y cómo se usa en un problema de agrupamiento considerando el dominio espacial-rango. El capítulo se termina con la formulación del corrimiento de media ponderado por la confianza del borde. El contenido de este capítulo está basado en el libro [56] de Silverman y en la publicación [20] de Comaniciu.

### 2.1. Estimación de densidad

Considere cualquier cantidad aleatoria  $X$  que tiene función de densidad de probabilidad  $f$ . La especificación de la función  $f$  da una descripción natural de la distribución de  $X$ , y permite encontrar las probabilidades asociadas con  $X$  de la relación:

$$P(a < X < b) = \int_a^b f(x)dx \quad \forall a < b \quad (2.1)$$

Suponga que tenemos un conjunto de datos observados y que son una muestra de una función de densidad de probabilidad. La estimación de densidad es la construcción de un estimado de la función de densidad de los datos observados.

La estimación de densidad *paramétrica* supone que los datos son extraídos de una familia de distribuciones paramétricas conocida, por ejemplo la distribución normal con

media  $\mu$  y varianza  $\sigma^2$ . Los datos de la densidad  $f$  se pueden estimar encontrando estimaciones de  $\mu$  y  $\sigma^2$  de los datos y substituyendo estas estimaciones en la fórmula de la densidad normal. La desventaja de la aproximación paramétrica es la restricción de pertenecer a una familia paramétrica, que algunas veces puede ser muy rígida. Esta rigidez se evita removiendo la restricción, obteniéndose lo que comúnmente se refiere como una aproximación *no paramétrica*, de esta manera, la aproximación tiende a ser más *no paramétrica* si se hacen suposiciones menos rígidas acerca de la distribución de los datos observados. Aunque se supondrá que la distribución tiene una densidad de probabilidad  $f$ , los datos hablarán por sí mismos en la determinación de la estimación de  $f$  que mejor los ajusta, a diferencia de lo que pudiera ser si  $f$  estuviera restringida a caer en una familia paramétrica dada.

La estimación de densidad puede dar una indicación valiosa de características tales como el sesgo y multimodalidad en los datos. En algunos casos producirá conclusiones que pueden ser consideradas como evidentemente verdaderas, mientras que en otros, todo lo que hará es apuntar el camino para un análisis y/o colección de datos adicional.

### 2.1.1. Métodos univariados

Muchas de las aplicaciones de estimación de densidad son en datos multivariados, pero todos los métodos multivariados son generalizaciones de métodos univariados; razón por la que vale la pena considerar primero el caso univariado.

Excepto donde otra cosa se establezca, se supondrá una muestra de  $n$  observaciones reales  $X_1, \dots, X_n$  cuya densidad se debe estimar. El símbolo  $\hat{f}$  denotará el estimador de densidad.

#### Histogramas

El estimador de densidad más antiguo y más usado es el histograma. Dado un origen  $x_0$  y una rebanada de ancho  $h$ , definimos las rebanadas del histograma como los intervalos  $[x_0 + mh, x_0 + (m + 1)h]$  para  $m \in \mathbb{Z}$ . El histograma se define por:

$$\hat{f}(x) = \frac{1}{nh}(\text{núm. de } X_i \text{ en la misma rebanada que } x) \quad (2.2)$$

El histograma se puede generalizar permitiendo que el ancho de la rebanada varíe. Suponga que tenemos alguna disección de la línea real en rebanadas, entonces la estimación se define por

$$\hat{f}(x) = \frac{1}{n} \times \left( \frac{\text{núm. de } X_i \text{ en la misma rebanada que } x}{\text{ancho de la rebanada conteniendo } x} \right) \quad (2.3)$$

Aquellos escépticos acerca de la estimación de densidad preguntan ¿por qué es necesario usar métodos más sofisticados que el histograma? El caso de tales métodos y las desventajas del histograma dependen del contexto. En términos de varias descripciones matemáticas de exactitud, el histograma se puede mejorar substancialmente, y su desventaja matemática se traduce en un uso ineficiente de los datos si el histograma se usa como estimación de densidad en procedimientos como análisis de grupo (cluster) y análisis discriminante no paramétrico. La discontinuidad de los histogramas causa una dificultad extrema si se requieren las derivadas de la estimación.

Para la presentación y exploración de datos, los histogramas son una clase de estimación de densidad extremadamente útil, particularmente en el caso univariado. Para la presentación gráfica de datos bivariados o trivariados presenta severas dificultades. Finalmente, se debe enfatizar que, en todos los casos, el histograma aun requiere elegir la cantidad de suavizado.

### Estimador ingenuo (naive)

De la definición de densidad de probabilidad, si la variable aleatoria  $X$  tiene densidad  $f$ , entonces

$$f(x) = \lim_{h \rightarrow 0} \frac{1}{2h} P(x-h < X < x+h) \quad (2.4)$$

Para cualquier  $h$  dada, podemos estimar  $P(x-h < X < x+h)$  por la proporción de muestras que caen en el intervalo  $(x-h, x+h)$ . Así, un estimador natural  $\hat{f}$  de la densidad está dado por la elección de un pequeño número  $h$ , estableciendo

$$\hat{f}(x) = \frac{1}{2hn} [\text{núm. de } X_1, \dots, X_n \text{ cayendo en } (x-h, x+h)] \quad (2.5)$$

el cual es llamado el estimador ingenuo.

Para expresar el estimador más claramente, se define la función de peso  $w$  por

$$w(x) = \begin{cases} \frac{1}{2} & \text{si } |x| < 1 \\ 0 & \text{de otra manera} \end{cases} \quad (2.6)$$

Es fácil ver que el estimador ingenuo (2.5) se puede escribir [56, p. 12]

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} w\left(\frac{x - X_i}{h}\right) \quad (2.7)$$

De la función de peso  $w$  se sigue que la estimación se construye colocando una “caja” de ancho  $2h$  y altura  $(2nh)^{-1}$  sobre cada observación y después sumando para obtener

la estimación. Nótese que la sumatoria de (2.7) sólo suma los pesos de las  $X_i$  que caen en  $(x - h, x + h)$ .

El estimador ingenuo puede ser visto como un intento de construir un histograma donde cada punto es el centro del intervalo de muestreo, liberando al histograma de una elección particular de posiciones de rebanadas. La elección del ancho de la rebanada se conserva y está determinada por el parámetro  $h$ , que controla la cantidad por la cual se suavizan los datos para producir la estimación.

De la definición de  $\hat{f}$  se observa que no es continua, sino que tiene brincos en los puntos  $X_i \pm h$  y tiene derivada cero en cualquier otro lugar. Esto da al estimado un carácter que no es solamente indeseable, sino que puede dar al observador no entrenado una impresión engañosa.

### El estimador kernel

Para generalizar el estimador ingenuo, reemplace la función peso  $w$  por una función kernel  $K$  que satisface la condición

$$\int_{-\infty}^{\infty} K(x)dx = 1 \quad (2.8)$$

Usualmente, aunque no siempre,  $K$  será una función de densidad de probabilidad simétrica, por ejemplo, la densidad normal, o la función peso  $w$  usada en el estimador ingenuo. Por analogía con la definición del estimador ingenuo, el estimador kernel  $K$  se define como [56, p. 15]

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (2.9)$$

donde  $h$  es el ancho de la ventana, también llamado el *parámetro de suavizado* o *ancho de banda*.

Como el estimador ingenuo, se puede considerar como una suma de “cajas” centradas en las observaciones, el estimador kernel es una suma de “lóbulos” (protuberancias) colocados en las observaciones. La función kernel  $K$  determina la forma de los lóbulos mientras que el parámetro  $h$  determina el ancho de la ventana.

El efecto de variar el ancho de la ventana, digamos cuando  $h$  tiende a cero es una suma de espigas en las observaciones de la función delta de Dirac, mientras que a medida que  $h$  aumenta, todo el detalle, espurio o de otra manera, se obscurece. Siempre que el kernel  $K$  sea no negativo en cualquier lugar y satisfaga la condición  $\int K(x)dx = 1$ , en otras palabras, que sea una función de densidad de probabilidad, se seguirá de la

definición, que  $\hat{f}$  será una densidad de probabilidad. Además,  $\hat{f}$  heredará todas las propiedades de continuidad y diferenciabilidad del kernel  $K$ .

Aparte del histograma, el estimador kernel es probablemente el más comúnmente usado en aplicaciones de estimación de densidad y es, ciertamente, el más estudiado matemáticamente. Sin embargo, sufre de una ligera desventaja cuando se aplica a datos de distribuciones con colas largas. Debido a que el ancho de la ventana está fijo en toda la muestra, hay una tendencia a aparecer ruido espurio en las colas de las estimaciones; si las estimaciones se suavizan lo suficiente para tratar con este ruido, entonces el detalle esencial, en la parte principal de la distribución, se enmascara.

### 2.1.2. Medidas de discrepancia

Se han estudiado varias medidas de la discrepancia del estimador de densidad  $\hat{f}$  de la densidad verdadera  $f$ . Cuando se considera la estimación en un solo punto, una medida natural es el error cuadrático medio (ECM o MSE del inglés Mean Square Error), definido por [56, p. 35]:

$$ECM_x(\hat{f}) = E \left( \hat{f}(x) - f(x) \right)^2 \quad (2.10)$$

se puede reescribir de la siguiente manera:

$$ECM_x(\hat{f}) = \left( E\hat{f}(x) - f(x) \right)^2 + var \hat{f}(x) \quad (2.11)$$

que es la suma del sesgo al cuadrado y la varianza en  $x$ .

La manera más usada de medir la exactitud global de  $\hat{f}$  como un estimador de  $f$  es el error cuadrático integrado medio (ECIM o MISE del inglés Mean Integrated Square Error) definido por:

$$ECIM(\hat{f}) = E \int \left( \hat{f}(x) - f(x) \right)^2 dx \quad (2.12)$$

Es útil notar las formas alternas de la ecuación anterior:

$$\begin{aligned} ECIM(\hat{f}) &= E \int \left( \hat{f}(x) - f(x) \right)^2 dx \\ &= \int ECM_x(\hat{f}) dx \\ &= \int \left( E\hat{f}(x) - f(x) \right)^2 dx + \int var \hat{f}(x) dx \end{aligned} \quad (2.13)$$



lo que permite interpretar al ECIM como la suma del sesgo cuadrático integrado y la varianza integrada.

Supongamos que el kernel  $K$  es una función simétrica que satisface lo siguiente:

$$\int K(t)dt = 1, \int tK(t)dt = 0 \text{ y } \int t^2K(t)dt = k_2 \neq 0 \quad (2.14)$$

y que la densidad desconocida  $f$  tiene derivadas continuas de todos los órdenes requeridos.

Usualmente, el kernel  $K$  será una función de densidad de probabilidad simétrica, por ejemplo la densidad normal, y la constante  $k_2$  será la varianza de la distribución con esa densidad del kernel.

Una aproximación para el sesgo cuadrático integrado para el ECIM está dada por [56, p. 39]:

$$\int \left( E\hat{f}(x) - f(x) \right)^2 dx = \int sesgo_h(x)^2 dx \approx \frac{1}{4}h^4k_2^2 \int f''(x)^2 dx \quad (2.15)$$

y la aproximación para la varianza es:

$$\int var \hat{f}(x) dx \approx n^{-1}h^{-1} \int K(t)^2 dt \quad (2.16)$$

Supóngase que queremos elegir  $h$  para hacer el ECIM lo más pequeño posible, comparando las dos aproximaciones anteriores se observa uno de los problemas fundamentales de la estimación de densidad. Si en un intento de eliminar el sesgo, se usa un valor muy pequeño de  $h$ , entonces la varianza integrada será grande. De otra manera, eligiendo un valor grande de  $h$  reducirá la variación aleatoria cuantificada por la varianza, a expensas de introducir un error sistemático, o sesgo, dentro de la estimación. Se debe hacer énfasis en que, cualquiera que sea el método de estimación de densidad que se use, la elección del parámetro de suavizado implica un compromiso entre el error sistemático y aleatorio [56, p. 40].

### El ancho de ventana y kernel ideal

El valor ideal de  $h$ , desde el punto de vista de minimizar la aproximación del ECIM:

$$ECIM(\hat{f}) = \int sesgo_h(x)^2 dx + \int var \hat{f}(x) dx \approx \frac{1}{4}h^4k_2^2 \int f''(x)^2 dx + n^{-1}h^{-1} \int K(t)^2 dt \quad (2.17)$$

se puede obtener derivando (2.17) con respecto a  $h$  e igualando la derivada a cero, como

se muestra a continuación. Es una aproximación del ECIM porque está en términos de las aproximaciones del sesgo y varianza. Derivando la aproximación del ECIM

$$\frac{d\left(\frac{1}{4}h^4k_2^2 \int f''(x)^2 dx + n^{-1}h^{-1} \int K(t)^2 dt\right)}{dh} = \frac{4}{4}h^3k_2^2 \int f''(x)^2 dx - n^{-1}h^{-2} \int K(t)^2 dt \quad (2.18)$$

e igualando a cero

$$h^3k_2^2 \int f''(x)^2 dx - n^{-1}h^{-2} \int K(t)^2 dt = 0 \quad (2.19)$$

$$h^5k_2^2 \int f''(x)^2 dx = n^{-1} \int K(t)^2 dt \quad (2.20)$$

$$h^5 = \frac{n^{-1} \int K(t)^2 dt}{k_2^2 \int f''(x)^2 dx} \quad (2.21)$$

finalmente, se obtiene el valor de  $h$  que minimiza el ECIM

$$h_{opt} = k_2^{-2/5} \left\{ \int K(t)^2 dt \right\}^{1/5} \left\{ \int f''(x)^2 dx \right\}^{-1/5} n^{-1/5} \quad (2.22)$$

La fórmula anterior para el ancho de ventana óptimo es algo decepcionante, ya que muestra que  $h_{opt}$  depende de la densidad desconocida que se está estimando. Sin embargo, se pueden obtener algunas conclusiones. Primera, el ancho de ventana ideal converge a cero a medida que el tamaño de la muestra aumenta, pero a una razón muy baja. Segunda, dado que el término  $\left\{ \int f''(x)^2 dx \right\}^{-1/5}$  mide, en algún sentido, la rapidez de las fluctuaciones en la densidad  $f$ , los valores más pequeños de  $h$  serán apropiados para densidades que fluctúan más rápido. Una aproximación natural es elegir  $h$  con referencia a alguna familia estándar de densidades, tal como las densidades normales.

Sustituyendo el valor de  $h$  por  $h_{opt}$  en (2.17), entonces el valor aproximado del ECIM será [56, p. 41]:

$$\frac{5}{4}C(K) \left\{ \int f''(x)^2 dx \right\}^{1/5} n^{-4/5} \quad (2.23)$$

donde la constante  $C(K)$  está dada por:

$$C(K) = k_2^{2/5} \left\{ \int K(t)^2 dt \right\}^{4/5} \quad (2.24)$$

La expresión (2.23) muestra que se debería elegir un kernel  $K$  con un valor pequeño

de  $C(K)$ , ya que ésto hará posible, teóricamente, obtener un valor pequeño del ECIM, si se puede elegir correctamente el parámetro de suavizado. Los kernels que deben considerarse son aquellos que son funciones de densidad de probabilidad, ya que aseguran que la estimación  $\hat{f}$  sea no negativa. Entonces el problema de minimizar  $C(K)$  se reduce a minimizar  $\int K(t)^2 dt$  sujeta a las restricciones de que  $\int K(t) dt$  y  $\int t^2 K(t) dt$  sean ambas iguales a uno. Este problema se resuelve haciendo que  $K(t)$  sea [56, p. 42]:

$$K_e(t) = \begin{cases} \frac{3}{4\sqrt{5}} (1 - \frac{1}{5}t^2) & -\sqrt{5} \leq t \leq \sqrt{5} \\ 0 & \text{de otra manera} \end{cases} \quad (2.25)$$

La notación  $K_e(t)$  se usa porque este kernel lo sugirió primero Epanechnikov [28] en estimación de densidad, de esta manera el kernel de Epanechnikov es óptimo entre los kernels no negativos, en el sentido de minimizar el ECIM más pequeño que se puede alcanzar.

### 2.1.3. Método de kernel multivariado

Los espacios característicos de una imagen muchas veces están caracterizados por grupos de datos muy irregulares cuyo número y formas no están disponibles. Esto sugiere fuertemente que una aproximación no paramétrica, proporcionando una detección confiable del máximo local de la densidad fundamental, i.e., las modas, debe emplearse en el análisis. El método del vecino más cercano no es recomendable para esta tarea por su propensión al ruido local que dificulta la detección de la moda. Además, la estimación obtenida no es una densidad de probabilidad debido a que su integral no es la unidad [56, p. 97] y sus extremos caen asintóticamente a una velocidad muy lenta. Para datos de tamaño bajo a mediano, la estimación por kernel es una buena elección práctica; es sencilla y para kernels que obedecen condiciones suaves, la estimación es asintóticamente no sesgada (significa que el promedio de las estimaciones se aproxima a la función desconocida), consistente en el sentido de la media cuadrada, es decir, el error cuadrático entre la estimación y la función desconocida es cero, y uniformemente consistente en probabilidad, esto es, el estimador converge en probabilidad al valor verdadero de la función desconocida. El detalle de las condiciones del kernel, antes mencionadas, está dado en [33].

La definición del estimador kernel como una suma de "lóbulos" centrados en las observaciones se puede generalizar para el caso multivariado. El estimador de densidad de kernel multivariado con kernel  $K$  y ancho de ventana  $h$  se define por [56, p. 76]:

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n K \left\{ \frac{1}{h}(\mathbf{x} - \mathbf{X}_i) \right\} \quad (2.26)$$

donde la función kernel  $K(\mathbf{x})$  es ahora una función, definida para  $\mathbf{x}$   $d$ -dimensional.

Como se mencionó anteriormente, el kernel óptimo que minimiza el ECIM es el Epanechnikov, que para el caso multivariado está dado por:

$$K_e(\mathbf{x}) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1 - \mathbf{x}^T \mathbf{x}) & \text{si } \mathbf{x}^T \mathbf{x} < 1 \\ 0 & \text{de otra manera} \end{cases} \quad (2.27)$$

donde  $c_d$  es el volumen de la esfera unitaria en el espacio euclidiano  $R^d$   $d$ -dimensional.

El uso de un kernel diferenciable permite definir la estimación del gradiente de densidad como el gradiente del estimador de densidad de kernel multivariado; así de la expresión (2.26):

$$\hat{\nabla} f(\mathbf{x}) \equiv \nabla \hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n \nabla K \left\{ \frac{1}{h}(\mathbf{x} - \mathbf{X}_i) \right\} \quad (2.28)$$

Sustituyendo (2.27) en (2.28):

$$\hat{\nabla} f_E(\mathbf{x}) = \frac{1}{n(h^d c_d)} \frac{d+2}{h^2} \sum_{\mathbf{X}_i \in S_h(\mathbf{x})} [\mathbf{X}_i - \mathbf{x}] = \frac{n_x}{n(h^d c_d)} \frac{d+2}{h^2} \left( \frac{1}{n_x} \sum_{\mathbf{X}_i \in S_h(\mathbf{x})} [\mathbf{X}_i - \mathbf{x}] \right) \quad (2.29)$$

donde la región  $S_h(\mathbf{x})$  es una hiperesfera de radio  $h$  teniendo el volumen  $h^d c_d$ , centrada en  $\mathbf{x}$ , y conteniendo  $n_x$  puntos de datos. El último término de (2.29):

$$M_h(\mathbf{x}) \equiv \frac{1}{n_x} \sum_{\mathbf{X}_i \in S_h(\mathbf{x})} [\mathbf{X}_i - \mathbf{x}] = \frac{1}{n_x} \sum_{\mathbf{X}_i \in S_h(\mathbf{x})} \mathbf{X}_i - \mathbf{x} \quad (2.30)$$

es lo que Fukunaga [33] llama el corrimiento de la media muestral.

La cantidad  $\frac{n_x}{n(h^d c_d)}$  es la estimación  $\hat{f}(\mathbf{x})$  de la densidad del kernel calculada en la hiperesfera  $S_h(\mathbf{x})$  (el kernel uniforme), y así podemos escribir (2.29) como:

$$\hat{\nabla} f_E(\mathbf{x}) = \hat{f}(\mathbf{x}) \frac{d+2}{h^2} M_h(\mathbf{x}) \quad (2.31)$$

donde  $M_h(\mathbf{x})$ :

$$M_h(\mathbf{x}) = \frac{h^2}{d+2} \frac{\hat{\nabla} f(\mathbf{x})}{\hat{f}(\mathbf{x})} \quad (2.32)$$

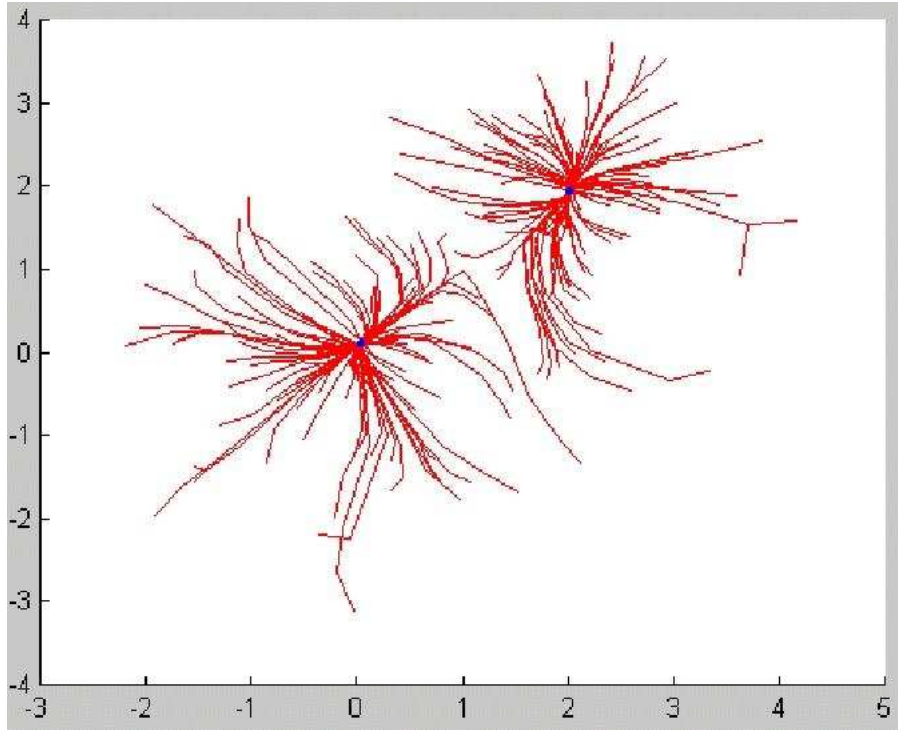


Figura 2.1: Trayectorias definidas por el corrimiento de media para una mezcla de gaussianas con media 0 y 2 y varianza 1.

Esta última expresión muestra que una estimación del gradiente normalizado se puede obtener calculando el corrimiento de media de la muestra con un kernel uniforme centrado en  $\mathbf{x}$ . El vector del corrimiento de media tiene la dirección del gradiente de la estimación de la densidad en  $\mathbf{x}$  cuando esta estimación se obtiene con el kernel Epanechnikov [20].

Si el gradiente es cero, correspondiendo a una densidad uniforme sobre la región  $S_h(\mathbf{x})$ , el promedio del corrimiento de media debe ser cero debido a la simetría de las observaciones cercanas a  $\mathbf{x}$ . Sin embargo, con un gradiente de densidad distinto de cero, apuntando en la dirección de incremento más rápido de la función de densidad de probabilidad, en el promedio más observaciones deben caer a lo largo de su dirección que en otro sitio dentro de  $S_h(\mathbf{x})$ . Correspondientemente, el promedio del corrimiento de media debe apuntar en esa dirección y tener una longitud proporcional a la magnitud del gradiente [33]. Dado que el vector de CM siempre apunta hacia la dirección del máximo incremento en la densidad, puede definir una trayectoria que conduce a un máximo de densidad local, i.e., a una moda de la densidad, como se aprecia en la Figura 2.1.

### 2.1.4. Convergencia para el kernel de Epanechnikov

En el corrimiento de media, la media local es desplazada hacia la región en que residen la mayoría de los datos y puede definir una trayectoria hacia un punto estacionario de la densidad estimada, como se mencionó anteriormente, estos puntos estacionarios son la modas de la densidad. En la búsqueda de las modas, el corrimiento de media tiene un comportamiento adaptable, el desplazamiento del corrimiento de media es grande para regiones de baja densidad y disminuye conforme el dato  $\mathbf{x}$  se aproxima a la moda. En el procedimiento del corrimiento de media se garantiza la convergencia a un punto cercano donde la estimación tiene gradiente cero. La convergencia ha sido probada por Cheng [13] y Comaniciu [20], la siguiente prueba de convergencia es tomada de esta última referencia.

Sea  $\mathbf{y}_j$ ,  $j=1, 2, \dots$  la secuencia de localidades sucesivas del procedimiento de corrimiento de media. Por definición tenemos para cada  $j = 1, 2, \dots$

$$\mathbf{y}_{j+1} = \frac{1}{n_j} \sum_{\mathbf{x}_i \in S_h(\mathbf{y}_j)} \mathbf{x}_i \quad (2.33)$$

donde  $\mathbf{y}_1$  es el centro de la ventana inicial y  $n_j$  es el número de puntos que caen en la ventana  $S_h(\mathbf{y}_j)$  centrada en  $\mathbf{y}_j$ .

La convergencia del corrimiento de media se ha justificado en [13] como una consecuencia de la relación (2.32). Sin embargo, mientras es verdad que el vector de CM,  $M_h(\mathbf{x})$ , tiene la dirección del gradiente de la estimación de densidad en  $\mathbf{x}$ , no es aparente que la estimación de densidad en las localidades  $\mathbf{y}_j$ ,  $j=1, 2, \dots$  sea una secuencia monótonica incremental. El movimiento en la dirección del gradiente garantiza el ascenso por pasos infinitesimales. El siguiente teorema afirma la convergencia para datos discretos.

**TEOREMA.** *Sea  $\hat{f}_E = \{\hat{f}_E(j)\}_{j=1,2,\dots} \equiv \{\hat{f}_E(\mathbf{y}_j)\}_{j=1,2,\dots}$  la secuencia de estimaciones de densidad obtenida usando el kernel Epanechnikov y calculada en los puntos  $\mathbf{y}_j$ ,  $j=1, 2, \dots$  definida por las localidades sucesivas del procedimiento de corrimiento de media con el kernel uniforme. La secuencia es convergente.*

**Prueba.** Dado que el conjunto de datos  $\mathbf{X}_i$ ,  $j=1, 2, \dots$  tiene cardinalidad finita  $n$ , la secuencia  $\hat{f}$  esta acotada. Además, se mostrará que es estrictamente monótonica incremental, i.e., si  $\mathbf{y}_j \neq \mathbf{y}_{j+1}$  entonces  $\hat{f}_E(j) < \hat{f}_E(j+1)$ , para todo  $j = 1, 2, \dots$

Sea  $n_j, n'_j, n''_j$  con  $n_j = n'_j + n''_j$  el número de puntos de datos que caen dentro de las ventanas  $d$ -dimensional  $S_h(\mathbf{y}_j)$ ,  $S'_h(\mathbf{y}_j) = S_h(\mathbf{y}_j) - S''_h(\mathbf{y}_j)$ , y  $S''_h(\mathbf{y}_j) = S_h(\mathbf{y}_j) \cap S_h(\mathbf{y}_{j+1})$ .

Sin pérdida de generalidad se puede suponer el origen localizado en  $\mathbf{y}_j$ . Usando la definición de la estimación de densidad (2.26) con el kernel de Epanechnikov (2.27) y notando que  $\|\mathbf{y}_j - \mathbf{X}_i\|^2 = \|\mathbf{X}_i\|^2$  se tiene:

$$\begin{aligned}
\hat{f}_E(j) = \hat{f}_E(\mathbf{y}_j) &= \frac{1}{nh^d} \sum_{\mathbf{X}_i \in S_h(\mathbf{y}_j)} K_e \left( \frac{\mathbf{y}_j - \mathbf{X}_i}{h} \right) \\
&= \frac{d+2}{2n(h^d c_d)} \sum_{\mathbf{X}_i \in S_h(\mathbf{y}_j)} \left[ 1 - \frac{\|\mathbf{X}_i\|^2}{h^2} \right]
\end{aligned} \tag{2.34}$$

Dado que el kernel  $K_e$  es no negativo se tiene también que

$$\begin{aligned}
\hat{f}_E(j+1) = \hat{f}_E(\mathbf{y}_{j+1}) &\geq \frac{1}{nh^d} \sum_{\mathbf{X}_i \in S'_h(\mathbf{y}_j)} K_e \left( \frac{\mathbf{y}_{j+1} - \mathbf{X}_i}{h} \right) \\
&= \frac{d+2}{2n(h^d c_d)} \sum_{\mathbf{X}_i \in S'_h(\mathbf{y}_j)} \left[ 1 - \frac{\|\mathbf{y}_{j+1} - \mathbf{X}_i\|^2}{h^2} \right]
\end{aligned} \tag{2.35}$$

Por lo tanto, sabiendo que  $n'_j = n_j - n''_j$  se obtiene

$$\hat{f}_E(j+1) - \hat{f}_E(j) \geq \frac{d+2}{2n(h^d c_d)h^2} \left[ \sum_{\mathbf{X}_i \in S_h(\mathbf{y}_j)} \|\mathbf{X}_i\|^2 - \sum_{\mathbf{X}_i \in S''_h(\mathbf{y}_j)} \|\mathbf{y}_{j+1} - \mathbf{X}_i\|^2 - n'_j h^2 \right] \tag{2.36}$$

donde el último término aparece debido a los diferentes límites de sumatoria. También por definición  $\|\mathbf{y}_{j+1} - \mathbf{X}_i\|^2 \geq h^2$  para todo  $\mathbf{X}_i \in S'_h(\mathbf{y}_j)$ , lo que implica que

$$\sum_{\mathbf{X}_i \in S'_h(\mathbf{y}_j)} \|\mathbf{y}_{j+1} - \mathbf{X}_i\|^2 \geq n'_j h^2 \tag{2.37}$$

Finalmente, empleando (2.37) en (2.36) y usando (2.33) se obtiene

$$\begin{aligned}
\hat{f}_E(j+1) - \hat{f}_E(j) &\geq \frac{d+2}{2n(h^d c_d)h^2} \left[ \sum_{\mathbf{X}_i \in S_h(\mathbf{y}_j)} \|\mathbf{X}_i\|^2 - \sum_{\mathbf{X}_i \in S_h(\mathbf{y}_j)} \|\mathbf{y}_{j+1} - \mathbf{X}_i\|^2 \right] \\
&= \frac{d+2}{2n(h^d c_d)h^2} \left[ 2\mathbf{y}_{j+1}^T \sum_{\mathbf{X}_i \in S_h(\mathbf{y}_j)} \mathbf{X}_i - n_j \|\mathbf{y}_{j+1}\|^2 \right] \\
&= \frac{d+2}{2n(h^d c_d)h^2} n_j \|\mathbf{y}_{j+1}\|^2
\end{aligned} \tag{2.38}$$

La última parte de la relación (2.38) es estrictamente positiva excepto cuando  $\mathbf{y}_j = \mathbf{y}_{j+1} = 0$ . Y siendo limitada y estrictamente monotónica incremental, la secuencia  $\hat{f}_E$  es convergente. Note que si  $\mathbf{y}_j = \mathbf{y}_{j+1}$  entonces  $\mathbf{y}_j$  es el límite de  $\hat{f}_E$ , i.e.,  $\mathbf{y}_j$  es el punto fijado por el procedimiento de corrimiento de media.

### 2.1.5. Agrupamiento por corrimiento de media

Un método para agrupar un conjunto de datos en diferentes clases puede ser asignando cada dato a la moda más cercana a lo largo de la dirección del gradiente en los puntos de observación. Para lograr esto, se puede correr cada dato por alguna cantidad proporcional al gradiente en la dirección del mismo, e iterativamente repetir el proceso sobre las datos transformados hasta que los grupos (clases) bien formados estén cerca de las modas. Después de cada iteración, cada uno de los datos se habrá movido más cerca de su moda o centro de clase [33].

El procedimiento del corrimiento de media es una técnica de optimización ascendente con el mejor paso, con un tamaño de paso variable que es proporcional a la magnitud del gradiente, que calcula el vector del corrimiento de media (2.32) para cada uno de los datos, corre o traslada el kernel por esta cantidad, y repite los cálculos hasta que se alcanza una moda. Así, los datos se dividen en grupos basándose solamente en sus trayectorias del corrimiento de media, y cuando dos puntos de datos convergen a la misma posición final, se considera que pertenecen al mismo grupo.

El algoritmo de corrimiento de media es el ascenso más empinado sobre la densidad de los datos. Cada centro de clase escala la cuesta en la superficie de densidad independientemente. Por lo tanto, si los datos o su densidad no cambian durante la ejecución del algoritmo, la convergencia de la evolución de los centros de clase es una consecuencia de la convergencia del ascenso más empinado para centros de clase individuales.

En un proceso con un kernel ancho, los datos convergen a una sola posición. En el otro extremo, cuando se trunca el kernel al grado de cubrir solo un dato en alguna posición, el dato inicial es un punto fijo del proceso y así no se lleva a cabo la fusión. Cuando el tamaño del kernel está entre estos dos extremos, los datos pueden tener trayectorias que los fusionan hacia un número variable de centros de clase.

A diferencia del k-means [55] y otros algoritmos de agrupamiento, que son probabilísticos por tener una inicialización aleatoria de los centros de los grupos, el agrupamiento por corrimiento de media con los centros inicializados con los datos es determinístico. Como un proceso determinístico, el resultado del agrupamiento proporciona una característica del conjunto de datos. El resultado o propósito del agrupamiento por corrimiento de media es usar un máximo local como una caracterización de los datos [13].



El algoritmo de corrimiento de media tiene la propiedad de ser un mapeo de gradiente, o similar a una búsqueda de modas de una función real. Comparado con los métodos de ascenso o descenso por gradiente, el corrimiento de media es más efectivo en términos de adaptar el tamaño del paso de ascenso que es igual a la magnitud del gradiente, el paso de ascenso es grande para regiones de baja densidad y se acorta conforme el dato  $\mathbf{x}$  se aproxima a la moda.

## 2.2. Información espacial

La segmentación y filtrado de una imagen conservando los bordes de alta calidad se pueden obtener aplicando el corrimiento de media en el dominio combinado espacial-rango. El método está basado en la misma idea de correr iterativamente una ventana de tamaño fijo hacia el promedio de los datos dentro de ella. Los detalles de la imagen se conservan debido al carácter no paramétrico del análisis, que no supone a priori alguna estructura particular de los datos.

La segmentación basada en agrupamiento (clustering) supone estructuras constantes por tramos; sin embargo, cuando se tienen cambios lentos de intensidad se puede generar un grupo alargado, cuya recuperación es difícil debido a su aumento de probabilidad de fusionarse con otros grupos. Un problema similar aparece cuando las regiones de densidad alta, en el espacio característico, no son necesariamente significativas en la imagen, estos tipos de datos tienen un impacto negativo en la segmentación basada en agrupamiento. Este impacto se puede eliminar tomando en cuenta la información espacial, i.e., procesando en el dominio espacial-rango.

### 2.2.1. Dominio espacial-rango

Una imagen típicamente se representa como un enrejado (lattice) 2-dimensional de vectores  $r$ -dimensionales (píxeles), donde  $r$  es 1 en el caso de niveles de gris, 3 en imágenes de color, o mayor que 3 en el caso multiespectral. El espacio del enrejado se conoce como el dominio *espacial*, mientras que el nivel de gris, color, o información espectral se representa en el dominio del *rango*. Los anchos de ventana del kernel en los dominios espacial y de rango se pueden concatenar para obtener un dominio espacial-rango de dimensión  $d = r + 2$ .

Al aplicar el algoritmo de corrimiento de media en el dominio conjunto espacial-rango, cada dato se asocia a un punto de convergencia que representa la moda local de la densidad en el espacio  $d$ -dimensional. En el proceso se definen el ancho de ven-

tana espacial ( $h_e$ ), y el ancho de ventana de intensidad ( $h_i$ ) correspondiendo al rango, tomando en cuenta simultáneamente tanto la información espacial y de rango.

La salida del filtro de corrimiento de media para un pixel de la imagen está definida como la información de rango alcanzada por el punto de convergencia. Este proceso logra un filtrado espacial de alta calidad que preserva las discontinuidades.

### 2.2.2. Filtrado

El filtrado es una operación fundamental en el procesamiento de imágenes y visión por computadora. El valor de la imagen filtrada, en una ubicación dada, es una función de los valores de entrada de la imagen en un vecindario pequeño de la misma ubicación. Las imágenes por lo general varían lentamente a través del espacio, de tal manera que los pixeles cercanos probablemente tienen valores similares, y por lo tanto es apropiado promediarlos. Los valores del ruido que corrompen estos pixeles cercanos están menos correlacionados que los valores de la imagen, así que el ruido disminuye con la promediación mientras se conservan los valores de la imagen.

El suavizado o filtrado por el reemplazo del pixel en el centro de una ventana por el promedio de los pixeles en la ventana, remueve el ruido pero también los bordes de la imagen, volviéndolos borrosos con el filtrado. Para resolver el problema de suavizar los bordes, se requieren técnicas de suavizado que preserven las discontinuidades, reduciendo de manera adaptable la cantidad de suavizado cerca de los cambios abruptos en la estructura local, esto es, en los bordes. El corrimiento de media es una técnica de suavizado que preserva los bordes, razón por la que se puede utilizar como operación de filtrado de una imagen, como se muestra en el siguiente algoritmo.

Sean  $\mathbf{x}_i$  y  $\mathbf{z}_i$  los puntos  $d$ -dimensionales de la imagen original y filtrada en el dominio espacial-rango. Para cada  $i = 1 \dots n$

1. Inicializar  $j = 1$  e  $\mathbf{y}_j = \mathbf{x}_i$ .
2. Calcular  $\mathbf{y}_{j+1} = \frac{1}{n_j} \sum_{\mathbf{x} \in S_h(\mathbf{y}_j)} \mathbf{x}$ ,  $j = j + 1$  hasta la convergencia.
3. Asignar  $\mathbf{z}_i = (\mathbf{x}_i^e, \mathbf{y}_{conv}^r)$ , donde los superíndices  $e$  y  $r$  corresponden al dominio espacial y al de rango.

La última asignación especifica que el dato filtrado en la ubicación espacial de  $\mathbf{x}_i$  tendrá las componentes de rango del punto de convergencia  $\mathbf{y}_{conv}$ . El número de puntos en la ventana  $S_h(\mathbf{y}_j)$  de radio  $h$  y centrada en  $\mathbf{y}_j$  es  $n_j$ .

Se obtiene una estimación de gradiente óptima asintóticamente cuando la distribución en el espacio conjunto es normal (esto es, gaussiana). El radio de la ventana de

búsqueda es una función del número de datos  $n$ , en nuestro caso los datos están lejos de ser normales. Por lo tanto, no se pueden imponer restricciones teóricas sobre los valores del ancho de ventana espacial,  $h_e$ , y del ancho de ventana de intensidad,  $h_i$ , que son dependientes de la tarea y en configuraciones prácticas su elección requiere que se maneje por este conocimiento.

Además del filtrado por corrimiento de media, otra técnica de suavizado con preservación de bordes es el filtrado bilateral [58]. Aunque ambos métodos se basan en el mismo principio: el procesamiento simultáneo en el dominio conjunto espacial-rango, el filtrado bilateral usa una ventana estática en el dominio conjunto, mientras que el corrimiento de media usa una ventana dinámica, moviéndose en la dirección de máximo incremento del gradiente. Por lo tanto, el filtrado por corrimiento de media se adapta mejor a la estructura local de los datos [20].

### 2.2.3. Segmentación por corrimiento de media

Para realizar la tarea de segmentación, los puntos de convergencia suficientemente cercanos en el dominio conjunto son considerados como los centros de clase, y todos aquellos puntos que fueron llevados por el corrimiento de media al mismo punto, i.e., a una moda de la densidad, se fusionan para obtener las regiones homogéneas de la imagen.

La segmentación por corrimiento de media en el dominio espacial-rango tiene el mismo diseño sencillo que el proceso de filtrado. Sean  $\mathbf{x}_j$ , con  $j = 1, \dots, n$  los datos de la imagen original,  $\mathbf{z}_j$ , con  $j = 1, \dots, n$  los puntos de convergencia, y  $E_j$  con  $j = 1, \dots, n$  un conjunto de etiquetas, entonces el procedimiento de segmentación es el siguiente:

1. Para cada  $j = 1, \dots, n$  calcular el vector del corrimiento de media para cada  $\mathbf{x}_j$  y almacenar el punto de convergencia en  $\mathbf{z}_j$ .
2. Identificar los grupos (clases)  $\mathbf{C}_p$ ,  $p = 1, \dots, m$  de los puntos de convergencia uniendo todos los  $\mathbf{z}_i$  que están más cerca de 0.5 uno de otro en el dominio conjunto.
3. Para cada  $j = 1, \dots, n$  asignar  $E_j = \{p \mid \mathbf{z}_j \in \mathbf{C}_p\}$ .
4. Eliminar las regiones espaciales que contienen menos de  $M$  píxeles.

El primer paso de la segmentación es un proceso de filtrado. Toda la información acerca del punto de convergencia  $d$ -dimensional se almacena en  $\mathbf{z}_i$ , no solo su parte del rango. Nótese que el número de grupos  $m$  está controlado por los radios del kernel  $h_e$  y  $h_i$ .

## 2.3. Corrimiento de media ponderado

La estimación por corrimiento de media (CM) se puede mejorar ponderando cada pixel dentro de una región por una función de su confianza de borde, tal que los pixeles que estén situados cerca de un borde (confianza de borde cercana a uno) influyan menos en la determinación del nuevo centro del agrupamiento. La ecuación del CM modificada, que incluye la ponderación de la confianza del borde y derivada de la ecuación (2.30) es:

$$M_h(\mathbf{x}) = \frac{1}{\sum (1 - \varphi_i)} \sum_{\mathbf{X}_i \in S_h(\mathbf{x})} (1 - \varphi_i) \mathbf{X}_i - \mathbf{x} \quad (2.39)$$

donde  $\varphi_i$  es la confianza del borde asociado a  $\mathbf{X}_i$ , la manera de calcular esta confianza se muestra en el siguiente capítulo.

El vector de corrimiento de media, tal como está expresado en (2.39), se usará en la etapa del filtrado para obtener regiones homogéneas en intensidad de las imágenes procesadas. La importancia de utilizar el procedimiento de corrimiento de media, radica en que es una herramienta versátil para el análisis del espacio característico y puede producir soluciones confiables. El primer paso en el análisis de un espacio característico con densidad fundamental  $f(\mathbf{x})$  es encontrar las modas de esta densidad; y el procedimiento del corrimiento de media es un manera de localizar las modas sin estimar la densidad. La propiedad de convergencia del corrimiento de media se debe a su magnitud adaptable que elimina la necesidad de procedimientos adicionales para elegir el tamaño de paso adecuado. Esto es una ventaja sobre los métodos tradicionales basados en gradiente. Por último, al filtrar con corrimiento de media ponderado, los pesos de los pixeles cercanos a un borde son pequeños, lo que permite realzar la propiedad de preservación de discontinuidad.

La descripción de la implementación del corrimiento de media, ponderado con la confianza del borde, se describe en el Apéndice C.

# Capítulo 3

## Mapas de confianza de bordes

Una de las técnicas más importantes en la extracción de características de una imagen es la detección de bordes, y dentro de la detección de bordes los métodos más usados están basados en la orientación del gradiente, realizando los siguientes pasos: estimación de la orientación y magnitud del gradiente, supresión no máxima y umbralización por histéresis. Estos métodos [37, 52, p. 75] se basan en la coherencia de los bordes cercanos a la dirección del gradiente y su problema es la falta de una manera de extraer una dirección de borde confiable. Este problema se puede resolver utilizando una medida de confianza del borde en los pasos involucrados en su detección como se propone en [48]. La estimación del gradiente se realiza usando dos máscaras de diferenciación empleadas como operadores de ventana, permitiendo también estimar la orientación del gradiente y con ésto generar una plantilla de borde ideal con la misma orientación. La medida de confianza se determina a través de la correlación de los datos con la plantilla de borde ideal. A continuación se describe el proceso para determinar los mapas de confianza de bordes, descripción que ha sido tomada de manera resumida y parcial del trabajo de Meer y Georgescu [48].

### 3.1. Medida de confianza

Una operación realizada con frecuencia en el procesamiento de imágenes es calcular el promedio ponderado de los datos en una ventana de  $(2m+1) \times (2m+1)$  deslizándola sobre la imagen. Usando los datos  $a_{ij}$  y los pesos  $w_{ij}$ , con  $i, j = -m, \dots, 0, \dots, m$ , como los elementos del  $i$ -ésimo renglón y  $j$ -ésima columna, se pueden definir las matrices de datos  $\mathbf{A}$  y de ponderación  $\mathbf{W}$ , siendo esta última la máscara empleada como operador de ventana. Estas dos matrices se combinan para obtener la salida asociada con el centro de la ventana, es decir en la ubicación de la muestra que corresponde a las coordenadas

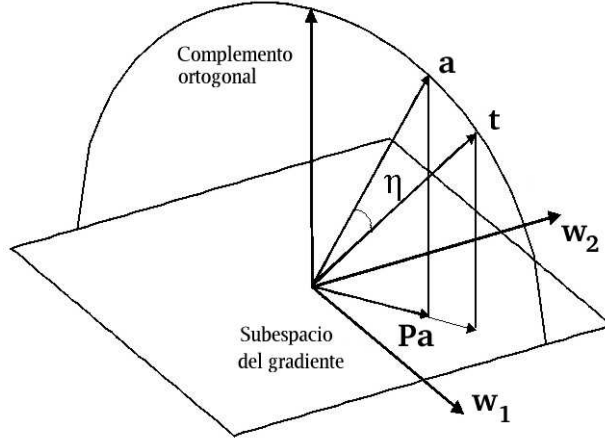


Figura 3.1: Determinación y definición de la medida de confianza

de la ventana  $i = j = 0$

$$salida = \text{traza}[\mathbf{W}^T \mathbf{A}] = \text{traza}[\mathbf{W} \mathbf{A}^T] \quad (3.1)$$

donde se han usado las propiedades de invarianza de la traza de una matriz. La salida del operador de ventana también se puede escribir como un producto interior vectorial, donde los vectores  $\mathbf{a} = \text{vec}[\mathbf{A}]$  y  $\mathbf{w} = \text{vec}[\mathbf{W}]$  se obtienen apilando las columnas de la matriz correspondiente:

$$salida = \mathbf{w}^T \mathbf{a} = \mathbf{a}^T \mathbf{w} \quad (3.2)$$

En  $R^{(2m+1)^2}$  el vector  $\mathbf{w}$  define un subespacio de una dimensión y sea  $W_{\perp}$  su complemento ortogonal  $[(2m+1)^2 - 1]$ -dimensional, esto es, el conjunto de todos los vectores ortogonales a  $\mathbf{w}$ . Ya que para cualquier  $\mathbf{b} \in W_{\perp}$  la salida del operador ventana es 0, tal que el dato es invisible para el operador ventana, se tiene:

$$salida = \mathbf{w}^T (\mathbf{a} + \mathbf{b}) = \mathbf{w}^T \mathbf{a} \quad (3.3)$$

mostrando que un número muy grande de vectores de datos (vecindarios de la imagen) producen la misma respuesta. Por ejemplo, con frecuencia se observa que el operador gradiente puede dar una respuesta espuria grande en un vecindario aparentemente no estructurado. En la práctica se requiere combinar la salida de varios operadores de ventana, por ejemplo, el gradiente se estima usando dos máscaras de diferenciación. El procedimiento que se describe a continuación para dos máscaras, se puede aplicar de la misma manera, a cualquier número y tipo de máscaras.

Sean  $\mathbf{w}_1$  y  $\mathbf{w}_2$  los vectores que corresponden a las dos máscaras de diferenciación, definiendo un hiperplano en  $R^{(2m+1)^2}$  y sea  $W_{\perp}$  el complemento ortogonal  $[(2m+1)^2-2]$ -dimensional de este plano. Como se observa en la Figura 3.1, las máscaras y los datos en la ventana se pueden representar por tres vectores en  $R^{(2m+1)^2}$ ; cada vector de las máscaras define una dimensión del hiperplano, y los datos son un vector arbitrario en el plano del complemento ortogonal del hiperplano. De la ecuación 3.3,  $W_{\perp}$  es el espacio nulo del operador gradiente. El proyector sobre el subespacio del operador gradiente es la matriz de  $(2m+1)^2 \times (2m+1)^2$ :

$$\mathbf{P} = \frac{\mathbf{w}_1 \mathbf{w}_1^{\top}}{\mathbf{w}_1^{\top} \mathbf{w}_1} + \frac{\mathbf{w}_2 \mathbf{w}_2^{\top}}{\mathbf{w}_2^{\top} \mathbf{w}_2} \quad (3.4)$$

Sin pérdida de generalidad, se puede suponer que el dato está normalizado a un vector unitario  $\|\mathbf{a}\| = 1$ , así su proyección sobre el plano del operador gradiente es el vector  $\mathbf{Pa}$ . La definición de  $\mathbf{w}_1$  y  $\mathbf{w}_2$  implica que la orientación de  $\mathbf{Pa}$  en el plano es la orientación estimada del gradiente  $\hat{\theta}$ , que es el ángulo entre la proyección del dato y uno de los vectores de las máscaras. El parámetro  $\hat{\theta}$  se puede usar para definir una plantilla de borde ideal  $\mathbf{t}$ , con la misma orientación del gradiente estimado. El vector plantilla siempre está situado en el plano  $\langle \mathbf{a}, \mathbf{Pa} \rangle$  en algún lugar fuera del hiperplano. De la Figura 3.1 se obtiene la definición de la medida de confianza para la presencia de un borde en el dato analizado:

$$\eta = |\mathbf{t}^{\top} \mathbf{a}| \quad (3.5)$$

donde  $\eta$  es el valor absoluto del coseno entre el ángulo que forman  $\mathbf{t}$  y  $\mathbf{a}$ , ya que ambos son vectores unitarios, y en el dominio de la imagen es el valor absoluto del coeficiente de correlación entre los datos y la plantilla normalizados. La medida de confianza incorpora información tanto de los datos como de la plantilla que no están en el subespacio del gradiente, siendo independiente de la magnitud del gradiente para la presencia de un modelo de borde en la ventana de procesamiento.

## 3.2. Estimación del gradiente

En el dominio discreto, solamente los datos de las muestras están disponibles y para obtener el gradiente en dos dimensiones, las dos derivadas parciales se deben calcular por diferenciación numérica. Muchas máscaras de diferenciación son separables, y los pesos se obtienen del producto exterior de dos secuencias unidimensionales  $s(i)$  y  $d(j)$ ,  $i, j = -m, \dots, 0, \dots, m$ ; así las máscaras de diferenciación están dadas por:

$$\mathbf{W} = \mathbf{s} \mathbf{d}^\top \quad (3.6)$$

Debido a condiciones de ruido en los datos, la diferenciación a lo largo de una dimensión (por ejemplo el eje horizontal,  $x$ , correspondiendo a la columna  $j$ ) se debe combinar con un esquema de suavizado en la otra dimensión (el eje vertical,  $y$ , correspondiendo al renglón  $i$ ). Sean  $d(j)$  con  $j = -m, \dots, 0, \dots, m$ , los pesos obtenidos por la diferenciación numérica del  $i$ -ésimo renglón de la matriz de datos  $\mathbf{A}$  en una ventana de  $n \times n$ , ( $n = 2m+1$ ). Así, el promedio ponderado es la estimación de la primera derivada en la ubicación  $(i, 0)$  de la ventana. Similarmente, sean  $s(i)$  con  $i = -m, \dots, 0, \dots, m$ , los pesos obtenidos por la operación de suavizado de la  $j$ -ésima columna. El resultado del promedio ponderado es el valor suavizado en la ubicación  $(0, j)$ . Ambas secuencias se definen usando los polinomios de Krawtchouk [49, 43, 31] para aproximar la estructura local de una superficie continua, los polinomios se eligen de filtros de diferenciación suavizados, que producen una ponderación binomial, definida por:

$$w(i) = \frac{1}{2^{2m}} \binom{2m}{m+i} = \frac{1}{2^{2m}} \frac{(2m)!}{(m-i)!(m+i)!} \quad (3.7)$$

donde  $i = -m, \dots, 0, \dots, m$ . La secuencia  $h(i; r, p)$  es el filtro para estimar la  $r$ -ésima derivada cuando se supone un polinomio de grado  $p$  para la estructura fundamental. Por lo tanto, la secuencia de suavizado está dada por  $s(i) = h(i; 0, 0) = w(i)$ , y la de diferenciación por  $d(j) = h(j; 1, 1) = (2j/m)w(j)$ . La combinación de dos filtros en un producto exterior proporciona operadores de ventana en 2D. Así, la ponderación de datos (usando los polinomios de Krawtchouk) y suavizado a lo largo de una coordenada, mientras se calcula la primera derivada a lo largo de la otra coordenada, produce un operador gradiente muy similar al detector de borde de Canny [11, 24]. Los coeficientes de los polinomios de Krawtchouk, para las secuencias de suavizado y diferenciación que se utilizan, se obtienen con una  $m = 2$  y son las siguientes:

$$\begin{aligned} s(i) &= [0.0625 \ 0.2500 \ 0.3750 \ 0.2500 \ 0.0625]^\top \\ d(j) &= [-0.1250 \ -0.2500 \ 0.000 \ 0.2500 \ 0.1250]^\top \end{aligned}$$

La máscara  $\mathbf{W}$  realiza la diferenciación numérica a lo largo de los renglones de los datos, seguida por el suavizado de los resultados, implementando  $\partial/\partial x$ . La diferenciación a lo largo de las columnas seguida por el suavizado, implementando  $\partial/\partial y$ , se obtiene con la máscara  $\mathbf{W}^\top = \mathbf{d} \mathbf{s}^\top$ . Esta definición corresponde a las coordenadas de ventana, donde el eje positivo  $x$  apunta hacia la derecha, y el eje positivo  $y$  apunta



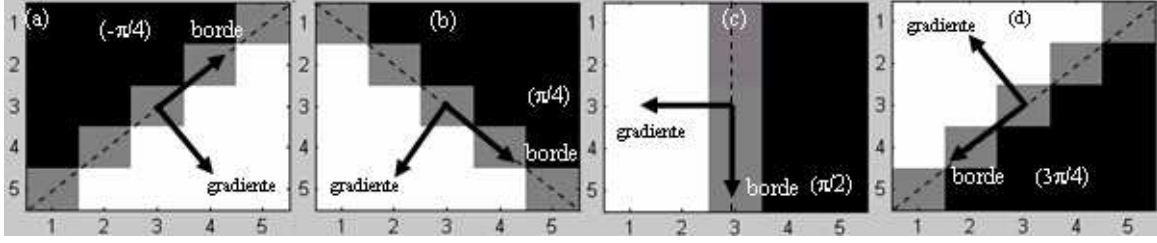


Figura 3.2: Ejemplos de orientación de la plantilla de borde ideal para  $m=2$ . El coeficiente de correlación entre la plantilla ideal y la matriz de datos, para el mismo ancho de ventana, determina la medida de confianza.

hacia abajo. Esta orientación de los ejes se muestra en la Figura 3.2, donde el sistema de coordenadas  $(x, y)$  es de mano izquierda, es decir, una rotación de  $+90^\circ$  desde el eje positivo  $x$  al eje positivo  $y$  es en el sentido de las manecillas del reloj.

La magnitud del gradiente estimado es:

$$\hat{g} = \sqrt{\text{traza}^2(\mathbf{W}^\top \mathbf{A}) + \text{traza}^2(\mathbf{WA})} \quad (3.8)$$

y la orientación del gradiente estimado está dada por:

$$\hat{\theta} = \tan^{-1} \left( \frac{\text{traza}[\mathbf{WA}]}{\text{traza}[\mathbf{W}^\top \mathbf{A}]} \right) \quad (3.9)$$

El modelo de borde empleado es el borde de paso ideal pasando por el centro del vecindario y orientado en  $-180^\circ \leq \hat{\theta}_e < 180^\circ$ . El valor de un pixel se calcula integrando a través de su sección transversal de área unitaria, y así, la forma de la región de transición depende de  $\hat{\theta}_e$ . El vector del gradiente siempre apunta hacia la parte más alta de intensidad de la región, como se observa en la Figura 3.2, y la orientación del borde se deriva de 3.9 de la siguiente manera:

$$\hat{\theta}_e = \hat{\theta} - 90 = -\tan^{-1} \left( \frac{\text{traza}[\mathbf{W}^\top \mathbf{A}]}{\text{traza}[\mathbf{WA}]} \right) \quad (3.10)$$

De esta manera, las plantillas son modelos de borde ideal con orientación  $\hat{\theta}_e$ , derivadas de la orientación del gradiente estimado. La medida de confianza ( $\eta$ ) se calcula para cada dato, correlacionando su estimación de borde y la plantilla de borde ideal correspondiente. El cálculo se lleva a cabo integrando a través de su sección transversal de área unitaria, y normalizando el modelo para una matriz de datos  $\mathbf{A}$  con media cero y norma de Frobenius igual a uno.

### 3.3. Mapa de confianza

Después de la estimación del gradiente, cada pixel de la imagen cuenta con una magnitud (o gradiente) de borde  $\hat{g}$ , y una orientación de borde  $\hat{\theta}_e$ . Con los rangos normalizados de los valores de la magnitud del gradiente se calcula su función de distribución acumulativa empírica, la cual es más conveniente de usar como criterio de selección para la presencia de un borde. Sean  $\hat{g}_1 < \dots < \hat{g}_k < \hat{g}_{k+1} < \dots < \hat{g}_N$  el conjunto ordenado de los distintos valores de magnitud del gradiente estimado, entonces para cada dato su magnitud de borde  $\hat{g}_k$  se reemplaza con la probabilidad:

$$\rho_k = \text{prob}[\hat{g} \leq \hat{g}_k] \quad (3.11)$$

Los rangos normalizados  $\rho_k$ , son los percentiles de la distribución acumulativa de los valores de magnitud del gradiente. Para cada pixel se tienen ahora dos cantidades disponibles:  $\rho$  y  $\eta$  con valores entre 0 y 1. El primero caracteriza la magnitud del gradiente estimado, y el último la confianza de estar en la presencia de un patrón de borde orientado de acuerdo a la orientación del gradiente estimado.

Finalmente, el mapa de confianza se calcula para cada pixel como una combinación lineal de la función de distribución acumulativa empírica y la medida de confianza del borde:

$$\varphi = \beta\rho + (1 - \beta)\eta \quad (3.12)$$

donde  $\beta$  es una constante entre 0 y 1 que controla la mezcla de la magnitud del gradiente,  $\rho$ , y la información del patrón local,  $\eta$ . La descripción para obtener el mapa de confianza, se describe en el Apéndice B, poniendo énfasis en los aspectos de la implementación.

# Capítulo 4

## Fusión de regiones

Una vez que se tiene la imagen filtrada, las regiones se deben delinear en el espacio de características, agrupando todas aquellas regiones que están más cercanas dentro del dominio espacial y del dominio de rango, es decir, todos aquellos pixeles que hayan convergido al mismo punto se fusionan y etiquetan con un valor escalar. La delineación de clases se puede refinar incorporando un grafo de adyacencias de región (GAR), que posteriormente se puede contraer con un algoritmo de *encontrar y unir*. La propuesta de fusión de regiones por contracción de un grafo está planteada en [15], y los conceptos sobre grafos y algoritmos para procesarlos fueron tomados del libro [54].

### 4.1. Contracción del grafo

Dados los parámetros de resolución del estimador por kernel, el número de regiones determinado por el procedimiento del CM ponderado puede ser arbitrariamente grande, produciendo una sobre segmentación de la imagen (más regiones de las que realmente tiene). Para fusionar a la vez las regiones adyacentes homogéneas que han sido divididas en partes separadas por el procedimiento del CM, se puede realizar una operación de contracción sobre el GAR.

La representación más directa de un grafo es por medio de una *matriz de adyacencias de regiones* (MAR) [54, p. 419]. En la MAR se almacena un valor booleano para establecer si hay una arista (una conexión) entre los vértices  $x$  e  $y$ , para construirla se toma la imagen filtrada y se etiqueta, considerando cada etiqueta (esto es, cada región) como un vértice y poniendo un valor de 1 en aquellos vértices con los que comparte una adyacencia, es decir, donde existe una arista y un 0 de otra manera. En el dominio de la imagen, el valor 1 en la MAR determina las regiones que comparten al menos un borde entre ellas, identificando de esta manera el vecindario de cada región.

La determinación de adyacencias entre vértices está basada en un proceso de conectividad 4 en la cuadrícula de la imagen. La representación de la matriz de adyacencias es satisfactoria solamente si el grafo a procesarse es denso: si  $V$  es el número de vértices, la matriz requiere  $V^2$  bits de almacenamiento y  $V^2$  pasos para inicializarla. Si el número de aristas es proporcional a  $V^2$ , entonces la representación es aceptable porque se requieren cerca de  $V^2$  pasos para leer las aristas en cualquier caso. Si el grafo es ralo, la inicialización de la matriz puede ser el factor dominante en el tiempo de ejecución del algoritmo, por lo que se debe buscar otro tipo de representación de las adyacencias.

La operación de fusión agrega aristas al grafo, ligando los vértices que corresponden a las regiones que: (i) tienen modas asociadas que están localizadas dentro de una distancia de separación de  $h_i/2$ , y (ii) satisfacen una condición de fuerza de frontera débil, como se menciona en [15]. De esta manera, se fusionan las regiones adyacentes cuyas modas asociadas, en el dominio del rango, están localizadas dentro de una distancia predefinida unas de otras. La fusión de las regiones se realiza, siempre y cuando se cumpla la condición de que la distancia, medida en modas entre dos regiones adyacentes, no sea mayor de 0.5. La distancia entre modas, o dicho de otra manera, la diferencia entre modas, se normaliza con respecto al radio del rango  $h_i$ , por lo que la condición de fusión se traduce a que la distancia entre modas no sea mayor a  $h_i/2$ . Si se cumple la condición anterior y la fuerza del borde entre regiones es menor que un umbral,  $\xi$ , esto es, tienen un borde débil entre ellas, indicando que realmente no hay una frontera entre las regiones, entonces las regiones se fusionan en una sola. La fuerza del borde se calcula promediando los valores de los píxeles del mapa de confianza a lo largo de la frontera que separa las regiones adyacentes.

## 4.2. Algoritmo encontrar y unir

Al realizarse la fusión de regiones sobre la MAR, el número de ellas disminuye por lo que se debe aplicar un procedimiento para reducir el grafo, esto se logra aplicando un algoritmo de *encontrar y unir* (union-find). En algunas aplicaciones se desea simplemente conocer si un vértice  $a$  está o no conectado a un vértice  $b$  en un grafo; el camino que los conecta puede no ser relevante. Los algoritmos que se han desarrollado para este problema son interesantes debido a que también pueden utilizarse para el procesamiento de conjuntos.

Los grafos se corresponden de manera natural con los conjuntos de objetos: los vértices corresponden a los objetos y las aristas significan “esta en el mismo conjunto que”. Otro término para definir estos conjuntos es de *clases de equivalencia*. Cada compo-

nente conexa corresponde a una clase de equivalencia diferente. El añadir una arista corresponde con la combinación de clases de equivalencia representadas por los vértices a conectar. El interés se centra en la pregunta “¿ $a$  es equivalente a  $b$ ?” o “¿ $a$  está en el mismo conjunto que  $b$ ?”. Esto corresponde claramente con la pregunta de los grafos “¿el vértice  $a$  está conectado al vértice  $b$ ?”

Para responder a la pregunta anterior se consideran métodos que sean dinámicos, es decir, pueden aceptar nuevas aristas mezcladas arbitrariamente con preguntas y contestar correctamente a las preguntas utilizando la información recibida. Por correspondencia con el problema de los conjuntos, la adición de una nueva arista se denomina operación *unir*, y las preguntas se denominan operaciones *encontrar* [54, p. 441]. El objetivo es tener una función que pueda verificar si dos vértices  $a$  y  $b$  pertenecen al mismo conjunto (o en representación de grafos, a la misma componente conexa) y, en caso de que así sea, que pueda unirlos en el mismo conjunto.

En lugar de construir una representación del grafo, es más eficaz utilizar una estructura orientada específicamente a la realización de las operaciones *encontrar* y *unir*. Esta estructura es un bosque de árboles, uno para cada componente conexa, se asume que cada componente conexa se identifica mediante un representante canónico. Se necesita poder encontrar si dos vértices pertenecen al mismo árbol y combinar dos árboles en uno. Las operaciones de *unir* y *encontrar* se implementan fácilmente usando la representación de “enlace padre” para los árboles, debido a que no se necesita recorrer el árbol hacia abajo, sino únicamente hacia arriba. En tales casos, sólo se necesita un enlace para cada vértice; el que lo conecta a su padre. Específicamente, se mantiene un arreglo *padre* que contiene, para cada vértice, el índice de su padre (0 si es la raíz de algún árbol). Se utiliza el procedimiento `encontrar(int a, int b, int union)` para implementar las dos operaciones de *unir* y *encontrar*. Si el tercer argumento es cero se tiene la operación *encontrar*, y si es distinto de cero, se tiene la operación *unir*. Para encontrar el padre de un vértice  $j$ , simplemente se pone  $j = \text{padre}(j)$ , y para encontrar la raíz del árbol a la que pertenece  $j$ , se repite esta operación hasta obtener cero. La implementación del procedimiento `encontrar` es [54, p. 444]:

```
int encontrar(int a, int b, int union)
{
    int i = a, j = b;
    while (padre(i) >0) i = padre(i);
    while (padre(j) >0) j = padre(j);
    if ((union != 0) && (i !=j)) padre(j) = i;
    return (i != j);
}
```

}

La función `encontrar` regresa cero si los dos vértices dados están en la misma componente. Si no es así y la bandera `union` está activada, se colocarán dichos vértices en la misma componente. El método es simple, se utiliza el arreglo `padre` para obtener la raíz del árbol que contiene cada vértice, y a continuación se checa si las raíces son las mismas. Para fusionar el árbol de raíz  $j$  con el de raíz  $i$ , simplemente se pone `padre(j) = i`. Por ejemplo, para probar si un vértice llamado  $a$  está en la misma componente que un vértice llamado  $b$  (sin introducir una arista entre ellos) se llama a la función `encontrar(a, b, 0)`.

El algoritmo descrito anteriormente tiene un mal comportamiento en el peor caso porque los árboles construidos pueden estar degenerados. Para resolver este problema se han sugerido varios métodos. Un método natural es intentar hacer lo más “razonable” cuando se fusionan dos árboles, en lugar de poner arbitrariamente `padre(j) = i`. Cuando se va a fusionar un árbol de raíz  $i$  con otro de raíz  $j$ , uno de los nodos debe permanecer como raíz y el otro (y todos sus descendientes) debe bajar un nivel en el árbol. Para minimizar la distancia entre la raíz y la mayoría de los nodos, parece lógico elegir como raíz el nodo que tenga el mayor número de descendientes. Esta idea, llamada *equilibrado de peso* [54, p. 446], se implementa fácilmente manteniendo el tamaño de cada árbol (número de descendientes de la raíz) en la entrada del arreglo `padre`, para cada nodo raíz, codificando con números negativos a todos los nodos raíz que puedan detectarse ascendiendo por el árbol en `encontrar`.

Idealmente, sería deseable que cada nodo apunte directamente a la raíz de su árbol. Sin embargo, para lograr este ideal habría que examinar al menos todos los nodos de uno de los dos árboles a fusionar, y esto podría representar una cantidad muy grande comparada con los relativamente pocos nodos situados en el camino a la raíz que suele examinar `encontrar`. Pero se puede hacer una aproximación al ideal haciendo que todos los nodos que se examinan apunten hacia la raíz. Este método, llamado *compresión de caminos*, es fácil de implementar haciendo otra pasada a través de cada árbol, después de haber encontrado la raíz, y fijando la entrada `padre` de cada vértice que se ha encontrado a lo largo del camino que apunta a la raíz.

La combinación de los métodos de equilibrado de peso y compresión de caminos asegura que los algoritmos se ejecuten más rápidamente. La siguiente implementación muestra que el código extra involucrado es un pequeño precio a pagar para prevenirse de los casos degenerados [54, p. 446].

```
int encontrar(int a, int b, int union)
{
```

```

int t, i = a, j = b;
while (padre(i) >0) i = padre(i);
while (padre(j) >0) j = padre(j);
while (padre(a) >0)
{ t = a; a = padre(a); padre(t) = i;}
while (padre(b) >0)
{t = b; b = padre(b); padre(t) = j;}
if ((union != 0) && (i !=j))
if (padre(j) <padre(i))
{padre(j) += padre(i) - 1; padre(i) = j;}
else
{padre(i) += padre(j) - 1; padre(j) = i;}
return (i != j);
}

```

Una vez obtenido el grafo reducido, se reetiqueta nuevamente la imagen, y se recalculan las modas de las regiones nuevas con el promedio de las modas de todas las regiones que se fusionaron, para dar origen a la nueva región. La operaciones de fusión de regiones y reducción del grafo se aplican iterativamente hasta que el número de regiones, entre una iteración y otra, permanece sin cambios.

### 4.3. Podado

Cuando se obtiene la fusión final de las regiones de la imagen, se aplica un proceso de poda para remover todas las regiones cuya área, establecida como un número de pixeles, sea inferior a un umbral,  $\mu$ , definido por el usuario. Para hacer esta remoción, se recorre la MAR uniendo regiones, cuya área es menor que el umbral, con su región *candidata* respectiva. Una región *candidata* es una región que tiene las siguientes propiedades: Es adyacente a la región que se esta podando y la distancia, medida como una diferencia de modas, a la región que se esta podando es un mínimo, o es la única región adyacente que tiene un área mayor que el umbral.

De la misma manera que en el grafo reducido, con la imagen reducida por el proceso de podado, se recalculan las modas de cada región, se asignan etiquetas a las nuevas regiones, y el proceso de podado se repite iterativamente hasta que el número de regiones ya no cambie o se alcance un número dado de iteraciones. Los información de la implementación acerca de la fusión y podado de regiones de la imagen se muestra en el Apéndice D.

# Capítulo 5

## Incorporación de información a priori

La aplicación de las técnicas mencionadas hasta este punto, producen regiones homogéneas en intensidad, reduciendo tanto como es posible el número de clases que conforman las diferentes estructuras neuroanatómicas del cerebro, no obstante, el número de ellas puede ser elevado, si se tiene en cuenta que las estructuras que se desean identificar son: Materia Gris (MG), Materia Blanca (MB) y Líquido Cefalorraquídeo (LCR). Para obtener el número de regiones deseado de la imagen segmentada, se propone el uso de conocimiento *a priori*, contenido en mapas de probabilidad para cada una de las clases de interés. Antes de utilizar los mapas de probabilidad, los datos de la imagen y la información *a priori* deben estar en el mismo espacio, esto requiere de una etapa de registro previo entre ellos, etapa realizada con el programa SPM99 [3], para posteriormente aplicar una regla de clasificación y lograr la segmentación de la imagen en el número de clases deseado.

### 5.1. Normalización espacial

Probablemente la forma más simple de registro es la corrección de movimiento de series de imágenes. Involucra encontrar la mejor transformación de cuerpo rígido de seis parámetros para minimizar la diferencia entre dos imágenes del mismo sujeto. Sin embargo, cuando se necesita mapear la imagen del cerebro de un sujeto hacia un espacio estándar, esto nos lleva al problema de la normalización espacial. La solución de este problema permite ampliar el análisis basado en voxel y facilita la comparación de diferentes sujetos y bases de datos. Para realizar la normalización espacial se utilizó el programa Statistical Parametric Mapping (SPM99)[3], debido a que contiene diferentes



plantillas del cerebro que permiten normalizar los datos de las imágenes hacia el espacio de Talairach. La entrada al SPM es el estudio de IRM que se desea normalizar, este estudio debe ser de la misma modalidad que la plantilla seleccionada en el programa. La salida que se obtiene es el estudio normalizado y los parámetros de la normalización. A continuación se describe la metodología que aplica el SPM para normalizar espacialmente las imágenes de cerebro de RM, la descripción es una combinación resumida de [7, 5, 6, 2].

La *Normalización espacial afín* describe el primer paso involucrado en registrar imágenes de sujetos diferentes dentro del mismo sistema coordinado. Seguida por la *Normalización espacial no lineal* para la corrección de diferencias totales en la forma de la cabeza que no pueden ser tomadas en cuenta por la normalización afín.

La transformación espacial que mapea cada voxel en la imagen, a su localización equivalente en las imágenes de probabilidad *a priori*, se logra registrando con imágenes plantilla en el mismo espacio estereotáctico (de Talairach y Tournoux) que las imágenes de probabilidad, las cuales fueron proporcionadas por el Montreal Neurological Institute [29, 30]. Las imágenes de probabilidad representan la probabilidad de que un voxel sea MG, MB o LCR después de que una imagen se ha normalizado al mismo espacio [4].

### 5.1.1. Normalización afín

La normalización espacial determina la transformación afín óptima de 12 parámetros estimados con una aproximación bayesiana *máximum a posteriori* (MAP). La aproximación de registro transforma espacialmente una imagen tal que iguala a otra (llamada la imagen plantilla dentro del SPM99) optimizando los parámetros que describen la transformación espacial. Estos parámetros óptimos se determinan minimizando la suma de las diferencias cuadradas entre la imagen plantilla y la imagen real, aplicando el algoritmo descrito en [32] que es similar al Gauss-Newton y se describe a continuación.

Suponga que  $e_i(\mathbf{p})$  es la función que describe la diferencia entre las imágenes real y plantilla en el voxel  $i$ , cuando el vector de parámetros del modelo tiene valores  $\mathbf{p}$ . Para cada voxel ( $i$ ), se puede usar una primera aproximación del teorema de Taylor para estimar el valor que esta diferencia tomará si los parámetros  $\mathbf{p}$  se incrementan por  $\mathbf{q}$  [5, 6]:

$$e_i(\mathbf{p} + \mathbf{q}) = e_i(\mathbf{p}) + q_1 \frac{\partial e_i(\mathbf{p})}{\partial p_1} + q_2 \frac{\partial e_i(\mathbf{p})}{\partial p_2} \dots \quad (5.1)$$

De esto, se puede establecer un conjunto de ecuaciones simultáneas (de la forma  $\mathbf{U}\mathbf{x} \simeq \mathbf{v}$ ) para estimar los valores que  $\mathbf{q}$  debe tomar para minimizar  $\sum_i e_i(\mathbf{p} + \mathbf{q})^2$ :

$$\begin{pmatrix} -\frac{\partial e_1(\mathbf{p})}{\partial p_1} & -\frac{\partial e_1(\mathbf{p})}{\partial p_2} & \dots \\ -\frac{\partial e_2(\mathbf{p})}{\partial p_1} & -\frac{\partial e_2(\mathbf{p})}{\partial p_2} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \end{pmatrix} \cong \begin{pmatrix} e_i(\mathbf{p}) \\ e_i(\mathbf{p}) \\ \vdots \end{pmatrix} \quad (5.2)$$

De este sistema de ecuaciones, podemos derivar un esquema iterativo para mejorar los parámetros estimados. Para la iteración  $n$ , los parámetros  $\mathbf{p}$  se actualizan como:

$$\mathbf{p}^{n+1} = \mathbf{p}^n + (\mathbf{U}^\top \mathbf{U})^{-1} \mathbf{U}^\top \mathbf{v} \quad (5.3)$$

donde  $\mathbf{v}$  es la parte derecha de (5.2), correspondiendo al residual de las diferencias, y  $\mathbf{U}$  es el primer factor de la parte izquierda, siendo el jacobiano de la derivada del residual  $v_i$  con respecto al parámetro  $p_j$ .

Para ajustar la imagen real ( $F$ ) a la imagen plantilla ( $G$ ) se usa una transformación afín de 12 parámetros (de  $p_1$  a  $p_{12}$ ), y dado que las imágenes pueden estar escaladas de manera diferente se incluye un parámetro adicional de escalamiento de intensidad ( $p_{13}$ ). El mapeo de transformación afín, a través de la matriz  $\mathbf{M}$  conteniendo los parámetros ( $p_1$  a  $p_{12}$ ), de la posición  $\mathbf{x}$  en una imagen a la posición  $\mathbf{y}$  en la otra está definido por  $\mathbf{y} = \mathbf{M}\mathbf{x}$  [7]:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \quad (5.4)$$

donde los elementos  $m_{ij}$  de la matriz  $\mathbf{M}$  son una función de los doce parámetros ( $p_1$  a  $p_{12}$ ). Los parámetros se optimizan usando la ecuación (5.3), y la función a minimizar [2, 7, 5] es:

$$\sum_i (F(\mathbf{M}x_i) - p_{13}G(x_i))^2 \quad (5.5)$$

El vector  $\mathbf{v}$  está dado por  $v_i = F(\mathbf{M}x_i) - p_{13}G(x_i)$ . La matriz  $\mathbf{U}$  se construye de las derivadas negativas, y se calculan de la siguiente manera. La razón de cambio de  $v_i$  con respecto al parámetro de escalamiento ( $p_{13}$ ) es simplemente  $-G(x_i)$  (la intensidad negativa de la imagen  $G$  en  $x_i$ ). Las derivadas de  $v_i$  con respecto a los parámetros de la transformación espacial ( $p_1$  a  $p_{12}$ ) se obtienen diferenciando  $F(\mathbf{M}x_i) - p_{13}G(x_i)$  con respecto a cada parámetro  $p_j$ . Las derivadas con respecto a los parámetros de translación son simplemente los gradientes de la imagen  $F$  en cada dirección ( $\partial f(\mathbf{y})/\partial y_i$ ), donde  $\mathbf{y} = \mathbf{M}\mathbf{x}$ .

Sin regularización, es posible introducir deformaciones innecesarias que sólo reducen la suma residual de cuadrados por una cantidad pequeña. Esto pudiera hacer inestable al algoritmo. La regularización se logra usando la regla de Bayes [6]:

$$p(u_p | v) = \frac{p(v | u_p)p(u_p)}{\int_j p(v | u_j)p(u_j)d_j} \quad (5.6)$$

Para nuestros propósitos,  $p(u_p)$  representa una distribución de probabilidad previa conocida de donde se obtienen los parámetros,  $p(v | u_p)$  es la verosimilitud de obtener los datos  $v$  dados los parámetros, y  $p(u_p | v)$  es la función a maximizar. La estimación *máxima a posteriori* (MAP) para los parámetros  $\mathbf{p}$  es la moda de  $p(u_p | v)$ . La optimización se puede simplificar suponiendo que todas las distribuciones de probabilidad son multidimensionales y normales (multi-normal), y puede por lo tanto describirse por un vector de medias y una matriz de covarianza.

Cuando se acerca al mínimo, la optimización casi se convierte en un problema lineal. Así, se puede suponer que los errores de los parámetros ajustados ( $\mathbf{p}$ ) se pueden aproximar localmente por una matriz de covarianza  $\mathbf{C}$ , y que los parámetros verdaderos se obtienen de una distribución multi-normal fundamental conocida de media ( $\mathbf{p}_0$ ) y covarianza ( $\mathbf{C}_0$ ) conocidas. Usando la función de densidad de probabilidad *a priori* de los parámetros, obtenemos una mejor estimación de los parámetros verdaderos tomando un promedio ponderado de  $\mathbf{p}_0$  y  $\mathbf{p}$  [7]:

$$\mathbf{p}_v = (\mathbf{C}_0^{-1} + \mathbf{C}^{-1})^{-1} (\mathbf{C}_0^{-1}\mathbf{p}_0 + \mathbf{C}^{-1}\mathbf{p})^{-1} \quad (5.7)$$

La matriz de covarianza estimada de los errores estándar para la solución MAP es entonces:

$$\mathbf{C}_v = (\mathbf{C}_0^{-1} + \mathbf{C}^{-1})^{-1} \quad (5.8)$$

$\mathbf{p}_v$  y  $\mathbf{C}_v$  son los parámetros que describen la distribución multi-normal  $p(u_p | v)$ .

Para emplear esta aproximación se requiere la estimación de  $\mathbf{C}$ , que es la matriz de covarianza estimada de los errores estándar de los parámetros ajustados; si las observaciones son independientes y cada una tiene desviación estándar unitaria, entonces  $\mathbf{C} = (\mathbf{U}^\top \mathbf{U})^{-1}$ . Las desviaciones estándar se desconocen, suponiendo que son iguales para todas las observaciones, su valor se estima de la suma de diferencias cuadradas [7, 5]:

$$\sigma^2 = \sum_{i=1}^I (F(\mathbf{M}\mathbf{x}_i) - p_{13}G(\mathbf{x}_i))^2 \quad (5.9)$$

Esto da una matriz de covarianza  $(\mathbf{U}^\top \mathbf{U})^{-1} \sigma^2 / (I - J)$ , donde  $I$  se refiere al número de localidades muestreadas en las imágenes y  $J$  se refiere al número de parámetros (13 en este caso), así la matriz de covarianza se puede estimar por:

$$\mathbf{C} = (\mathbf{U}^\top \mathbf{U})^{-1} \sigma^2 / \nu \quad (5.10)$$

donde  $\nu$  se refiere a los grados de libertad [2]. Usualmente no se cumple que  $\nu = I - J$ , tal que se hace una estimación de los grados de libertad como se describe en [7].

La distribución *a priori* de los parámetros ( $\mathbf{p}_0$  y  $\mathbf{C}_0$ ) se determinó con transformaciones afines de 51 imágenes T1 de RM cerebral [7]. Combinando las ecuaciones (5.3), (5.7) y (5.10), se obtiene el siguiente esquema:

$$\mathbf{p}_v^{n+1} = (\mathbf{C}_0^{-1} + \alpha)^{-1} (\mathbf{C}_0^{-1} \mathbf{p}_0 + \alpha \mathbf{p}_v^n - \gamma)^{-1} \quad (5.11)$$

donde  $\alpha = \mathbf{U}^\top \mathbf{U} \nu / \sigma^2$  y  $\gamma = \mathbf{U}^\top \mathbf{v} \nu / \sigma^2$ .

La solución óptima ya no es aquella que minimiza la suma de cuadrados de los residuales, ya que el objetivo de la optimización es obtener un ajuste con los errores más pequeños, descritos por la matriz de covarianza de los parámetros estimados,  $(\mathbf{C}_0^{-1} + \alpha)^{-1}$ . La calidad del ajuste está reflejada en el determinante de esta matriz, tal que la solución óptima deberá alcanzarse cuando el determinante se minimice.

### 5.1.2. Normalización espacial no lineal

Con la normalización no-lineal se determinan los parámetros que describen las diferencias de forma global entre la imagen y la plantilla, que no fueron tomadas en cuenta por la normalización afín. El objetivo de la normalización espacial no lineal es distorsionar (warp) las imágenes tal que las regiones homólogas de diferentes cerebros se muevan tan cerca unas de otras como sea posible. Para abarcar el rango de posibles distorsiones no lineales se requiere de un elevado número de parámetros. Algunas aproximaciones reducen el número de parámetros que modelan la distorsión no lineal usando deformaciones que consisten de una combinación lineal de funciones base, tal que, la transformación desde las coordenadas  $\mathbf{x}$ , a las coordenadas  $\mathbf{y}$  es:

$$y_d = x_d + \sum_j t_{jd} b_{jd}(\mathbf{x}) \quad (5.12)$$

donde  $t_{jd}$  es el  $j$ -ésimo coeficiente para la dimensión  $d$ , y  $b_{jd}(\mathbf{x})$  es la  $j$ -ésima función base en la posición  $\mathbf{x}$  para la dimensión  $d$ . Las funciones base usadas son la transformada coseno discreta (TCD) de tres dimensiones [2, 5].

Los errores asociados con el ajuste de la normalización no lineal, probablemente serán muy grandes mientras mayor sea el número de parámetros, así el uso de funciones base reduce los errores. Otra posibilidad para reducir los errores es el uso de restricciones, tal como conservar un mapeo uno a uno entre la transformación de la imagen y la plantilla, pero aún con esto, los errores siguen siendo de consideración. Una aproximación bayesiana similar a la ya descrita podría intentar alcanzar un compromiso óptimo entre estas dos posibilidades.

De nuevo, la optimización involucra minimizar la suma de diferencias cuadradas entre la imagen real ( $F$ ) y una imagen plantilla ( $G$ ). Las imágenes pueden estar escaladas de manera diferente, tal que es necesario un parámetro adicional ( $o$ ) para acomodar esta diferencia [5, 6]. La función a optimizar es:

$$\sum_i (F(\mathbf{y}_i) - oG(\mathbf{x}_i))^2 \quad (5.13)$$

Ya que la expresión a minimizar es de la misma forma que la anterior (5.5), la aproximación descrita en la sección anterior se usa para optimizar los parámetros  $t_1$ ,  $t_2$ ,  $t_3$  y  $o$ . Esto requiere las derivadas de la función  $F(\mathbf{y}_i) - oG(\mathbf{x}_i)$  con respecto a cada parámetro, lo cual se puede obtener aplicando la regla de la cadena:

$$\frac{dF(\mathbf{y})}{dt_{jd}} = \frac{dF(\mathbf{y})}{dy_d} \frac{dy_d}{dt_{jd}} \quad (5.14)$$

En esta expresión,  $dF(\mathbf{y})/dy_d$  es simplemente la derivada en la dimensión  $d$  de la imagen  $F$ , y  $dy_d/dt_{jd}$  simplemente evalúa a  $b_{jd}(\mathbf{x})$ .

## 5.2. Clasificación

En el contexto de la segmentación de imágenes, la clasificación es un proceso que calcula la probabilidad de cada uno de los puntos de pertenecer a ciertas clases, alternativamente se puede operar sobre grupos de puntos en lugar de tomar puntos individuales. El proceso de clasificación es un esquema guiado por punto (pixel o voxel), y en particular usaremos la forma alterna de operar sobre grupos de puntos, esto es regiones de la imagen. El problema de clasificación está especificado por: un conjunto de regiones  $\mathbf{R}_n$  y un conjunto finito de clases  $\Omega$  para cada región  $\mathbf{R}_n$ . El proceso es como sigue: determinar para cada región su membresía de clase más probable, determinar las regiones semilla de la imagen y, finalmente, calcular una métrica para clasificar cada región.

### 5.2.1. Probabilidad de pertenencia

La elección de propiedades o, equivalentemente, la definición de una clase, es el tema fundamental del problema de clasificación. En la aproximación no paramétrica, las muestras se separan de acuerdo a la función de densidad, considerada como la frontera natural que separa las modas de las distribuciones. Para obtener la clasificación de las principales estructuras neuroanatómicas se hace uso del conocimiento *a priori*, contenido en mapas de probabilidad para cada una de las clases de interés (MG, MB y LCR) de la imagen. Una vez que los mapas de probabilidad, se han transformado al mismo espacio de los datos, la clasificación de las imágenes de RM cerebrales, requiere esencialmente calcular las probabilidades de pertenencia a un tipo de tejido cerebral para cada una de las regiones encontradas durante el procesamiento de los datos, y asignar, de acuerdo a una regla de clasificación, las regiones a la clase de tejido con mayor probabilidad.

La etapa de filtrado del proceso de segmentación, da como resultado una partición de la imagen en regiones de intensidad homogénea, sin embargo, dichas regiones aun no están asociadas con alguna de las clases de tejido en que se debe clasificar la imagen. La clasificación de las regiones se realiza utilizando el teorema de Bayes, interpretando cada uno de los términos que lo conforman en el contexto del problema que nos ocupa. El cambio más relevante en la aplicación del teorema de Bayes, consiste en permitir que el proceso de clasificación sea guiado por las regiones de la imagen y no por los voxeles, es decir, cada región de la imagen, en su conjunto, deberá ser asignada a un tipo de tejido cerebral particular. El cálculo de los términos del teorema, en el contexto del problema de clasificación de regiones de la imagen, es el siguiente.

Dada una tarea de clasificación de  $M$  clases,  $w_1, \dots, w_M$ , y  $\mathbf{R}_n$  regiones por asignar a alguna de las clases, cada una constituida por voxeles ( $\mathbf{v}$ ) de la misma intensidad, el cálculo de la probabilidad para una región está basado en la determinación de la probabilidad de que la región pertenezca a la clase  $w_i$ , de este modo, las  $M$  probabilidades condicionales de clasificación de la región son:  $P(w_i | \mathbf{v}, \mathbf{v} \in \mathbf{R}_n) \equiv P(w_i | \mathbf{R}_n)$ ,  $i = 1, 2, \dots, M$ . Para calcular  $P(w_i | \mathbf{R}_n)$  se usa la información que se tiene del conocimiento previo de las clases, expresado por medio de mapas probabilísticos del cerebro.

Las imágenes de probabilidad se derivaron de imágenes de RM de un gran número de sujetos. Las imágenes originales se segmentaron en imágenes binarias de MG, MB y LCR, y las imágenes de probabilidad son las medias de estas imágenes binarias, tal que contienen valores entre 0 y 1. Estas imágenes representan la probabilidad previa de que un voxel sea MG, MB o LCR [5]. El cálculo de la probabilidad a priori de la clase  $P(w_i)$  está basado en el conocimiento de las imágenes de probabilidad ( $B$ ), y se calcula

de la siguiente manera

$$P(w_i) = \frac{\sum_{\mathbf{v} \in B_i} B_i(\mathbf{v})}{\sum_{k=1}^M \sum_{\mathbf{v} \in B_k} B_k(\mathbf{v})} \quad (5.15)$$

donde  $B_i(\mathbf{v})$  es el valor de probabilidad en el voxel  $\mathbf{v}$ ,  $\sum B_i$  es la integral sobre la  $i$ -ésima imagen de probabilidad y la clase  $w_i$  corresponde a la imagen de probabilidad  $B_i$ , tal que  $P(w_i)$  es la proporción de cada mapa probabilístico al total de los mapas.

Para cuantificar la información que aportan cada una de las regiones ( $\mathbf{R}_n$ ) de la imagen, se calcula su verosimilitud  $p(\mathbf{v}, \mathbf{v} \in \mathbf{R}_n | w_i) \equiv p(\mathbf{R}_n | w_i)$  como la proporción de la región en cada uno de los mapas probabilísticos

$$p(\mathbf{R}_n | w_i) = \frac{\sum_{\mathbf{v} \in \mathbf{R}_n} B_i(\mathbf{v})}{\sum_{k=1}^M \sum_{\mathbf{v} \in \mathbf{R}_n} B_k(\mathbf{v})} \quad (5.16)$$

donde ahora  $\sum B_i$  es la integral sobre la región  $\mathbf{R}_n$  para la clase  $w_i$ , calculándose así la verosimilitud de  $w_i$  con respecto a  $\mathbf{R}_n$  para indicar la clase más probable y poder identificar la clase verdadera a la que pertenece la región. Una vez calculadas la probabilidad previa de la clase y la verosimilitud, se sustituyen en el teorema de Bayes para determinar la probabilidad de que la región pertenezca a una clase particular

$$P(w_i | \mathbf{R}_n) = \frac{p(\mathbf{R}_n | w_i)P(w_i)}{\sum_{k=1}^M p(\mathbf{R}_n | w_k)P(w_k)} \quad (5.17)$$

En una tarea de clasificación, la regla de decisión de Bayes para una región  $\mathbf{R}_n$ , es asignarla a la clase  $w_i$  si

$$P(w_i | \mathbf{R}_n) > P(w_k | \mathbf{R}_n) \quad \forall i \neq k \quad (5.18)$$

El teorema de Bayes permite calcular la probabilidad *a posteriori*, es decir, qué tan probables son los diferentes valores posibles de  $w_i$  para que dada una región de la imagen pertenezca a la clase  $w_i$ . Como sugiere esta ecuación, podemos usar la información proporcionada por los mapas probabilísticos *a priori* para determinar, tanto las probabilidades condicionales de clase o verosimilitudes, y las probabilidades *a priori* o prevalencias, necesarias para calcular  $P(w_i | \mathbf{R}_n)$ . Así, para cada una de las regiones homogéneas, encontradas en la etapa de fusión de regiones de la imagen, se calcula su probabilidad de pertenencia, y se clasifican de acuerdo a la regla de decisión de Bayes.

### 5.2.2. Etiquetado de las regiones

Como ya se dijo, la segmentación de imagen se refiere a la descomposición de una imagen en regiones, pero estrictamente hablando no clasifica los tipos de tejido. La clasificación es el paso que trata con el etiquetado de las regiones. Para resolver el problema del etiquetado de las regiones de la imagen se debe etiquetar cada voxel como perteneciendo a cada una de las clases de interés que se desean encontrar en el cerebro.

La meta del etiquetado de una imagen es asignar una etiqueta, es decir, un significado a cada región de la imagen para lograr una interpretación apropiada de la misma. Así, el problema del etiquetado de la imagen está especificado por el conjunto de regiones  $\mathbf{R}_n$  detectadas, que deben ser asignadas a una de  $M$  posibles categorías semánticas del conjunto  $\Omega = \{\omega_1, \dots, \omega_M\}$  de etiquetas. Idealmente, quisiéramos asignar a cada región una etiqueta, pero esto en general no es posible, y así tendremos, para cada región un conjunto de probabilidades de cada etiqueta:

$$\{P(\mathbf{R}_n \in \omega_1), P(\mathbf{R}_n \in \omega_2), \dots, P(\mathbf{R}_n \in \omega_M)\} \quad (5.19)$$

$$\sum_{i=1}^M P(\mathbf{R}_n \in \omega_i) = 1 \quad (5.20)$$

donde  $P(\mathbf{R}_n \in \omega_i)$  denota la probabilidad de que la etiqueta  $\omega_i$  sea la correcta asignación para la región  $\mathbf{R}_n$ .

El proceso de segmentación de imágenes cerebrales descrito, produce regiones homogéneas en intensidad sin que sea posible, hasta este punto del proceso, clasificar cada región en algún tipo de tejido cerebral. Para llevar a cabo la clasificación es necesario especificar, tanto el número y tipo de cada clase que se tiene como propósito extraer de las imágenes. La especificación del número y tipo de cada clase pudiera hacerse de manera automática, por ejemplo, buscando en la imagen las regiones que sean representativas de los tejidos a delinear, y considerarlas como semilla de la clase que representan. Debido a que el interés del proceso de segmentación es la identificación de LCR, MG, MB y una clase, a la que se denomina como fondo, conteniendo cualquier otro tipo de tejido diferente a los mencionados; con ésto queda especificado el número y tipo de clases a extraer de la imagen. Por lo tanto, el problema del etiquetado de las regiones de la imagen consiste en asignar cada  $\mathbf{R}_n$ , a uno de cuatro ( $M = 4$ ) posibles tipos de tejido del conjunto  $\Omega = \{\omega_1, \dots, \omega_4\}$ , donde  $\{\omega_1, \dots, \omega_4\} = \{\text{fondo, LCR, MG, MB}\}$ .

La selección de la región semilla para cada clase, se realiza buscando la región más grande de cada clase de entre todas las regiones, clasificadas previamente con el cálculo



de la probabilidad de pertenencia para cada tipo de tejido. De esta manera, las cuatro semillas obtenidas están caracterizadas por la moda de intensidad encontrada por el algoritmo del CM. La asignación final del etiquetado (tipo de tejido) de las regiones se hace calculando una medida de distancia, que establece que tan lejana está la intensidad de la región bajo análisis, de la intensidad de la región semilla.

Una medida simple de distancia es la diferencia absoluta entre la moda de intensidad de la región ( $IR$ ) y la moda de intensidad que caracteriza la región semilla ( $IS$ ). Para cada región  $\mathbf{R}_n$ , su distancia a la etiqueta  $\omega_i$  de la región semilla es:

$$d(\mathbf{R}_n \in \omega_i) = |IR_n - IS_i| \quad (5.21)$$

Esto entonces necesita ser cambiado a una probabilidad, lo cual puede hacerse de la siguiente manera:

$$P(\mathbf{R}_n \in \omega_i) = 1 - \frac{d(\mathbf{R}_n \in \omega_i)}{\sum_{k=1}^4 d(\mathbf{R}_n \in \omega_k)} \quad (5.22)$$

el cálculo de esta probabilidad permite que a  $\mathbf{R}_n$  se le asigne la etiqueta que maximiza su probabilidad, es decir,  $P(\mathbf{R}_n \in \omega_i) = \max_k P(\mathbf{R}_n \in \omega_k)$ .

Los mapas probabilísticos con la información a priori, de la misma manera que en [47, 44], fueron tomados del atlas digital distribuido con el software SPM99. Debido a que los mapas de probabilidad sólo están disponibles para MG, MB y LCR, se estimó un cuarto mapa para fondo y tejido no cerebral, restando la suma de los tres mapas de un mapa con todos sus elementos puestos igual al máximo de la suma; finalmente, los cuatro mapas se normalizan para que la suma de probabilidades del mismo pixel en todos los mapas sea uno. El Apéndice E describe la manera de calcular la probabilidad de cada región, para luego determinar las regiones semilla y finalmente, realizar la clasificación de todas las regiones.

### 5.3. Resumen

Una vez establecidos los fundamentos teóricos que se utilizan en la aproximación de segmentación propuesta, a continuación se mencionan las etapas principales que se siguen en la propuesta:

1. Calcular el mapa de confianza de bordes, con el propósito de medir la confianza para la presencia de un borde en los datos del volumen a segmentar. Los pasos para obtener el mapa de confianza son:

- a) Determinar los polinomios de Krawtchouk para estimar las derivadas por medio de las máscaras de diferenciación  $\mathbf{W}$ .
- b) Calcular el gradiente de la imagen utilizando las máscaras de diferenciación para ambas coordenadas X e Y.
- c) Determinar la fuerza de borde, expresada como la magnitud del gradiente de la imagen.
- d) Determinar los percentiles,  $\rho$ , de la distribución acumulativa de los valores de la fuerza de borde normalizados.
- e) Calcular la medida de confianza,  $\eta$ , para cada dato con base en la estimación del ángulo del borde y la plantilla de borde ideal respectiva.
- f) Calcular el mapa de confianza de bordes,  $\varphi$ , para cada dato como una combinación lineal de los percentiles ( $\rho$ ) y la medida de confianza ( $\eta$ ), afectado por un factor de mezcla ( $\beta$ ) que controla la cantidad de información del percentil (gradiente) y de la medida de confianza que contiene el mapa de confianza. El mapa de confianza no sólo depende del factor de mezcla, también depende del umbral de borde ( $\delta$ ). La elección de ambos parámetros se hizo procesando una sola imagen sintética de cerebro, variando entre 0 y 1 los parámetros, determinando la probabilidad de pertenencia de las regiones resultantes y evaluando, cualitativa y cuantitativamente, la calidad de la segmentación. Los resultados de variar los dos parámetros se observan en la Figura 6.6; la calidad de la segmentación se mejora si se permite que la medida de confianza tenga mayor peso en el mapa de confianza, pero como también influye el umbral de borde, el mejor índice de similitud se obtuvo con la combinación de factor de mezcla = 0.3 y umbral de borde = 0.4. Con estos valores, el mapa de confianza tiene 30 % de información del gradiente y 70 % de medida de confianza. De manera general se puede decir que la calidad de la segmentación se mejora, dando mayor peso a la confianza del borde que al gradiente.

2. Filtrado de la imagen por corrimiento de media, ponderado con el mapa de confianza, hasta alcanzar la convergencia a una moda, o el número de iteraciones dado para cada dato.

- a) Para el dato que se está procesando ( $\mathbf{x}$ ), determinar su vecindario ( $\mathbf{X}_i$ ) de análisis en función del ancho de ventana espacial ( $h_e$ ).

- b) Para cada dato del vecindario, determinar la distancia Euclidiana de separación con el dato en proceso, si la diferencia es menor que el ancho de ventana espacial, se realiza un análisis de intensidad.
  - c) En el análisis de intensidad se evalúa la diferencia en intensidad entre el dato y su vecino. Si la diferencia es menor o igual que el ancho de ventana de intensidad ( $h_i$ ), el vecino se pondera con el mapa de confianza ( $\varphi$ ) y se incluye en el cálculo del vector de corrimiento de media,  $M_h(\mathbf{x})$ .
  - d) Repetir los pasos (a), (b) y (c) de esta etapa, hasta que el dato ( $\mathbf{x}$ ) alcance su moda o bien, se complete el número máximo de iteraciones.
3. Análisis de adyacencia de regiones, fusionando y podando, hasta que dejen de cumplirse las condiciones para fusión y poda de regiones .
- a) Etiquetar la imagen filtrada asignando el mismo valor escalar a todos los pixeles que convergieron a la misma moda, formando una región homogénea en intensidad, y así para todas las regiones de la imagen filtrada.
  - b) Con la imagen etiquetada, crear una matriz de adyacencia de regiones.
  - c) Fusionar las regiones adyacentes que cumplan las siguientes condiciones: la diferencia en modas es menor de  $h_i/2$  y la medida de confianza del borde es menor a un umbral de fuerza de borde ( $\xi$ ).
  - d) Recalcular las modas de las nuevas regiones con el promedio de la intensidad de las regiones que les dieron origen, volver a etiquetar las regiones resultantes de la fusión.
  - e) Los pasos (b), (c) y (d) de esta etapa se repiten iterativamente hasta que ya no es posible fusionar más regiones.
  - f) Podar las regiones pequeñas de la imagen fusionada cuyo número de pixeles sea menor a un umbral preestablecido ( $\mu$ ), uniéndolas con la región adyacente de área mayor al umbral y con la cual tiene una diferencia de modas mínima. Recalcular las modas, reetiquetar la imagen podada y repetir el podado iterativamente, hasta que ya no existan regiones pequeñas.
4. Normalización espacial entre las imágenes y los mapas de probabilidad a priori.
- a) Determinar la matriz de transformación que lleva a las imágenes hacia un espacio estándar definido por una imagen plantilla que se aproxima al espacio de Talairach.

- b) Transformar los mapas de probabilidad hacia el espacio de los datos de la imagen, por medio de la matriz de transformación.
5. Clasificación de las regiones a algún tipo de tejido cerebral.
- a) Calcular el mapa que corresponde a la clase de fondo de la imagen a partir de los mapas de probabilidad para MG, MB y LCR.
  - b) Utilizando los mapas de probabilidad *a priori*, determinar la probabilidad de pertenencia de cada región de la imagen, etiquetando la región con la clase de mayor probabilidad.
  - c) Calcular la semilla de la región que caracteriza a cada una de las clases, considerando los criterios de tamaño máximo de la región con esa etiqueta y máxima probabilidad en la etiqueta.
  - d) Reetiquetar las regiones, asignando cada región a la etiqueta de la semilla más cercana, aplicando un criterio de mínima diferencia en modas de intensidad entre región y semilla.

# Capítulo 6

## Resultados y discusión

Para probar la técnica de segmentación propuesta se utilizaron dos tipos de datos: sintéticos y reales, obtenidos del Internet Brain Segmentation Repository (IBSR) [39] y del McConnell Brain Imaging Centre (BIC) [1]. Un tipo de datos sintéticos corresponde a 5 volúmenes de datos, simulando por medio de un modelo tridimensional muy sencillo, al cerebro; y otro tipo a un volumen de IRM simulando un cerebro normal. Los datos reales consisten de 20 volúmenes de cerebro normal y sus correspondientes segmentaciones manuales, los cuales fueron proporcionados por el Center for Morphometric Analysis del Massachusetts General Hospital y están disponibles en el IBSR.

### 6.1. Datos sintéticos

Este tipo de datos son imágenes sencillas utilizadas en la etapa inicial de desarrollo y prueba del método de segmentación. El modelo 3D de los datos sintéticos consiste en dos esferas concéntricas, donde la interior representa la materia blanca y la exterior a la materia gris; cada volumen de datos está contaminado con diferentes niveles de ruido y gradientes de intensidad, y formado por 52 rebanadas de  $256 \times 256$ , como puede observarse en la Figura 6.1, donde se muestra una ventana de datos de  $161 \times 161$  de una rebanada de cada volumen.

Los modelos sintéticos tienen diferente grado de complejidad para identificar las clases que contienen, dependiendo de la cantidad de ruido e inhomogeneidad con que están contaminados; así en la Figura 6.1, también se muestran los histogramas de intensidad correspondientes, donde se aprecia que la esfera 1 es el modelo más sencillo contaminado solo con ruido y no tiene gradiente, es decir, sus intensidades no están traslapadas. En la esfera 2, además del ruido, se observa un poco de traslape en las intensidades, aumentándose aun más este traslape en la esfera 3, hasta que el traslape

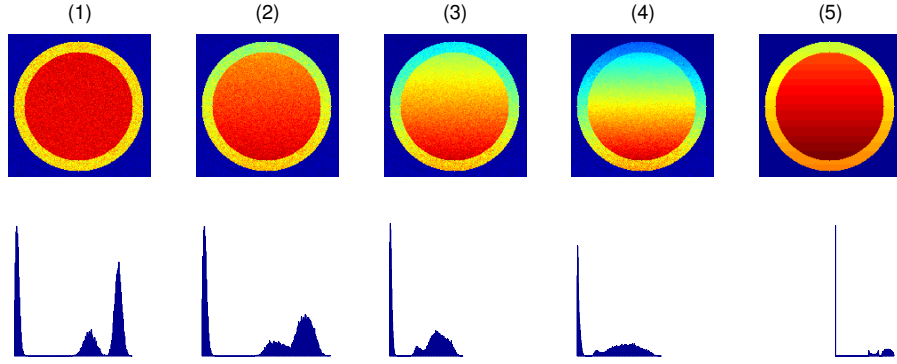


Figura 6.1: Volúmenes de datos sintéticos y sus histogramas de intensidad. La esfera 1 corresponde a datos con ruido y sin gradiente de intensidad. La esfera 2 está contaminada con ruido y un gradiente pequeño. La esfera 3 tiene ruido y un gradiente mayor. La esfera 4 tiene ruido y un gradiente aun mayor que la 3. La esfera 5 no tiene ruido y tiene un gradiente igual que en la esfera 2.

Tabla 6.1: Valores de los parámetros del proceso de segmentación para los tipos de datos segmentados.

Parámetro	Símbolo	Fantoma esférico	Cerebro Real y Sintético
Factor de mezcla	$\beta$	0.2	0.3
Umbral de borde	$\delta$	0.4	0.4
Fuerza de borde	$\xi$	0.7	0.8
Radio espacial	$h_e$	6	6
Radio de intensidad	$h_i$	25	9
Tamaño mínimo de región	$\mu$	30	30
Ventana del gradiente	$m$	2	2

es completo en la esfera 4. A mayor traslape de intensidades, mayor dificultad para estimar las funciones de densidad de los datos y para separar las clases presentes en la imagen.

Aun cuando en el proceso de segmentación intervienen varios parámetros, que deben ajustarse conforme a los datos, una vez determinados, permanecen constantes a lo largo de todo el proceso. La segmentación de los datos sintéticos se realizó en cada uno de los 5 volúmenes, sin variar ninguno de los parámetros del proceso, es decir, se mantuvieron fijos para los 5 grupos de datos. Los valores de cada uno de ellos se muestran en la Tabla 6.1, tanto para los datos sintéticos como para los reales.

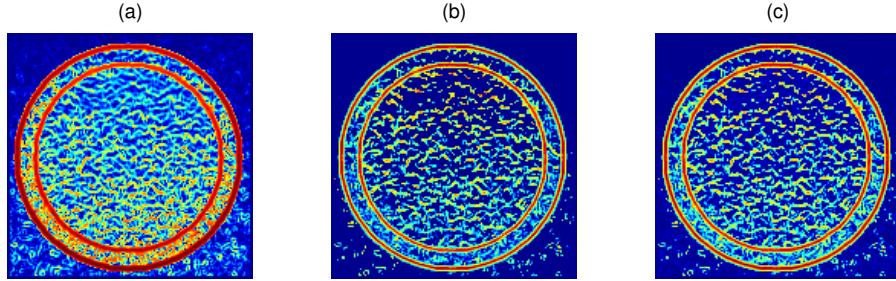


Figura 6.2: Determinación del mapa de confianza. (a) distribución acumulativa, (b) medida de confianza y (c) mapa de confianza de una rebanada de la esfera 4 con las intensidades traslapadas.

### 6.1.1. Mapa de confianza

El primer paso del proceso de segmentación es el cálculo del mapa de confianza, teniendo por objeto medir la confianza para la presencia de un borde en los datos del volumen a segmentar [41]. La detección de los bordes de la imagen se realiza estimando la magnitud y orientación del gradiente en una ventana de datos, cuyo tamaño está determinado por el parámetro  $m$  (ventana del gradiente), el valor de  $m = 2$  se mantuvo constante para todos los datos que se procesaron, dando como resultado un tamaño de ventana de  $5 \times 5$ . Con la magnitud del gradiente se determina la función de distribución acumulativa empírica, obteniéndose los rangos normalizados; y con la estimación de la orientación del gradiente se elige la plantilla del borde ideal para cada dato. Si el valor del rango normalizado es mayor que el umbral del borde ( $\rho_k > \delta$ ) para un dato dado, se le calcula la medida de confianza, de otra manera no se considera parte de un borde y su medida de confianza es cero. La medida de confianza del dato se calcula como el coeficiente de correlación entre el dato y su plantilla de borde ideal.

En la Figura 6.2 se muestra un ejemplo de la imagen de la distribución acumulativa y de la medida de confianza, así como el mapa de confianza que es el resultado de combinar las imágenes anteriores. La combinación de la distribución acumulativa y la medida de confianza, en el mapa de confianza, está en proporción con el parámetro del factor de mezcla ( $\beta$ ), es decir, si el factor de mezcla tiende a cero el mapa de confianza queda determinado por la medida de confianza exclusivamente, y si tiende a uno, por la función de distribución (i.e. por el gradiente). Así, en la Figura 6.2, el mapa de confianza tiene mayor semejanza con la medida de confianza debido a que  $\beta = 0,2$ , esto es, el mapa de confianza tiene un 80 % de la medida de confianza y un 20 % del gradiente.

### 6.1.2. Filtrado

El segundo paso consiste en filtrar la imagen original, conservando los bordes del mapa de confianza. El filtrado se realiza aplicando el algoritmo de CM a datos 3-dimensionales: dos dimensiones espaciales, y una de rango o intensidad. Como ya se mencionó, el CM define una trayectoria de los datos hacia sus modas de la función de densidad, corriendo cada dato, a lo largo de la trayectoria, con un paso de tamaño equivalente a su media muestral en cada iteración. Para calcular la media muestral se determina el vecindario del dato a analizar, considerando como parte del vecindario a todos aquellos pixeles que estén a una distancia Euclidiana  $\leq$  al radio espacial ( $h_e$ ); en seguida, para cada dato del vecindario, se evalúa su diferencia en intensidad con el dato que se está analizando, y si la diferencia está dentro del radio de resolución en intensidad ( $h_i$ ), entonces el vecino se toma en cuenta para el cálculo de la media.

Para preservar los bordes en el filtrado de la imagen, se utiliza el mapa de confianza como una ponderación en el cálculo de la media muestral, esto es, para que el dato tenga un peso relevante en la determinación de la media, no debe formar parte de un borde, como queda establecido al usar la ecuación 2.39 en la determinación del vector de corrimiento de media. Los resultados del filtrado de los datos sintéticos se muestran en la Figura 6.3, donde las imágenes Esfera 1 a Esfera 5 son una rebanada de cada uno de los modelos sintéticos, y se grafican en color falso para apreciar con mayor facilidad la variación en intensidad que existe dentro de cada clase; así se observa que la Esfera 4 es la que tiene mayor variación, siendo el modelo más complicado para la segmentación, como ya se había mencionado. En los mapas de confianza, Mapa 1 a 4, se observa con facilidad el contenido de ruido en las imágenes, sin embargo, los bordes que definen las fronteras entre las regiones de las clases están bien remarcados, esta característica también se presenta en el mapa de la Esfera 5, que no está contaminada con ruido y sólo contiene un poco de traslape en las intensidades. En las imágenes filtradas, tercer renglón en la Figura 6.3, se nota el resultado del filtrado por la ausencia de ruido en ellas, pero también se nota la excelente preservación de los bordes entre regiones.

El algoritmo de CM aplicado como etapa de filtrado, pareciera insensible al contenido de ruido en los datos porque las imágenes CM2 y CM5 se ven iguales, y sólo difieren, considerando los datos originales, en que la primera tiene ruido y la segunda no, conteniendo ambas un gradiente de intensidad similar. A simple vista, la imagen CM1 se observa como la mejor imagen segmentada, en el sentido de que al sustituir los datos por las modas a las cuales convergieron, el ruido desaparece y las regiones son más homogéneas que las de los datos originales. Los valores de intensidad son diferentes porque en CM1 los valores originales se sustituyeron por las modas encontradas por el



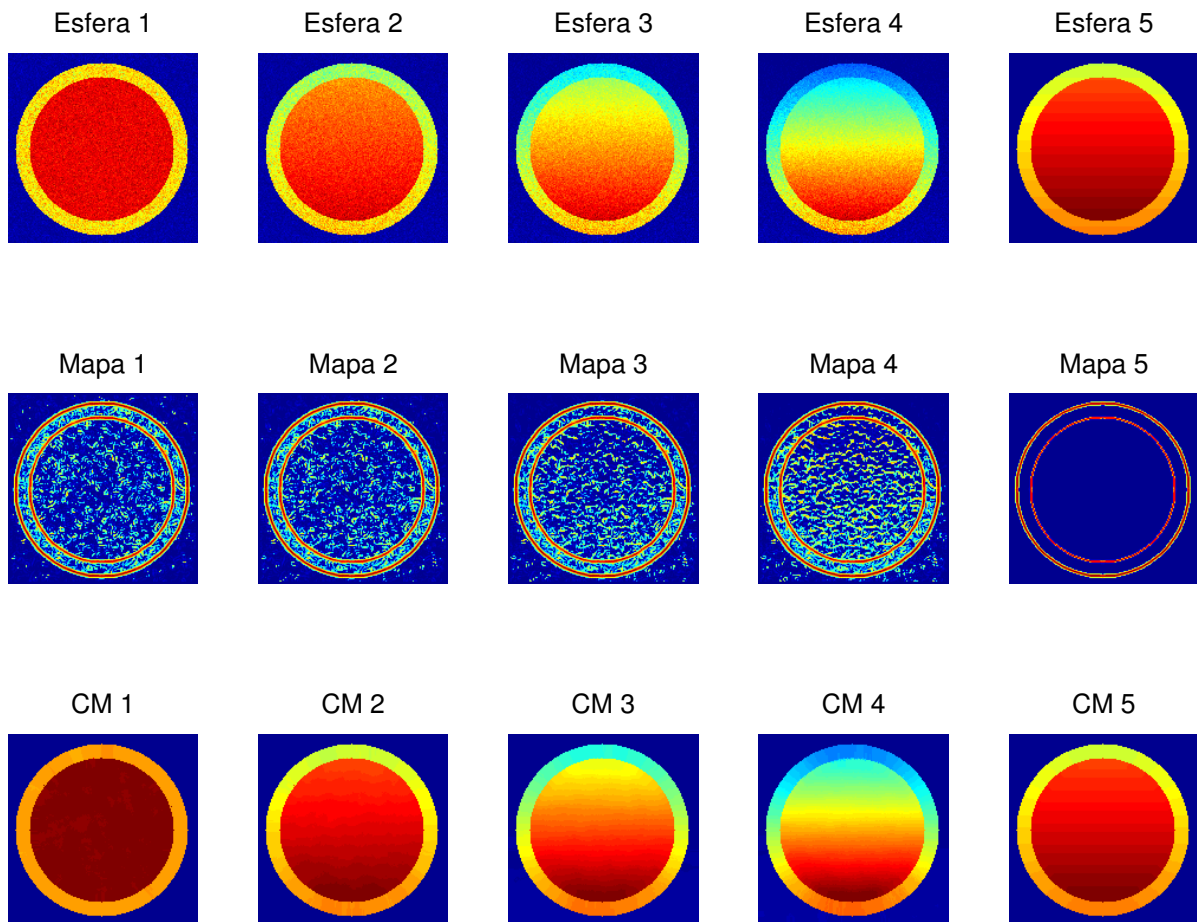


Figura 6.3: Filtrado de los datos sintéticos. Rebanadas de cada modelo sintético, donde el primer renglón representa los datos originales, el segundo los mapas de confianza correspondientes, y el último, los datos filtrados por el proceso de segmentación.

Tabla 6.2: Número de regiones presentes, en una rebanada de los datos sintéticos, en las etapas del análisis de adyacencias.

Esfera	Etiquetado	Fusión	Poda
1	354	10	3
2	251	12	3
3	393	16	3
4	1665	83	3
5	89	3	3

CM. En las imágenes filtradas CM3 y CM4 de la Figura 6.3, al parecer la ganancia, con respecto a sus datos originales, es sólo la disminución del ruido porque las regiones no son muy homogéneas en intensidad. Esta falta de homogeneidad pudiera considerarse como una desventaja del algoritmo, y lo que es aún peor como una contradicción, debido a que el objetivo de un método de segmentación por regiones es separar dichas regiones por un criterio de homogeneidad, y a simple vista no se observa suficiente homogeneidad en CM3 y CM4 que permita delimitar sus regiones. Recordando que la calidad o resolución del proceso de segmentación está altamente influenciada por los parámetros del ancho de banda del kernel en el estimador, es precisamente el ancho de banda espacial el que restringe el tamaño de las regiones homogéneas, realizando una búsqueda local y no global de la moda de la función de densidad. A mayor ancho de banda espacial mayor el área de las regiones homogéneas, lo que puede tener un impacto negativo en la segmentación si existe una variación de intensidad suave entre regiones. Para evitar en lo posible este problema se restringe espacialmente la búsqueda de la moda, aunque esto produzca una gran cantidad de regiones, haciendo necesario agregar alguna manera de fusionar las regiones encontradas en la imágenes filtradas. La forma de fusionar es a través del análisis de adyacencia de las regiones, definiendo las restricciones que permiten decidir sobre si se fusionan o no dos regiones adyacentes.

### 6.1.3. Análisis de adyacencia

Así, el tercer paso en el proceso de segmentación tiene que ver con el análisis de adyacencia de regiones, para fusionar y disminuir el número de regiones encontradas en la imagen filtrada. La etapa de filtrado por CM produce una gran cantidad de regiones, para tener una idea más clara acerca del número de regiones involucradas en el análisis de adyacencias, la Tabla 6.2 muestra el número de regiones contenidas en una rebanada dada, para cada uno de los cinco volúmenes, a lo largo del análisis.

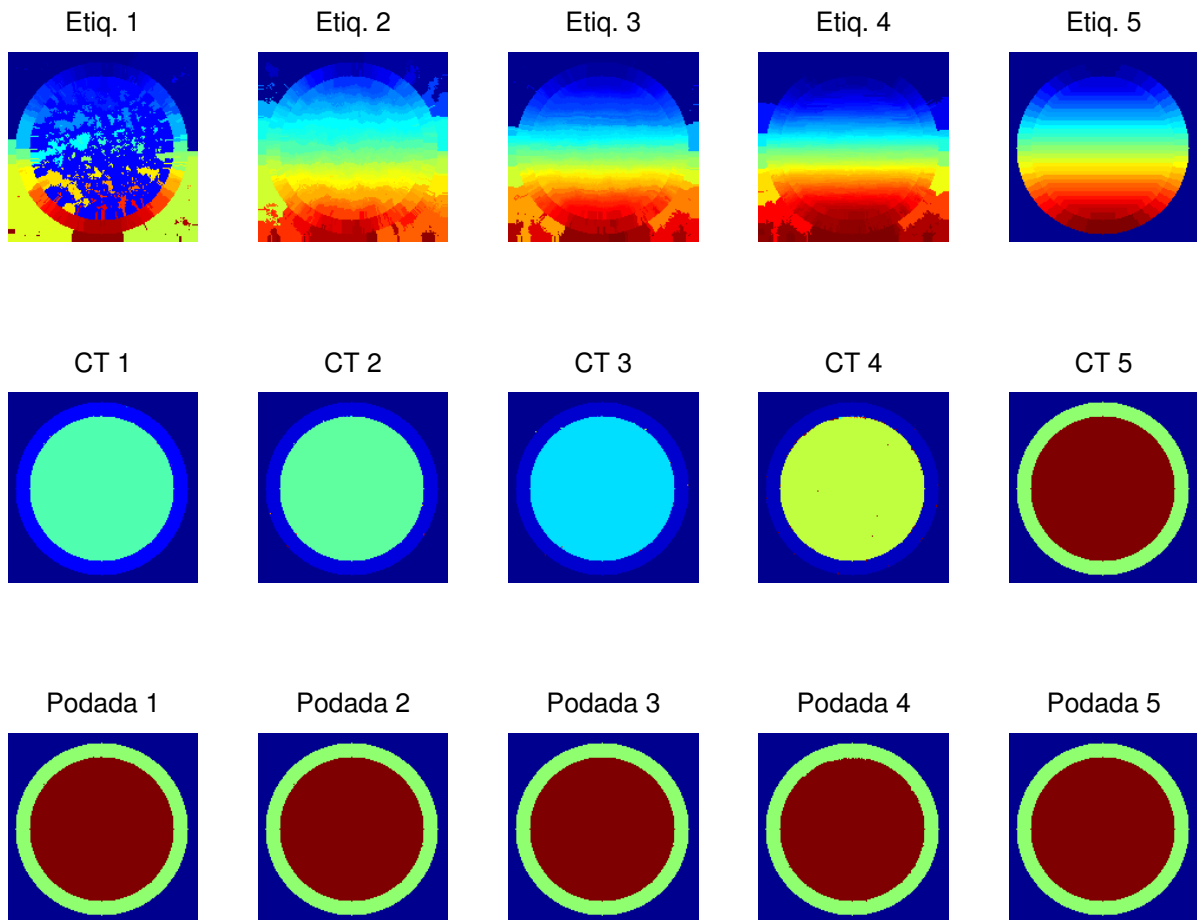


Figura 6.4: Análisis de adyacencia de regiones en esferas sintéticas. Primer renglón corresponde a las imágenes etiquetadas para una rebanada de cada modelo, cada color representa un etiqueta o equivalentemente, una región. Segundo renglón, imágenes después de fusionar las imágenes Etiq. 1 a 5, observándose una reducción sustancial en el número de regiones, y las imágenes Podada 1 a 5 son el resultado final de la segmentación, donde se podaron la regiones pequeñas, representadas por puntos difíciles de observar, de las imágenes CT1 a 5.

Para realizar el análisis de adyacencias, se requiere que a cada región homogénea, encontrada en el paso de filtrado, se le asigne una etiqueta para crear el grafo, quedando representado por su matriz de adyacencias de regiones (MAR). Las imágenes etiquetadas son obtenidas a partir de las imágenes producidas en el paso de filtrado por corrimiento de media. Las regiones homogéneas en intensidad están formadas por todos los pixeles que convergieron a la misma moda y para realizar el análisis de adyacencia, lo primero que se debe hacer es etiquetar cada región con un valor escalar, que corresponde a la etiqueta que identifica la región. El etiquetado se lleva a cabo para cada uno de los pixeles de la imagen filtrada, al primer pixel se le asigna el valor 0 como etiqueta y a continuación se analizan sus 8 vecinos, los vecinos que convergieron a la misma moda, se etiquetan con el mismo valor. De esta manera, la etiqueta 0 identifica a la primer región, la etiqueta 1 a la segunda y así sucesivamente para todas las regiones contenidas en la imagen filtrada. El resultado del etiquetado, para una rebanada de cada una de las cinco esferas, se muestra en el primer renglón de la Figura 6.4, así por ejemplo, la imagen Etiq. 1 requirió los valores de 0 a 353 para etiquetarla, el despliegue en color falso, muestra con un color diferente cada etiqueta. El número de etiquetas utilizado por cada imagen, se muestra en la columna Etiquetado de la Tabla 6.2.

De la imagen etiquetada se toma la información de adyacencias entre regiones, que es la información que permite establecer las regiones que están relacionadas entre sí y se almacena en la MAR como una estructura de datos que hace posible resolver cuestiones de alcanzabilidad; de esta matriz, se obtiene el vecindario de una región para intentar fusionar la región con alguna otra que tengan un borde en común. El criterio que determina si se fusionan o no dos regiones es que su diferencia en intensidad, normalizada con respecto al ancho de banda en intensidad, sea menor a 0.5 y que la fuerza del borde entre ellas, tomada del mapa de confianza, sea menor al umbral de fuerza de borde ( $\xi$ ). La condición de intensidad busca regiones que sean similares, y la condición de fuerza busca las regiones con bordes débiles entre ellas, así, si dos regiones son similares en intensidad y comparten un borde débil entonces se fusionan. Cuando se fusionan dos regiones se crea una nueva, modificándose el grafo y apareciendo nuevos bordes. El algoritmo de *encontrar y unir* es el responsable de efectuar la reducción del grafo y de agregarle los nuevos bordes.

Una vez realizado el análisis de adyacencias para todas las regiones, se reetiqueta la imagen y se calcula la moda de las regiones nuevas, con el promedio de las modas de las regiones que se fusionaron. Las operaciones de fusión, reducción del grafo, y actualización de la imagen resultante, se repiten iterativamente hasta que ya no es posible fusionar más regiones. El resultado de esta parte del análisis se muestra en el

segundo renglón de imágenes CT 1 a 5 de la Figura 6.4, para una rebanada de cada volumen sintético.

Después de haber aplicado las operaciones de fusión, la reducción del número de regiones homogéneas es significativa, sin embargo, aun está excedido en cuatro de las cinco imágenes, como se aprecia en la Figura 6.4. Las imágenes que siguen teniendo un número de regiones mayor al esperado son CT 1 a 4, aunque las imágenes se observan ya bastante homogéneas, contienen pequeñas regiones que son difíciles de observar precisamente por su tamaño, las cuales no pudieron ser unidas por la fusión. Conociendo de antemano las regiones contenidas en los modelos sintéticos, se concluye que estas pequeñas regiones deben pertenecer forzosamente a una de las regiones de mayor superficie, la forma de fusionar las regiones pequeñas es a través de un proceso de poda.

Así, el podado tiene como objetivo remover las regiones que tiene una superficie menor a un umbral preestablecido ( $\mu$ ), expresado como un número de píxeles. La poda toma las imágenes etiquetadas después de aplicar la fusión, como sus datos de entrada para construir la MAR, y empieza a recorrerla buscando las regiones menores al umbral. Cuando encuentra una región pequeña, con la información de la MAR determina sus regiones vecinas, y establece la región candidata con la cual fusionarla. La región candidata debe satisfacer las siguientes propiedades: es la más parecida en intensidad, o es la única con superficie mayor al umbral. En esta parte del proceso se vuelve a aplicar el algoritmo de *encontrar y unir* para reducción del grafo. Una vez recorrida toda la MAR, al igual que en la aplicación de la fusión, se reetiqueta la imagen y se recalculan sus modas, repitiéndose el proceso iterativamente hasta que el número de regiones permanece estable. El resultado del podado se muestra en el tercer renglón de la Figura 6.4 que contiene las imágenes Podada 1 a 5, donde se puede observar, que la delineación final, en todas las imágenes, ha encontrado exactamente el número de regiones esperadas. La implementación de las funciones utilizadas en el análisis de adyacencias se encuentra en el Apéndice D.

Como ya se mencionó, sólo uno de los cinco modelos sintéticos quedó con el número de regiones igual al esperado, la esfera 5, después de aplicar las operaciones de fusión, y esto pudiera parecer un error del método de segmentación propuesto. Sin embargo, la aplicación de la etapa final de podado demuestra su utilidad para corregir este aparente error y lograr la identificación de las regiones que integran la imagen. Recordando, que la esfera 5 es similar a la esfera 2, en cuanto a contenido de inhomogeneidad y que difieren en el contenido de ruido, la primera no contiene ruido y la segunda si, se puede concluir que en ausencia de ruido y con una inhomogeneidad moderada, la fusión es capaz de unir adecuadamente las imágenes, sin la necesidad de aplicar el podado.

El resultado final del proceso de segmentación son las imágenes podadas, las cuales contienen solamente dos clases (tres si consideramos el fondo), lo que demuestra que el proceso es capaz de identificarlas y separarlas, no obstante la presencia de ruido y el traslape (aun el total) de intensidades en las clases de las imágenes originales.

## 6.2. Cerebro normal simulado

El segundo volumen sintético fue obtenido del BIC y corresponde a un fantoma digital. Este fantoma consiste de imágenes ponderadas T1, el tamaño del volumen es de  $181 \times 217 \times 131$  con un espesor de rebanada de 1 mm, con ruido de 3% relativo al tejido más brillante y una inhomogeneidad de intensidad del 20%. Para este volumen, la “verdad de tierra” también se encuentra disponible en el BIC para efectos de comparación cuantitativa; para medir la similitud entre la imagen segmentada con la metodología propuesta y la verdad de tierra, considerada como la referencia, se evaluó el coeficiente de Tanimoto para el volumen. Es importante mencionar que los índices de Tanimoto son métricas de desempeño global que han sido criticados por su falta de sensibilidad local para el error; sin embargo, han llegado a ser la métrica estándar en el análisis de segmentación de IRM, y para propósitos de comparación de desempeño aun se deben de usar.

### 6.2.1. Parámetros del proceso

La segmentación de una imagen requiere de la aplicación de una serie de pasos para lograr la identificación de las estructuras anatómicas del cerebro. En cada uno de estos pasos intervienen diferentes parámetros, afectando en mayor o menor grado la delineación final de la imagen, por lo que es importante hacer una buena selección de parámetros. Todos los parámetros ya fueron presentados en los resultados del fantoma esférico, ahora se mostrará el efecto de su variación en el proceso.

El primer paso del proceso tiene que ver con el cálculo del mapa de confianza, que mide la confianza para la presencia de un borde en los datos del volumen a segmentar utilizando dos parámetros: el umbral de borde ( $\delta$ ) y el factor de mezcla ( $\beta$ ). El primero, determina si se calcula o no la confianza de borde para una dato dado. Si el valor del gradiente normalizado del dato, es decir, su percentil en la función de distribución acumulativa de magnitud del gradiente, es mayor a este umbral de borde, entonces se calcula su confianza de borde, y en caso contrario se le asigna un cero como valor de confianza. Una vez que se tienen calculados los valores de confianza, el factor de mezcla

determina la proporción de magnitud del gradiente y confianza de borde que se incluyen en el mapa de confianza. Ambos parámetros toman valores entre 0 y 1.

Con la finalidad de observar el efecto producido por los parámetros de umbral de borde y factor de mezcla en el mapa de confianza, se experimentó con una misma rebanada del cerebro simulado, obteniéndose las imágenes mostradas en la Figura 6.5. Cuando el umbral de borde tiene un valor de cero, es equivalente a permitir cualquier valor de magnitud del gradiente del pixel para considerarlo parte de un borde, es decir todos los pixeles forman parte de un borde, debido a esto la primer imagen (esquina superior izquierda) de la Figura 6.5 muestra un gran cantidad de bordes. A medida que se aproxima el valor del umbral de borde a 1 la cantidad de bordes disminuye, llegándose al extremo de no tener bordes en la imagen, dado que la restricción para calcular la confianza del borde es imposible de cumplirse.

El renglón superior de la Figura 6.5 corresponde a un factor de mezcla igual a cero, esto es, el mapa de confianza contiene exclusivamente los valores de confianza de borde de los datos, descartando por completo la información del gradiente. El aumento gradual del valor del factor de mezcla, hace que aumente el contenido de gradiente en el mapa y disminuya el contenido de confianza del borde. El otro extremo del factor de mezcla, cuando vale 1, corresponde al mapa de confianza formado exclusivamente con el gradiente de la imagen; es por esto que en el renglón inferior no existe diferencia en las imágenes.

El factor de mezcla y umbral de borde, como ya se mencionó, sólo participan en el cálculo del mapa de confianza, sin embargo, aun cuando ya se conoce el impacto que tienen en el mapa de confianza, aun no se tiene un criterio para seleccionarlos. Para realizar la selección, observamos tanto el efecto cualitativo como cuantitativo del mapa de confianza al final del proceso. La Figura 6.6 muestra las imágenes de probabilidad de pertenencia obtenidas con los mapas de confianza de la Figura 6.5, mostrando el efecto cualitativo. El efecto cuantitativo es observado en el índice de similitud, dado por los coeficientes de Tanimoto, entre la imagen de pertenencia y la segmentación manual del experto de la imagen correspondiente.

En la Figura 6.7 se muestran los índices de similitud agrupados de acuerdo al valor del parámetro. Así, el primer grupo se refiere a un valor cero en el factor de mezcla, seguido por el grupo con valor 0.3 y así sucesivamente. Dentro de cada grupo se muestran cinco barras apiladas correspondiendo a las clases de Fondo, LCR, MG y MB. La primera barra en el grupo representa la suma de las cuatro clases mencionadas para un valor de umbral de borde de cero, la segunda con valor de 0.3 y así sucesivamente. No se observa una diferencia muy marcada en los índices de similitud para realizar la

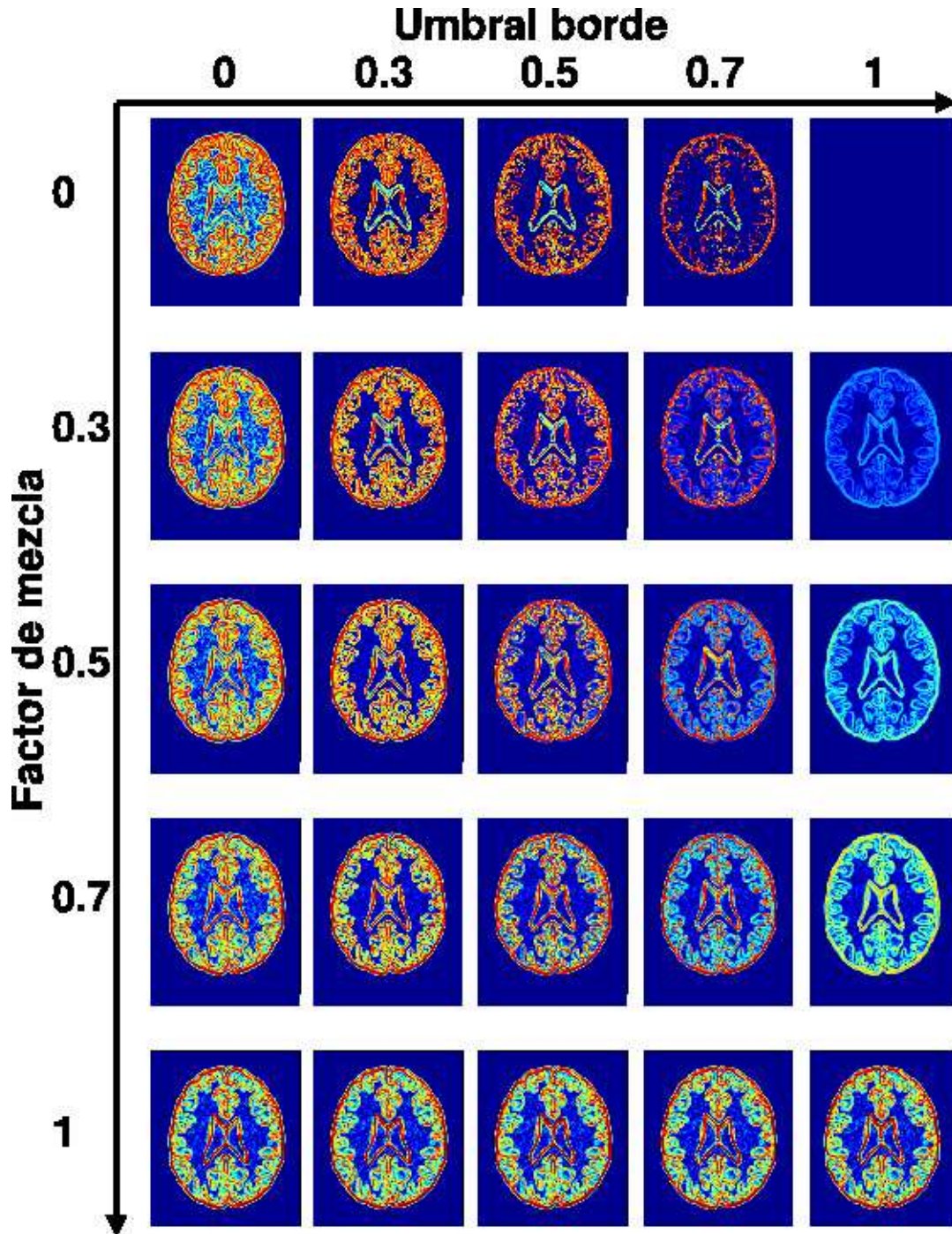


Figura 6.5: Mapas de confianza. Los parámetros umbral de borde y factor de mezcla se variaron de la misma manera entre 0 y 1 produciendo mapas distintos para una misma rebanada del volumen del cerebro simulado.



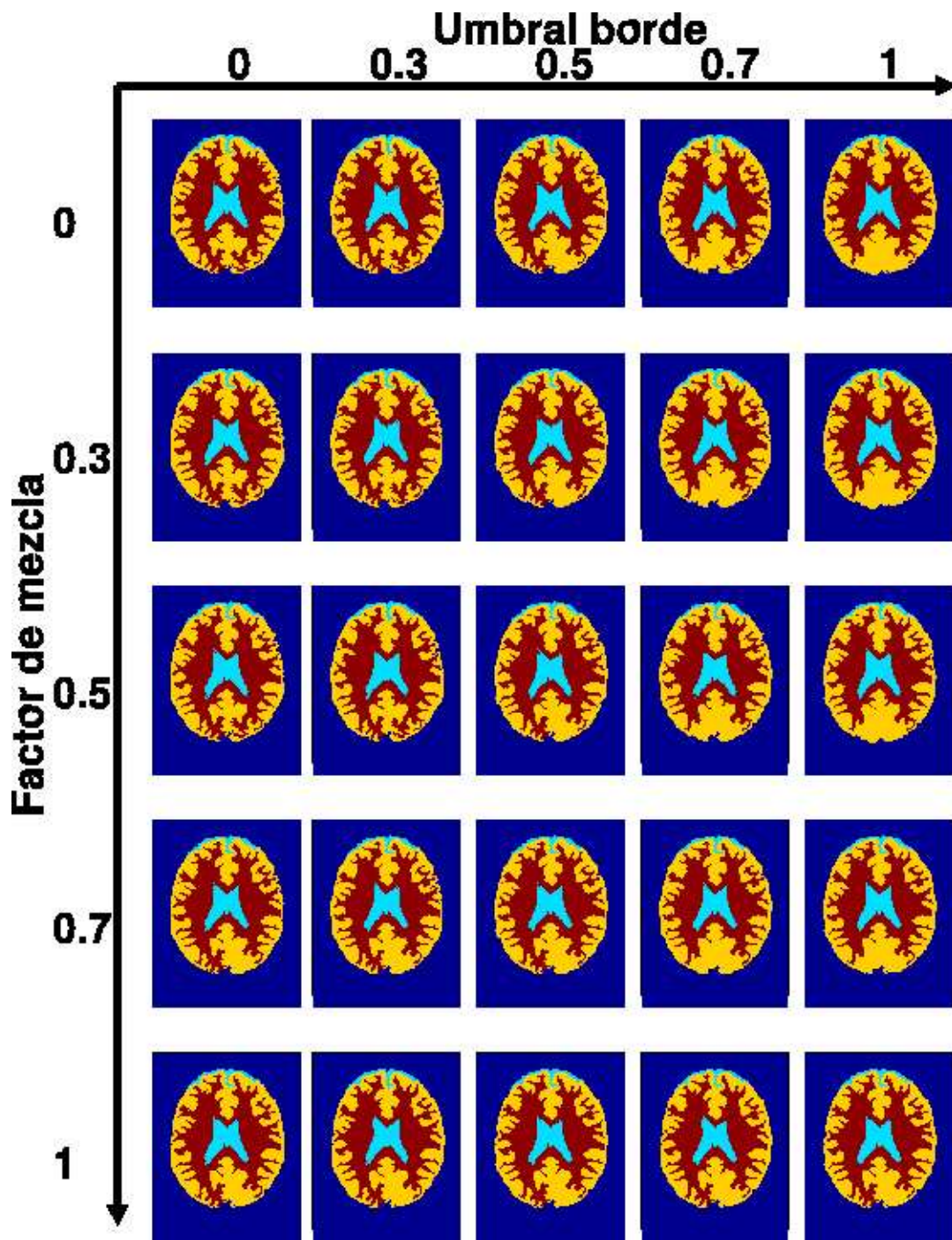


Figura 6.6: Imágenes clasificadas con probabilidad de pertenencia variando el umbral de borde y el factor de mezcla.

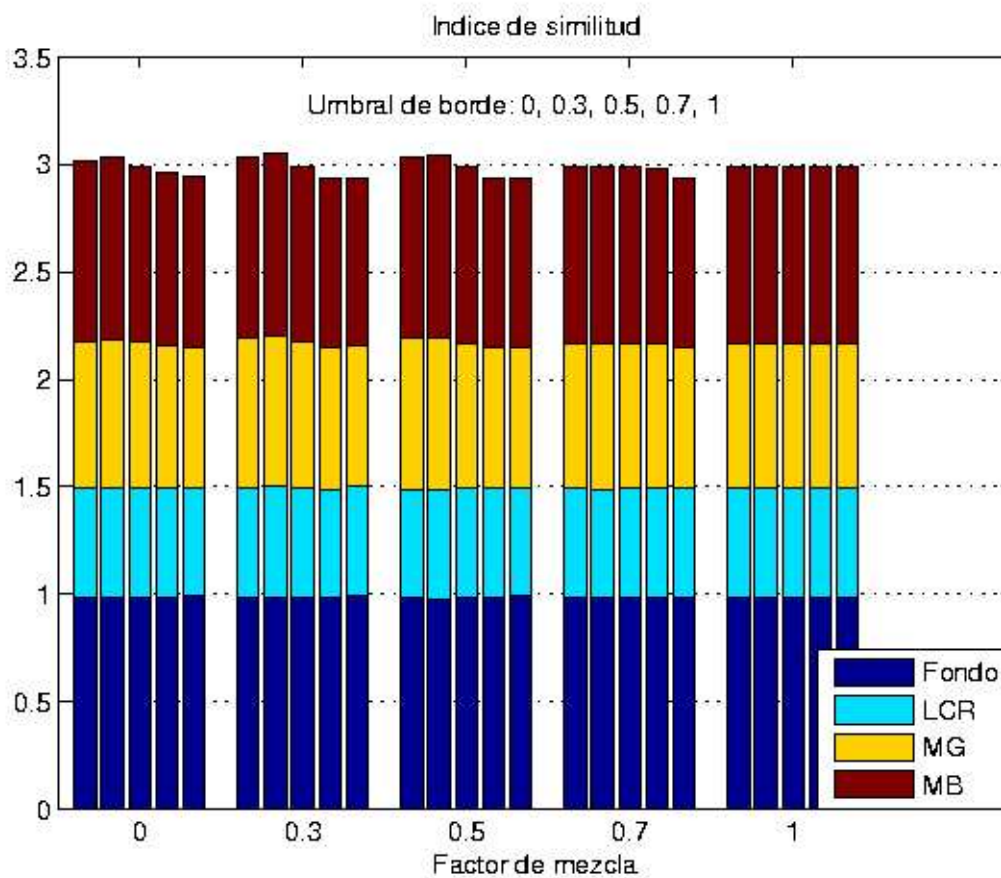


Figura 6.7: Índice de similitud variando el factor de mezcla y el umbral de borde. Cada grupo de barras corresponde al factor de mezcla indicado en la base del grupo, y dentro del grupo, cada barra corresponde al valor de umbral de borde indicado en la leyenda mostrada arriba de las barras.

selección de parámetros, ya que los valores apilados varían de 2.931 a 3.051 presentando una variación de apenas 0.12, lo que hace necesario, que adicionalmente, se tome en cuenta la información cualitativa de la Figura 6.6, donde se observa que con umbral de borde = 0.5 se empieza a clasificar erróneamente la región de materia blanca. Por lo tanto, considerando la información cualitativa y cuantitativa se eligió un factor de mezcla de 0.3 y un umbral de borde de 0.4, ya que con estos valores se aprecia una buena delineación de las estructuras anatómicas de interés del cerebro y el índice de similitud sumado para las cuatro clases es un máximo. En la Figura 6.7 no se observa el valor de umbral de borde de 0.4, debido a que se buscó simplificar la representación de la información y mostrar la tendencia de los resultados a intervalos regulares; sin embargo, en los experimentos el mejor índice de similitud se alcanzó con un umbral de borde = 0.4.

El paso de filtrado por CM es la parte principal del proceso de segmentación planteado, y los parámetros que intervienen en esta etapa son: los radios espacial ( $h_e$ ) y de intensidad ( $h_i$ ) del kernel utilizado en la estimación de la función de densidad. La Figura 6.8 muestra los efectos de ambos radios en el filtrado de la imagen. En dicha Figura, el eje horizontal representa al radio de intensidad y se observa que la calidad de la segmentación depende considerablemente de la resolución en intensidad: un radio  $h_i$  pequeño produce muchas regiones, sobresegmentando así la imagen, mientras que un radio de intensidad mayor puede producir pocas regiones, subsegmentando la imagen.

Por lo que respecta al radio espacial, representado en el eje vertical, se observa que cuando  $h_e$  aumenta, solamente las características de la imagen con soporte espacial grande se representan en la imagen filtrada. Esto puede observarse con mayor facilidad en las imágenes con pocas regiones, por ejemplo las dos columnas más a la derecha, donde se observa que las regiones que prevalecen son las mayores, a medida que se incrementa el radio espacial. Un efecto adicional del radio espacial, que no es observado en la gráfica, pero que es importante resaltar es que tiene un impacto mayor en el tiempo de ejecución del filtrado: a mayor radio espacial, mayor tiempo de procesamiento.

La elección de los radios del kernel se apoya en las imágenes de probabilidad de pertenencia y sus índices de similitud. De las imágenes de probabilidad de la Figura 6.9 se observa que el radio de intensidad debe ser menor de 32 para que se mantenga la calidad de la segmentación. Por lo que respecta al radio espacial se observa que algunas partes de la MB que no aparecen en radios menores, aparecen cuando aumenta el radio, lo que puede ser indicativo de que la calidad de segmentación se mejora con radios espaciales grandes. Sin embargo, se debe recordar que a mayor radio espacial, mayor tiempo de proceso computacional, por lo que se debe establecer un compromiso entre

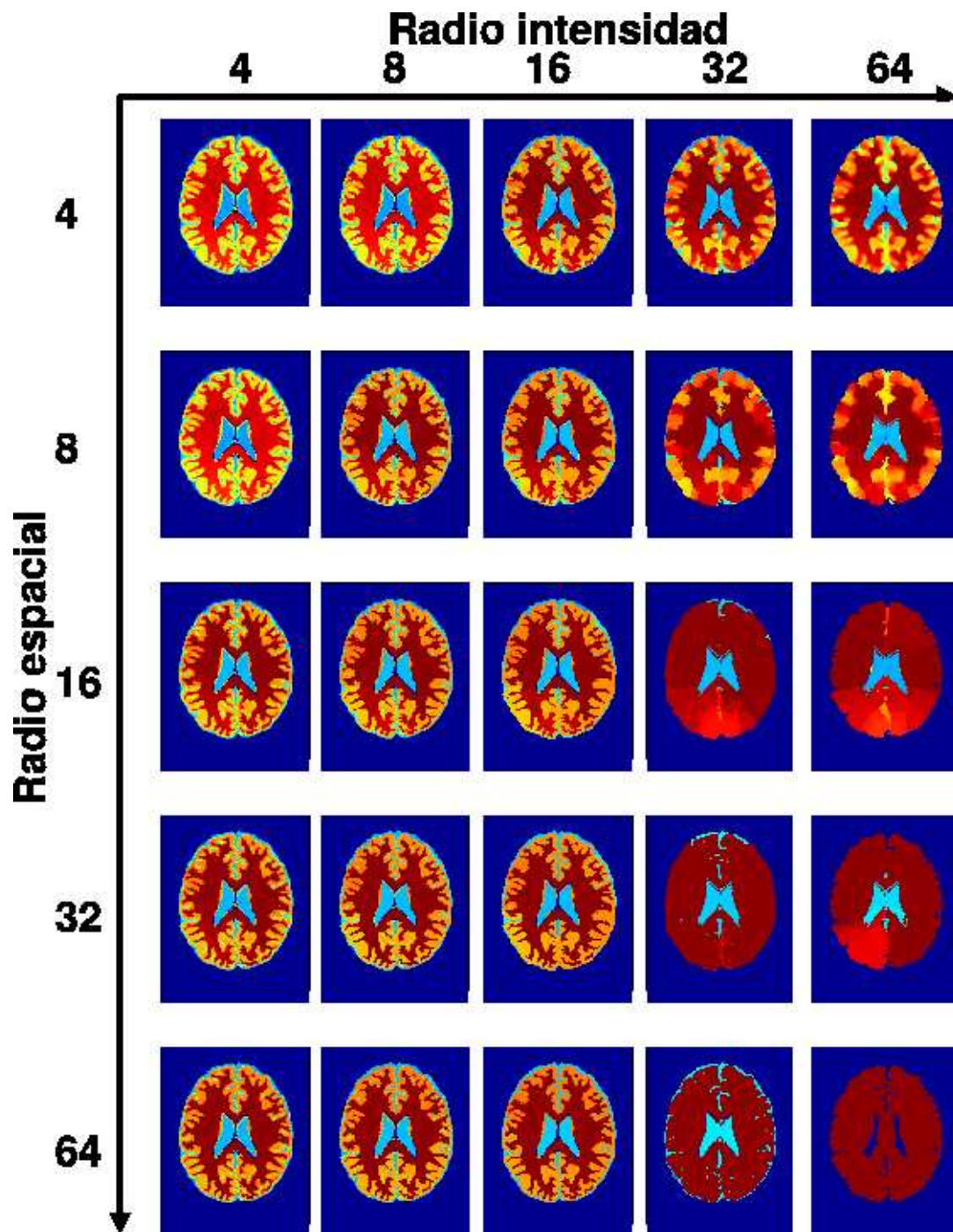


Figura 6.8: Imagen filtrada por CM variando los radios espacial y de intensidad del kernel.

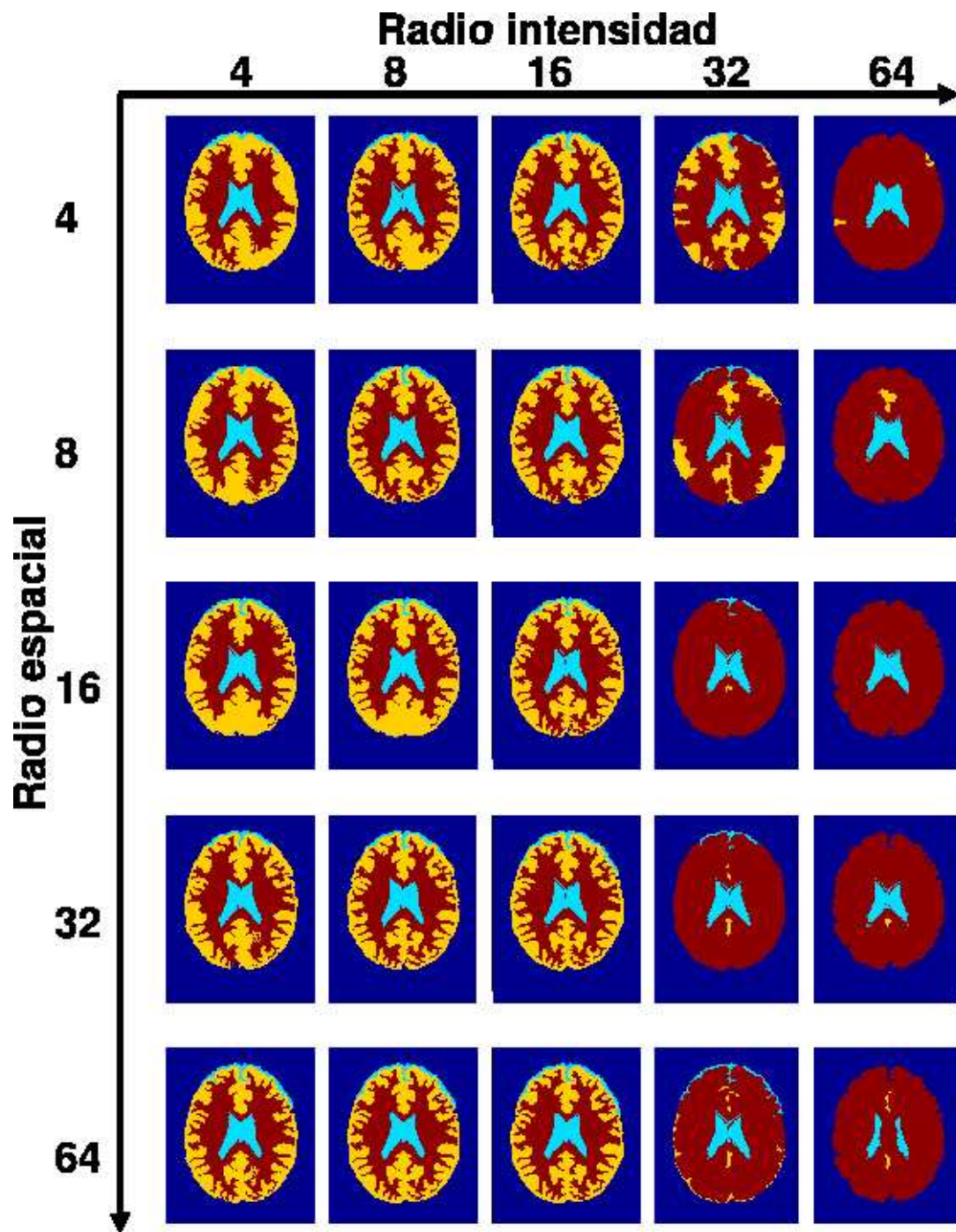


Figura 6.9: Imágenes clasificadas con probabilidad de pertenencia variando el radio espacial y el radio de intensidad del kernel.

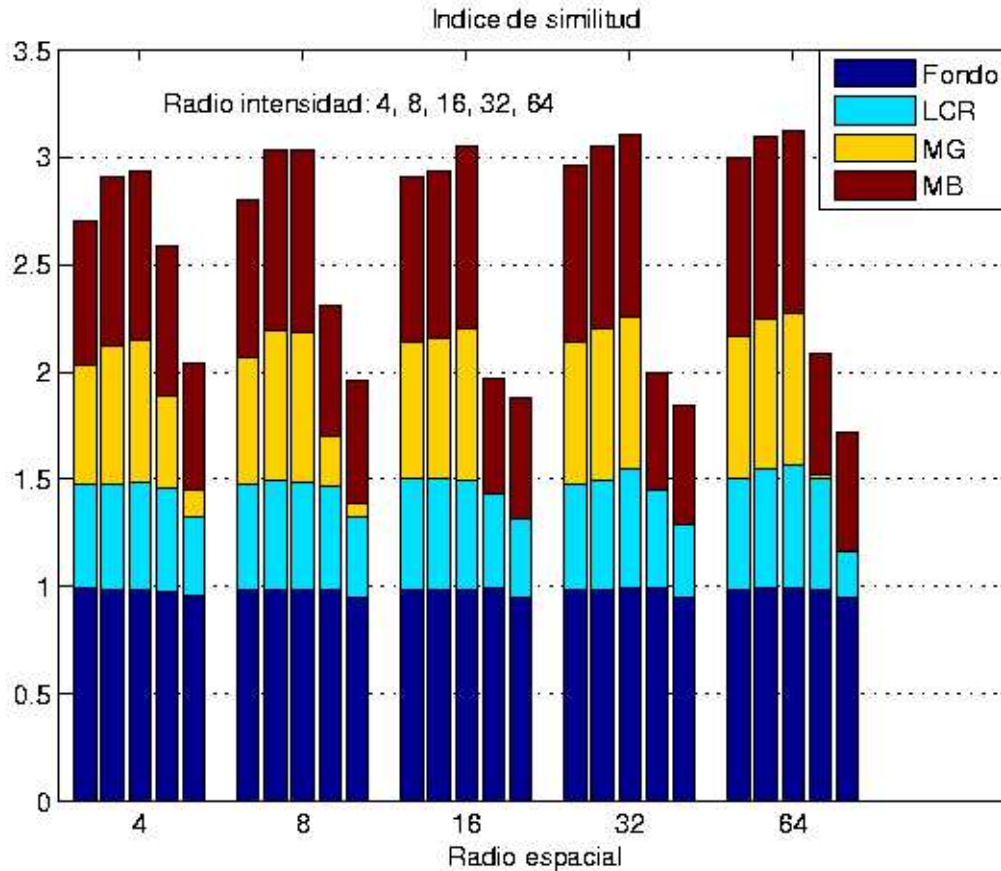


Figura 6.10: Índices de similitud entre imágenes de probabilidad de pertenencia y segmentación manual, variando los radios espacial y de intensidad. Cada grupo de barras corresponde al radio espacial indicado en la base del grupo, y dentro del grupo, cada barra corresponde al valor del radio de intensidad indicado en la leyenda mostrada arriba de las barras.

calidad de segmentación y tiempo de proceso.

Observando los índices de similitud de la Figura 6.10, para las imágenes de la Figura 6.9, es evidente que un radio de intensidad de 32 o mayor (4ª y 5ª barras en todos los grupos) produce los índices más bajos, y para todos los grupos de los diferentes valores de radio espacial, los índices de similitud mayores están en los radios de intensidad de 8 y 16, basándose en lo anterior se eligió un radio de intensidad = 9, y principalmente para acelerar el tiempo de proceso, un radio espacial = 6.

El tercer paso del proceso de segmentación planteado, es el análisis de adyacencias para la fusión de regiones, por medio de la aplicación iterativa de las operaciones de fusión como ya se explicó. Los resultados de aplicar la fusión, a una imagen dada,

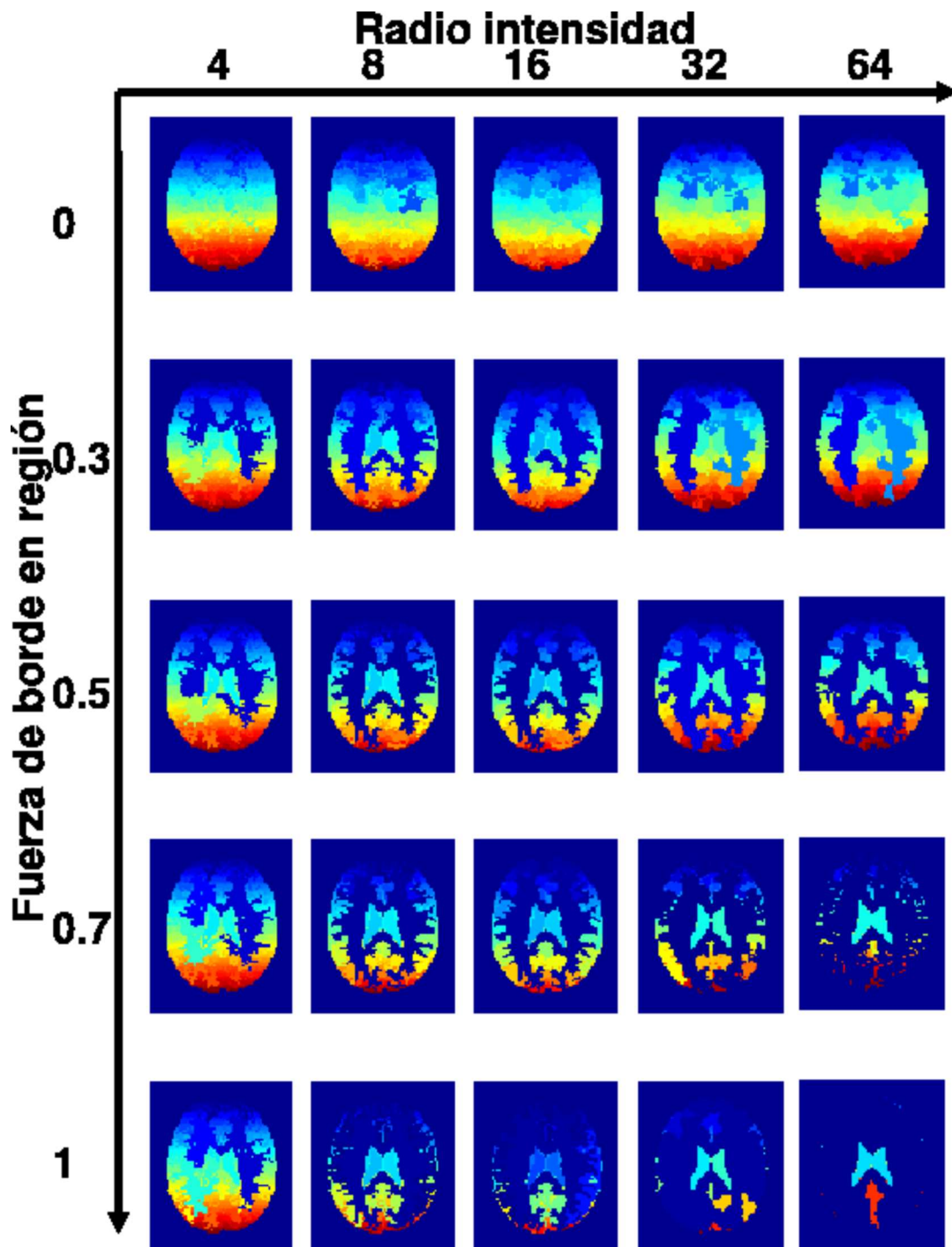


Figura 6.11: Imágenes obtenidas después de aplicar las operaciones de fusión sobre la imagen filtrada, variando los parámetros de fuerza de borde entre regiones y radio de intensidad.

variando los parámetros que participan en ella: el umbral de fuerza de borde ( $\xi$ ) entre regiones vecinas y el radio de intensidad ( $h_i$ ) se muestran en la Figura 6.11; en esta figura, nuevamente se observa, que la mejor calidad cualitativa de la segmentación de imagen se obtiene con radios (anchos de banda) de intensidad menores de 32. Ya que la fuerza de borde es el promedio de la confianza de los pixeles que están situados en el borde o frontera entre dos regiones, el darle un valor de cero a este umbral, significa que ningún borde se considera débil y todos los bordes se deben conservar, es decir, no se permite la fusión de ninguna de las regiones, como puede observarse en el renglón superior de la Figura 6.11. Al ir aumentando el umbral de fuerza de borde se aprecia la mejora en la calidad de la segmentación; ya que empiezan a identificarse con mayor claridad las estructuras correspondientes a LCR, MG y MB. Sin embargo, cuando el umbral toma el valor extremo de 1, la delineación de las regiones se degrada, debido a que ahora todos los bordes se consideran débiles y la fusión de regiones queda sólo en función del ancho de banda.

La información cualitativa y cuantitativa mostrada en las Figuras 6.12 y 6.13 respectivamente, muestra que la elección del valor de umbral de fuerza de borde debe ser cercano a uno, por lo que el valor elegido para este parámetro fue de 0.8.

El último parámetro por analizar es el tamaño mínimo de región utilizado en la fase de podado del paso de análisis de adyacencias. Este parámetro permite fusionar aquellas regiones que son pequeñas, comparadas con un tamaño mínimo de región ( $\mu$ ) predefinido, con regiones de tamaño mayor y teniendo una diferencia de moda de intensidad mínima. El tamaño mínimo de región tiene poca influencia en la calidad de la segmentación como se observa en las imágenes podadas de la Figura 6.14; ya que existe poca diferencia en las imágenes podadas con tamaños de región diferentes.

Las Figuras 6.15 y 6.16 también evidencian la poca influencia que tiene el parámetro de tamaño mínimo de región. Aunque es interesante observar que cuando el radio de intensidad es pequeño, la imagen clasificada por pertenencia cualitativamente es de poca calidad, sin embargo, la calidad se mejora a medida que se aumenta el umbral de tamaño mínimo, como se muestra en la primer columna a la izquierda de la Figura 6.15. El valor elegido para el tamaño mínimo de región fue de 30 pixeles, con ese valor no se apreciaron cualitativamente regiones pequeñas empujadas en otras, y además con ese valor no se comprometen las estructuras anatómicas pequeñas.

Los valores seleccionados, para cada uno de los parámetros, se eligieron de acuerdo al efecto que tienen en el paso de la segmentación en que participan, como ya se mencionó. De todos los parámetros que son utilizados en el proceso, los que tienen mayor relevancia son los anchos de banda del kernel y de manera especial el ancho de banda de intensidad,



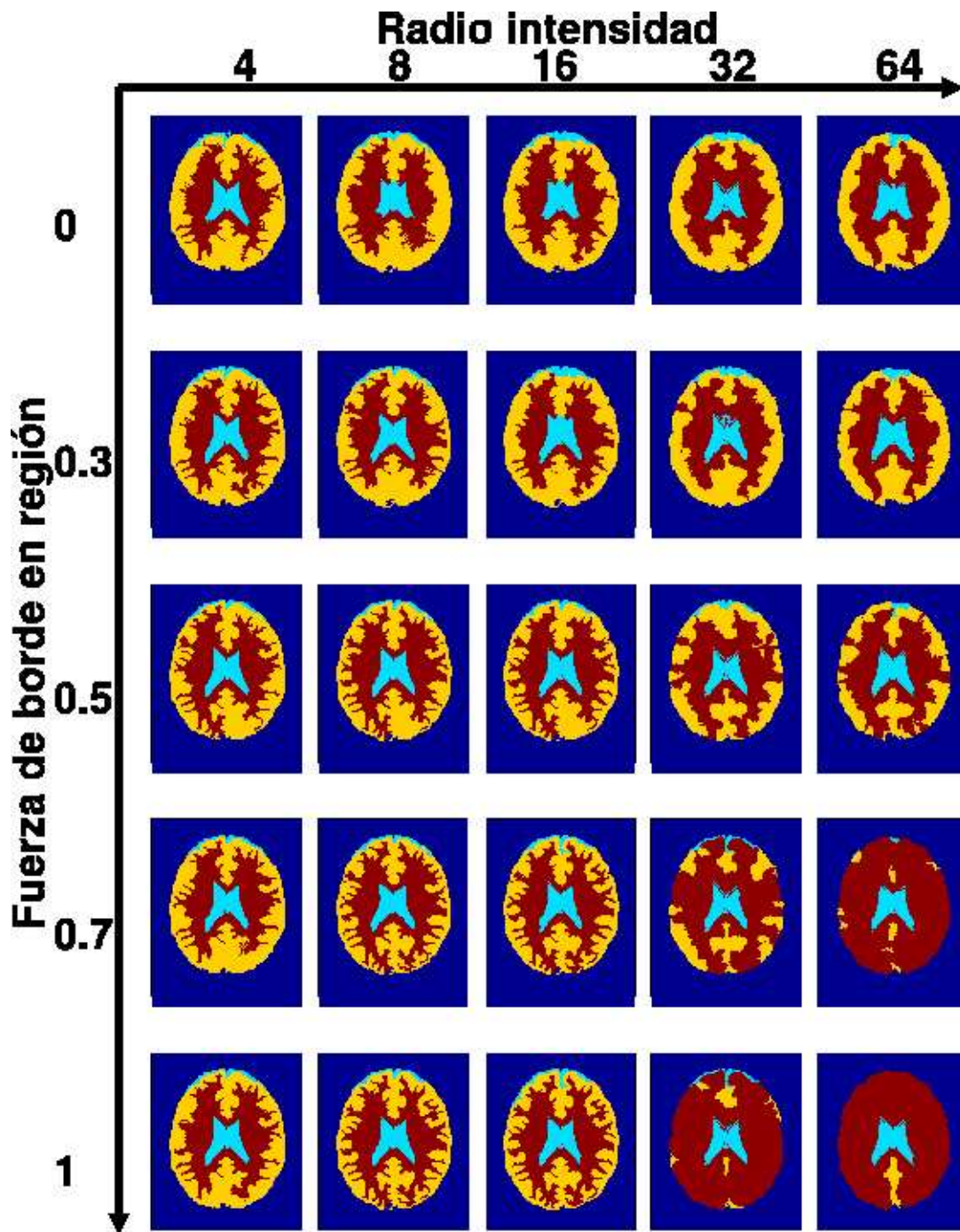


Figura 6.12: Imágenes clasificadas con probabilidad de pertenencia variando los parámetros de umbral de fuerza de borde entre regiones y radio de intensidad.

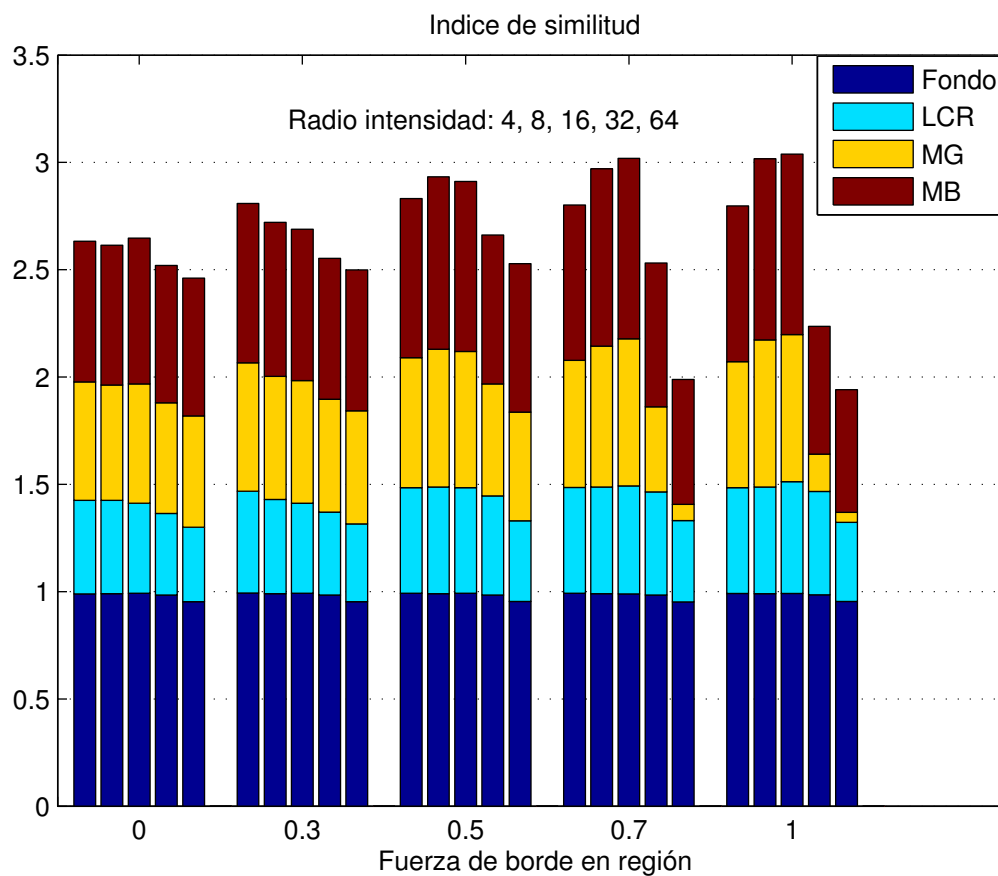


Figura 6.13: Índices de similitud entre imágenes clasificadas por pertenencia y segmentación manual, variando el umbral de fuerza entre región y el radio de intensidad. Cada grupo de barras corresponde al umbral de fuerza de borde indicado en la base del grupo, y dentro del grupo, cada barra corresponde al valor del radio de intensidad indicado en la leyenda mostrada arriba de las barras.

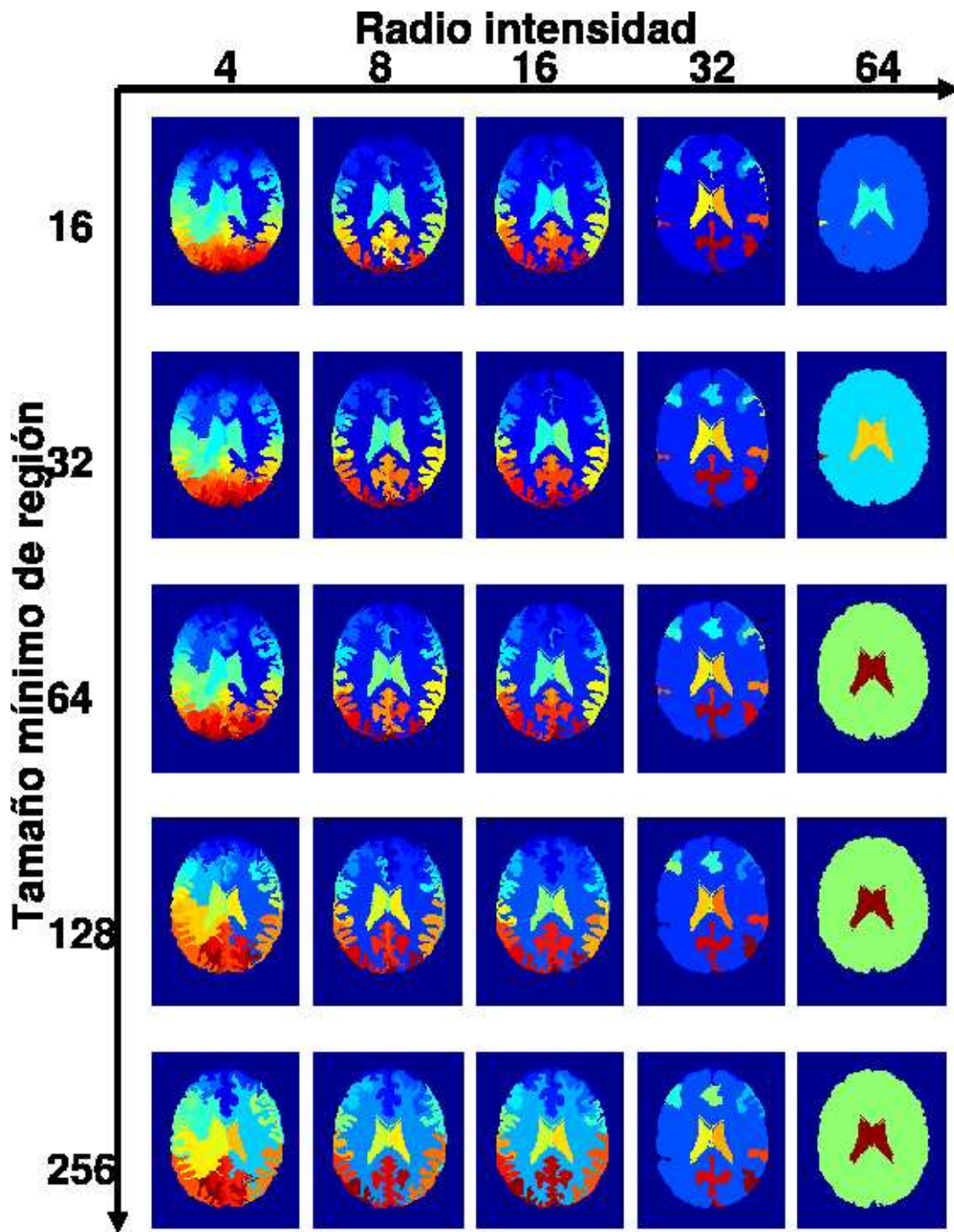


Figura 6.14: Imágenes resultado de la fase de podado variando los parámetros de tamaño mínimo de región y radio de intensidad.

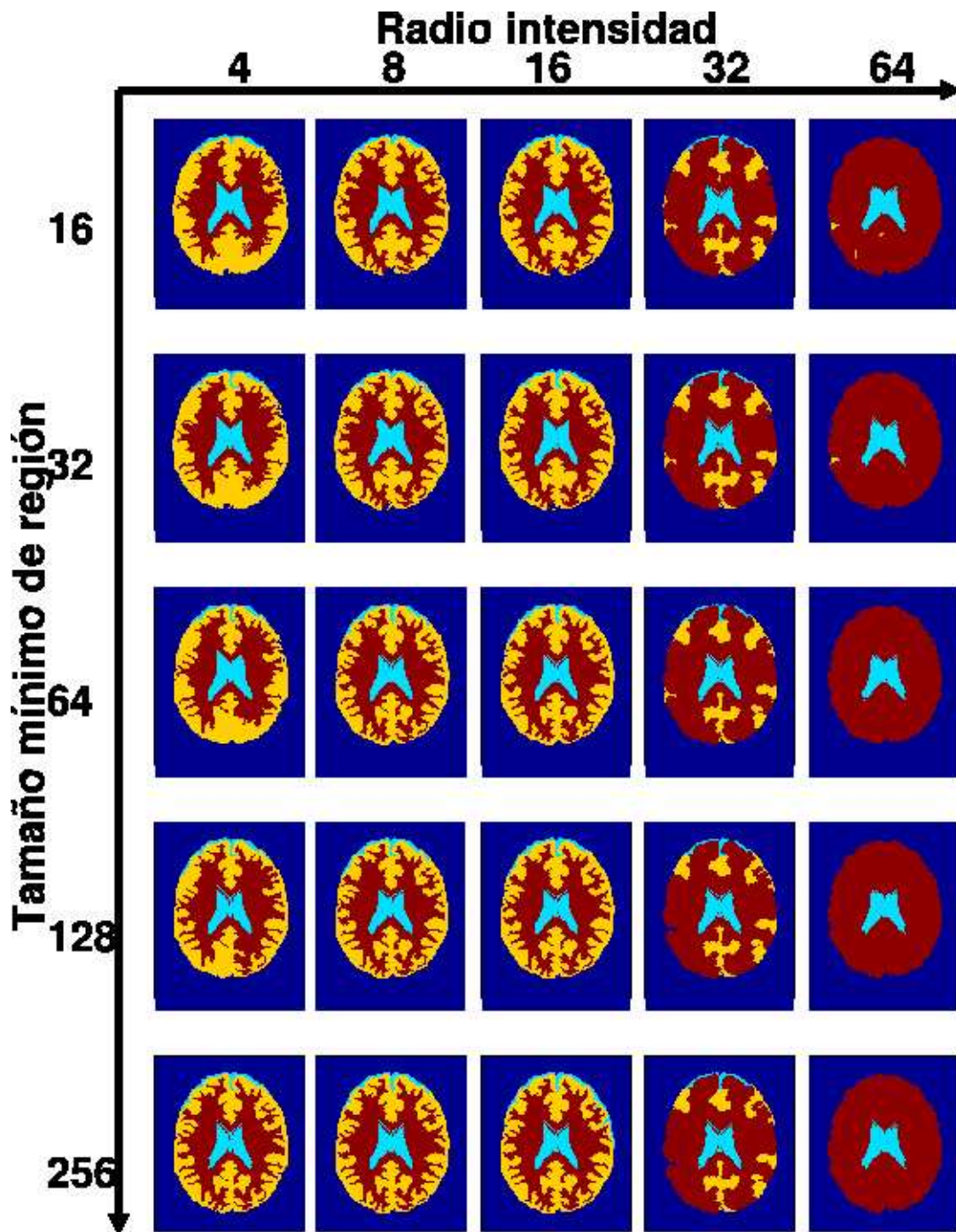


Figura 6.15: Imágenes clasificadas con probabilidad de pertenencia variando los parámetros de radio de intensidad y tamaño mínimo de región.

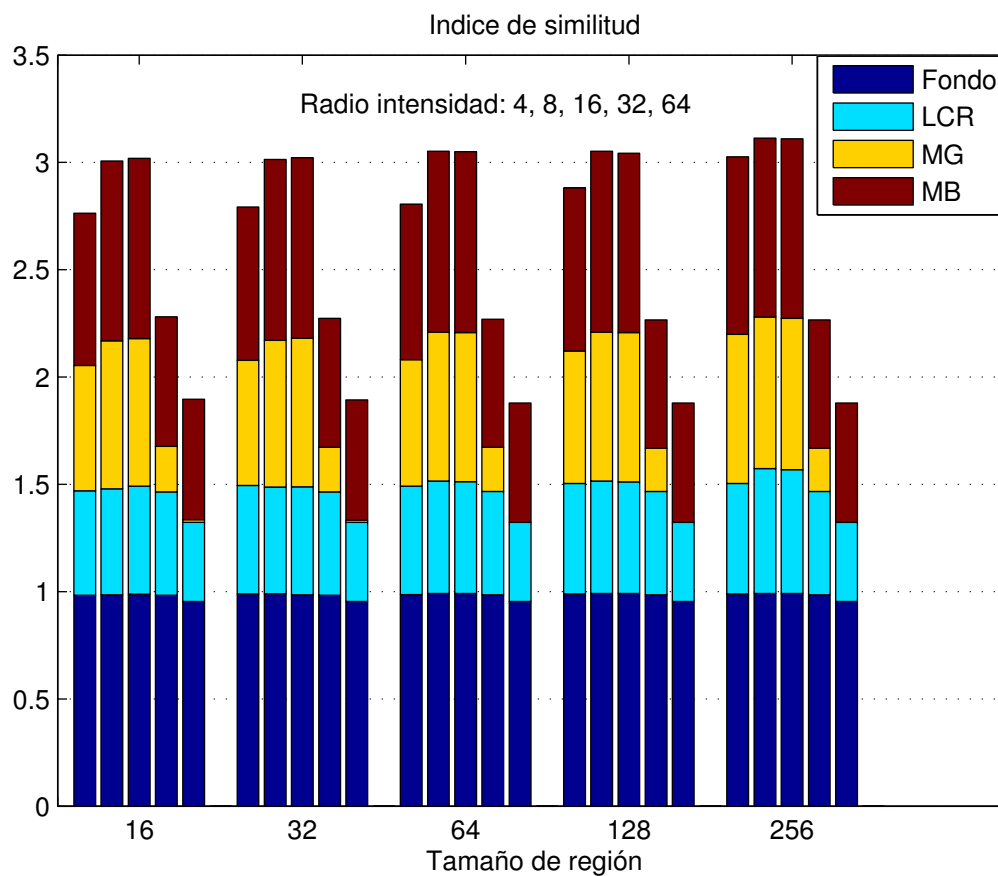


Figura 6.16: Índices de similitud entre imágenes clasificadas por pertenencia y segmentación manual, variando el tamaño mínimo de región y el radio de intensidad. Cada grupo de barras corresponde al tamaño mínimo de región indicado en la base del grupo, y dentro del grupo, cada barra corresponde al valor del radio de intensidad indicado en la leyenda mostrada arriba de las barras.

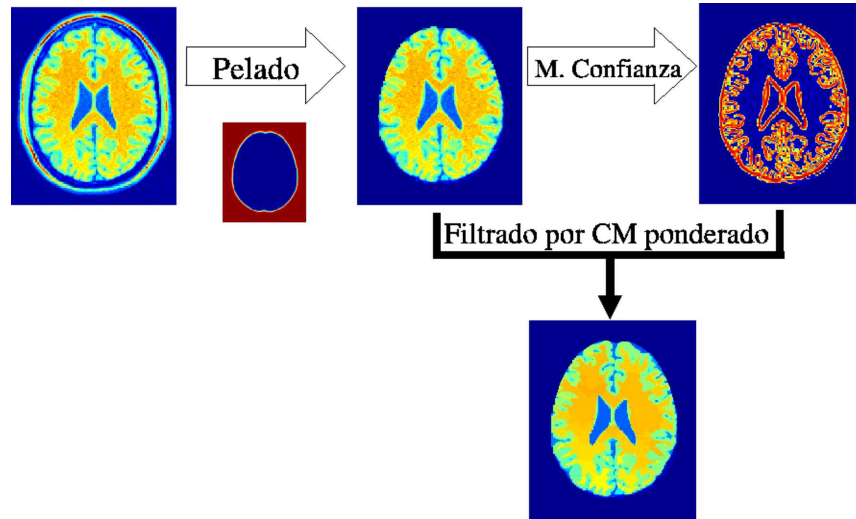


Figura 6.17: Filtrado de la imagen. El proceso inicia con el pelado de los datos, seguido por el cálculo del mapa de confianza y finalmente, el filtrado de los datos aplicando el algoritmo de CM ponderado por el mapa de confianza.

puesto que es el que tiene el mayor impacto en la calidad de la regiones delineadas. En resumen, los valores seleccionados para procesar el volumen simulado del cerebro son: factor de mezcla ( $\beta$ ) = 0.3, umbral de borde ( $\delta$ ) = 0.4, radio espacial ( $h_e$ ) = 6, radio de intensidad ( $h_i$ ) = 9, umbral de fuerza de borde entre región ( $\xi$ ) = 0.8 y tamaño mínimo de región ( $\mu$ ) = 30. Los resultados obtenidos, al procesar el volumen del cerebro simulado con estos parámetros, se presentan a continuación.

### 6.2.2. Filtrado del volumen simulado

El paso de filtrado de los datos inicia con la remoción de estructuras que no son de interés como son piel y hueso, esto se lleva a cabo aplicando una máscara obtenida de la verdad de tierra, que incluye fondo, grasa, músculo, piel y hueso; todo esto englobado en la clase denominada fondo. Esta remoción de regiones que no son de interés se le conoce comúnmente como el “pelado” de los datos. Posteriormente, se procede al cálculo del mapa de confianza de los datos, los cuales se filtran aplicándoles el algoritmo de CM, ponderado por el mapa de confianza para tener un filtrado con preservación de bordes. El proceso del paso de filtrado descrito se muestra en la Figura 6.17, donde se puede apreciar en la imagen filtrada, la desaparición del ruido y la presencia de regiones con intensidades más homogéneas, generadas a partir de las modas encontradas por el CM.

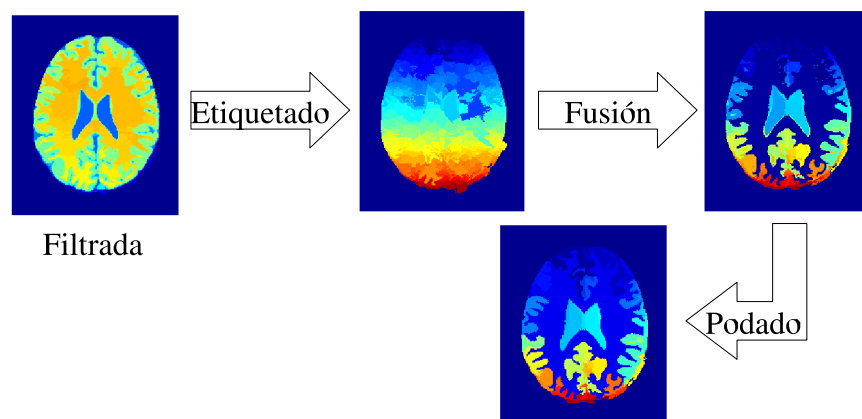


Figura 6.18: Análisis de adyacencias de región. En la imagen filtrada se etiquetan las regiones homogéneas, las cuales son unidas por la aplicación iterativa de operaciones de fusión y finalmente se podan las regiones pequeñas.

### 6.2.3. Análisis de adyacencias de región

El filtrado de los datos produce una gran cantidad de regiones con intensidades homogéneas, las cuales deben analizarse en el contexto de su vecindario y determinar si deben o no fusionarse con regiones vecinas. El objetivo del análisis de adyacencias es fusionar regiones que compartan un borde débil entre ellas y cuya diferencia en intensidad, medida como la distancia normalizada entre sus modas de intensidad, no sea mayor a un umbral preestablecido, como ya se explicó anteriormente.

La secuencia que se sigue en el paso de análisis de adyacencias de región se muestra en la Figura 6.18. Los datos de entrada a este paso son los de la imagen filtrada que contiene las regiones homogéneas encontradas por el algoritmo CM; a cada una de estas regiones se le asigna una etiqueta, representada por un valor escalar, para estar en condiciones de poder aplicar las operaciones de fusión y unir aquellas regiones que cumplan con los criterios expresados por el umbral de fuerza de borde entre región y el radio espacial. Después de que el número de regiones ya no cambia con la aplicación de la fusión, se observa cómo las estructuras anatómicas de interés empiezan a delinearse en la imagen. El refinamiento de las regiones encontradas por la fusión se logra aplicando un proceso de poda, eliminando las regiones muy pequeñas, lo que permite que se tenga una mejor definición de la segmentación de la imagen.

Hasta aquí, los pasos del proceso de segmentación son idénticos en ambos tipos de datos sintéticos, y para el fantoma esférico fueron suficientes para identificar las clases contenidas en el modelo. Para los datos del cerebro no es suficiente tener las regiones homogéneas, sino que es necesario asociar cada una de ellas con algún tipo de

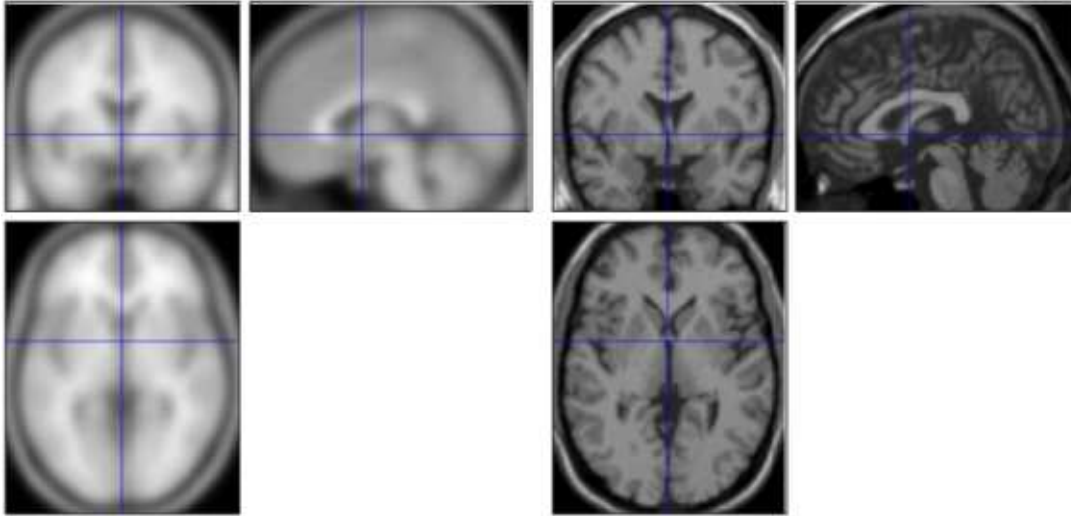


Figura 6.19: Normalización espacial entre los datos del cerebro simulado y la plantilla del SPM.

tejido anatómico cerebral. Para satisfacer esta necesidad se deben aplicar dos pasos más en el proceso de segmentación: la normalización espacial entre los datos y los mapas de probabilidad *a priori*, y la clasificación de regiones con el cálculo de sus probabilidades de pertenencia.

#### 6.2.4. Registro de datos y mapas de probabilidad

Antes de clasificar una imagen, es necesario determinar la transformación espacial que mapea cada voxel en la imagen a su ubicación correspondiente en las imágenes de los mapas de probabilidad *a priori*. El mapeo se realizó registrando el volumen de datos, con imágenes plantilla que están en el mismo espacio estereotáctico que los mapas de probabilidad. Estos mapas representan la probabilidad *a priori* de que un voxel pertenezca a LCR, MG o MB.

Es posible obtener un registro razonable de las imágenes de cerebros normales con una imagen plantilla de la misma modalidad, determinando una matriz de transformación afín de doce parámetros. Tanto las plantillas y los mapas probabilísticos se distribuyen con el programa SPM, y con este mismo programa se realizó el registro entre los datos y la plantilla. Dado que los mapas están en el mismo espacio que la plantilla, los datos quedan automáticamente registrados con los mapas. El resultado del mapeo se muestra en la Figura 6.19. En esta Figura las tres imágenes de la izquierda representan la plantilla, y las de la derecha son los datos.

El SPM entrega una matriz de transformación que permite transformar los datos



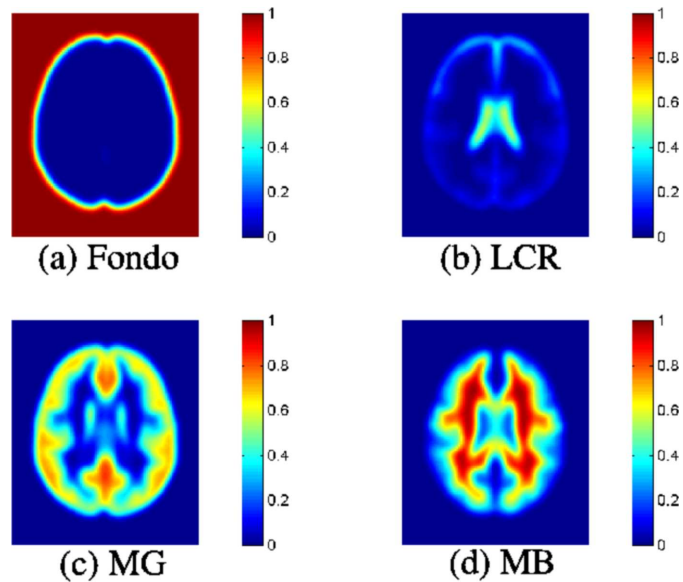


Figura 6.20: Mapas de probabilidad correspondientes a una rebanada del volumen simulado de cerebro. La imagen (a) representa la probabilidad de que la clase sea fondo, (b) que sea líquido cefalorraquídeo, (c) que se materia gris, y (d) que sea materia blanca.

de la imagen (imagen objeto) hacia el espacio de la plantilla (imagen objetivo). Así, cuando se registran los datos con la plantilla, se obtiene la matriz de transformación  $\mathbf{M}_t$ . Dadas las coordenadas del voxel en la plantilla, al multiplicarlas por la matriz  $\mathbf{M}_t$ , se obtiene como resultado las coordenadas correspondientes en la imagen de datos. Estas últimas coordenadas se usan para determinar el valor de intensidad en la imagen de datos, y asignárselo a la coordenadas correspondientes en la plantilla, transformando de esta manera la imagen de datos al espacio de la plantilla.

La transformación de interés para nuestros fines es exactamente en sentido contrario al descrito arriba, es decir, se requiere una transformación de la plantilla hacia el espacio de los datos, y así poder determinar de manera directa el valor de probabilidad en los mapas para la ubicación de un voxel en la imagen de datos. Si se multiplican las coordenadas del dato, por la inversa de la matriz de transformación  $\mathbf{M}_t$ , se obtienen las coordenadas correspondientes en los mapas de probabilidad, lográndose así la transformación en el sentido deseado. En la Figura 6.20 se muestra un ejemplo del resultado de efectuar esta transformación, quedando en posibilidad de determinar directamente el valor de probabilidad en los mapas para un dato de la imagen a clasificar, puesto que los mapas probabilísticos están ahora en el mismo espacio coordenado que los datos.

### 6.2.5. Clasificación de regiones por probabilidad de pertenencia

Las imágenes fusionadas están formadas por varias regiones homogéneas, que no corresponden necesariamente al número de clases que se desea identificar. Así, en el paso final del proceso de segmentación se toma como entrada la imagen fusionada (podada), y con la ayuda de los mapas probabilísticos, se calculan sus probabilidades posteriores, asignando cada región al tipo de tejido con mayor probabilidad de pertenencia dentro del mapa probabilístico correspondiente.

En el SPM solo están disponibles los mapas para LCR, MG y MB, a partir de estos mapas se tuvo que determinar un cuarto mapa, que se denominó como Fondo, para representar la clase de todo aquello que no se considere materia cerebral. Para calcular el mapa de Fondo se sumaron los tres mapas existentes,  $\mathbf{A} = (\text{LCR} + \text{MG} + \text{MB})$ , después se generó una matriz de las mismas dimensiones de los mapas,  $\mathbf{Fondo} = \text{tamaño}(\mathbf{A})$  y cuyo contenido fue el valor máximo de la suma. Posteriormente, el mapa de fondo se calculó, restándole a la matriz del máximo, la suma de los mapas,  $\mathbf{Fondo} = \text{máximo}(\mathbf{A}) - \mathbf{A}$ ; y finalmente, cada mapa se normalizó con respecto al máximo de la suma, asegurando que la suma de probabilidades, para un mismo voxel, en los cuatro mapas sea igual a uno. En la Figura 6.20, la imagen (a) Fondo representa el cuarto mapa calculado de la manera antes mencionada.

En el cálculo de las probabilidades de pertenencia de cada una de las regiones, se consideran todos los píxeles de la región en su conjunto, y no los píxeles de manera individual, esto es, las probabilidades se calculan por región; se hace de esta manera para no perder las regiones que han sido encontradas hasta este momento por el proceso, y únicamente se pretende asignar las regiones a la clase en la que su probabilidad sea máxima.

Para calcular las probabilidades de pertenencia, se deben conocer las prevalencias de las clases, disponibles en el conocimiento *a priori* contenido en los mapas de probabilidad. Las prevalencias,  $P(w_i)$ , para las clases de Fondo, LCR, MG y MB se calculan utilizando (5.15) considerando la proporción de cada mapa para la rebanada correspondiente. Para calcular la verosimilitud de cada clase,  $w_i$ , con respecto a los datos conformando una región,  $p(\mathbf{R}_n | w_i)$ , se toman en cuenta exclusivamente a los datos pertenecientes a la región bajo análisis, como se indica en (5.16). Una vez calculadas las probabilidades *a priori*  $P(w_i)$  para cada clase por rebanada, y las probabilidades condicionales de clase  $p(\mathbf{R}_n | w_i)$  por región, se tiene toda la información necesaria para calcular la probabilidad de pertenencia,  $P(w_i | \mathbf{R}_n)$  con (5.17), para cada clase dada

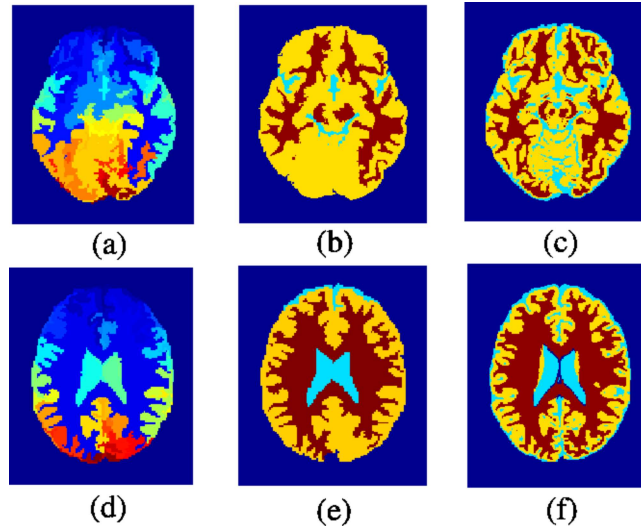


Figura 6.21: Clasificación por pertenencia. Las imágenes podadas (a) y (d), son la entrada para el cálculo de la probabilidades de pertenencia, y después de aplicar la regla de decisión se obtienen las imágenes (b) y (e), que deben ser comparadas con la segmentación de referencia (c) y (f) para obtener su similitud.

una región y aplicar la regla de decisión (5.18) para asignar la región a la clase  $w_i$  que maximiza la probabilidad de pertenencia.

En la Figura 6.21 se muestra un par de rebanadas clasificadas con la probabilidad de pertenencia. Las imágenes podadas, (a) y (d), resultantes del análisis de adyacencias, contienen las regiones que deben ser asignadas a algún tipo de tejido, por lo que se determina la probabilidad de pertenencia de cada región, asignando cada una de ellas a la clase que maximiza su probabilidad; dando como resultado las imágenes (b) y (e). La comparación visual de las imágenes clasificadas por el método, (b) y (e), con las imágenes de referencia, (c) y (f), pone de manifiesto que el método clasifica mal algunas regiones, y al calcular el índice de similitud se obtiene un valor bajo, indicando que las imágenes no son muy parecidas, lo cual es un resultado no deseado.

Para corregir los errores de la clasificación, cuando se realiza a través del cálculo de las probabilidades de pertenencia, se aplica una fase de regularización empleando una técnica de crecimiento de región semilla. La Figura 6.22 muestra los mapas probabilísticos para las clases de Fondo, LCR, MG y MB, correspondientes a la rebanada del renglón superior de la Figura 6.21, de los cuales se extraen las regiones semilla (imágenes (c), (d), (g) y (h) en la Figura 6.22) que caracterizan a cada una de las cuatro clases de interés. Las regiones semilla se seleccionaron eligiendo la región con el mayor número de píxeles en cada clase, como criterio de mayor importancia, seguido por la probabilidad

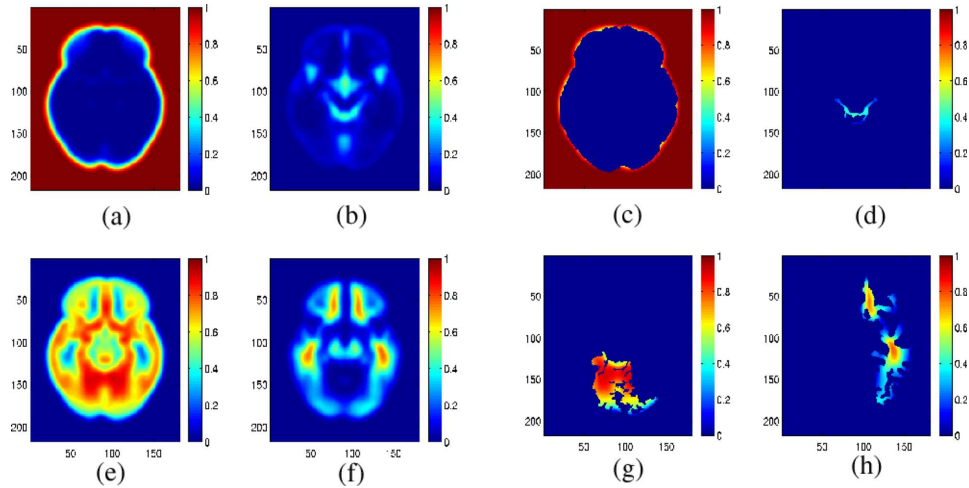


Figura 6.22: Mapas probabilísticos y semillas. (a) mapa de fondo, (b) mapa de LCR, (e) mapa de MG, (f) mapa de MB, (c) semilla para fondo, (d) semilla para LCR, (g) semilla para MG y (h) semilla para MB.

de pertenencia más alta a cada una de las clases; de esta manera, las regiones semilla son las más grandes de cada clase.

Con las regiones podadas obtenidas al final del paso tres del proceso, que corresponde al análisis de adyacencia de regiones, se genera una imagen, asignando a cada pixel de la imagen, la moda de intensidad de la región a la que pertenece, de esta manera se obtiene una imagen segmentada con las modas de intensidad, y ejemplo de ellas son las imágenes (a) y (d) de la Figura 6.23. En estas imágenes se observan con mayor facilidad las regiones que fueron mal clasificadas por el cálculo de las probabilidades de pertenencia (imágenes (b) y (e) de la Figura 6.21), y de manera global se observa que todas las regiones de interés son preservadas por el algoritmo del CM. Aun cuando las regiones no son idealmente homogéneas, sí se observa que están bien delineadas, y sólo se necesita aplicar un criterio para etiquetar adecuadamente cada región. Este criterio es el crecimiento de región semilla.

Como ya se mencionó, para obtener las regiones semilla para cada clase se toman en cuenta el tamaño de la región y su probabilidad de pertenencia, posteriormente se realiza un análisis sobre las regiones de la imagen, para unir las a la región semilla más cercana aplicando un criterio de mínima diferencia en modas de intensidad entre región y semilla. El resultado de aplicar este criterio a las imágenes (a) y (d) de la Figura 6.23 se muestra en las imágenes (b) y (e), y a simple vista son más parecidas a las imágenes de referencia (c) y (f).

Para concluir la presentación de resultados del cerebro simulado, en la Figura 6.24

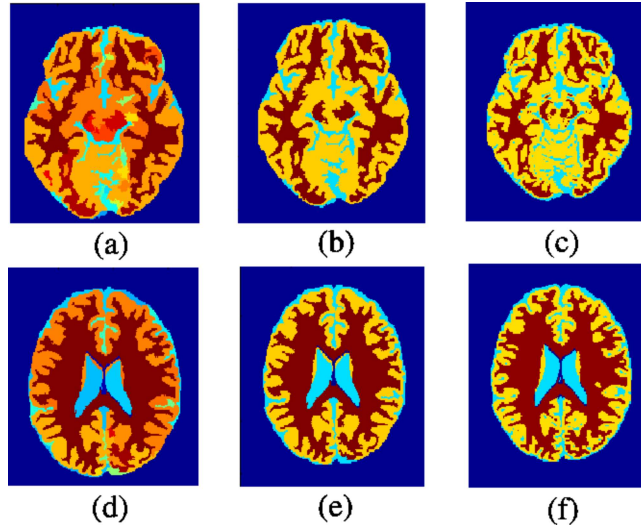


Figura 6.23: Clasificación por crecimiento de semilla. Las imágenes (a) y (d) se generan con las modas de intensidad encontradas y muestran las regiones que deben ser clasificadas, al aplicar el crecimiento de semilla se producen las imágenes (b) y (e), que tienen una similitud mayor con las imágenes segmentadas de referencia (c) y (f).

se muestran varias rebanadas en corte axial del volumen simulado del cerebro, donde en la columna (a) se despliegan los datos originales, la columna (b) son las imágenes segmentadas con las modas de intensidad, la columna (c) es la clasificación con la metodología propuesta, y la (d) las imágenes de referencia para evaluar la técnica propuesta. La evaluación de la segmentación del volumen se realiza calculando los índices de similitud, expresados como los coeficientes de Tanimoto, entre la segmentación automática y las imágenes de referencia consideradas como la “verdad de tierra”.

Determinar qué tan buena o mala es una segmentación, haciendo sólo un análisis visual, puede no ser un asunto fácil, y es altamente subjetivo porque depende del observador. Por lo tanto es conveniente tener una manera de evaluar la similitud entre imágenes sin incurrir en las desventajas antes mencionadas. Para medir la similitud, entre la imagen segmentada por la metodología propuesta y la segmentación de referencia, se determinaron los coeficientes de Tanimoto para cada rebanada. El coeficiente de Tanimoto está inspirado en la comparación de conjuntos. Si  $X$  e  $Y$  son dos conjuntos, y las cardinalidades de estos conjuntos son  $n_X$ ,  $n_Y$  y  $n_{XY}$  para  $X$ ,  $Y$  y  $XY$  respectivamente, el coeficiente de Tanimoto [26] entre dos conjuntos se define como:

$$TC_{XY}(k) = \frac{n_{X \cap Y}(k)}{n_X(k) + n_Y(k) - n_{X \cap Y}(k)} = \frac{n_{X \cap Y}(k)}{n_{X \cup Y}(k)}$$

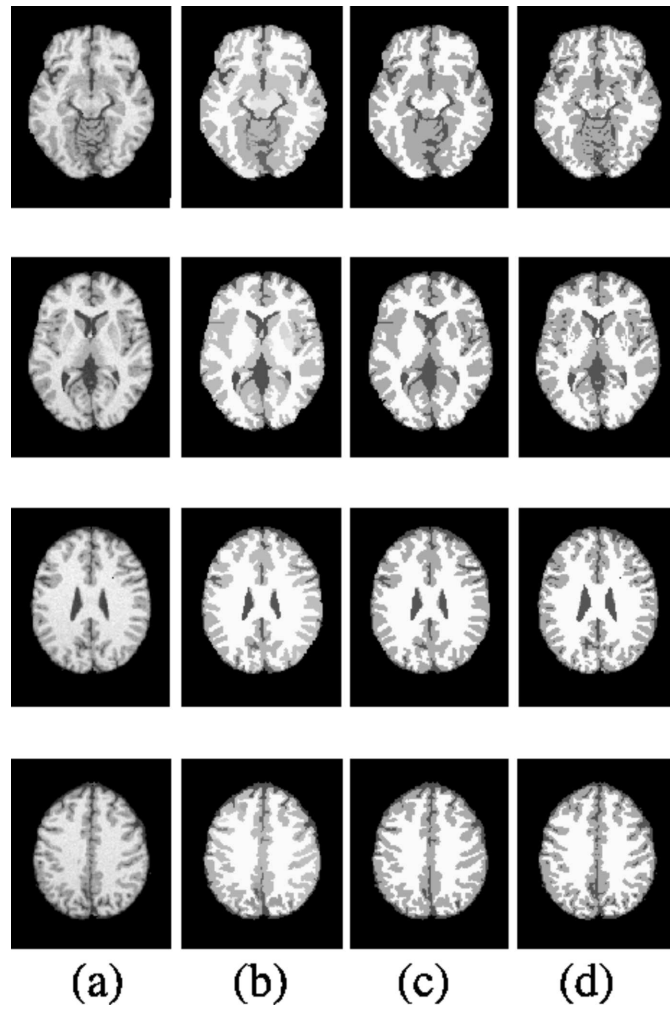


Figura 6.24: Segmentación del volumen sintético del cerebro. Rebanadas segmentadas del volumen sintético obtenido del BIC, mostrando en la columna (a) las imágenes originales, (b) las imágenes segmentadas con las modas, (c) las imágenes con las regiones clasificadas en tipos de tejido, y (d) las imágenes de referencia.

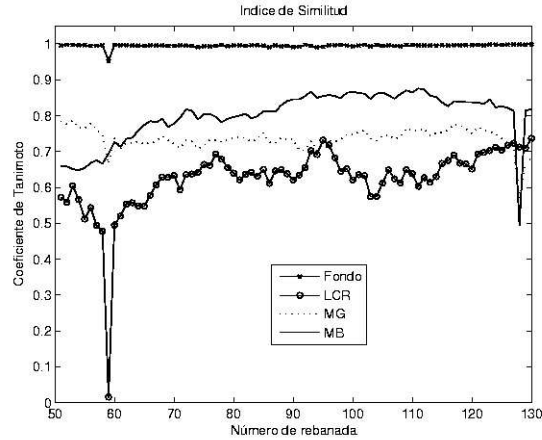


Figura 6.25: Índice de similitud entre el volumen simulado y la referencia. El índice de similitud se calcula a través de las rebanadas del volumen.

Tabla 6.3: Índices de similitud expresados como coeficientes de Tanimoto entre el cerebro simulado y la referencia.

	Fondo	LCR	MG	MB
Promedio	0.993	0.622	0.737	0.805
Desviación estándar	0.006	0.089	0.029	0.074

El coeficiente de Tanimoto entre la imagen segmentada por la aproximación propuesta y la referencia es la razón del número de elementos que se tienen en común como clase  $k$ , al número de elementos diferentes clasificados como clase  $k$ . Esta métrica se aproxima a un valor de uno para resultados que son muy similares y cercana a cero cuando son muy diferentes.

Del cerebro simulado se procesaron 80 rebanadas axiales, de la 51 a la 130, dichas rebanadas fueron elegidas porque contienen la información más representativa del tejido cerebral. La variación de los índices de similitud a través de las rebanadas procesadas del volumen se grafica en la Figura 6.25. De esta gráfica se observa que el índice de similitud se comporta mejor en las rebanadas centrales del volumen que en los extremos, porque se observa que en dos rebanadas se cae el índice, esto tal vez debido a que en los extremos la información de la imagen es menos representativa de las estructuras anatómicas de interés.

Los índices de similitud promedio entre el volumen del cerebro simulado del BIC, segmentado por la metodología propuesta, y el volumen de referencia para los datos se muestran en la Tabla 6.3.

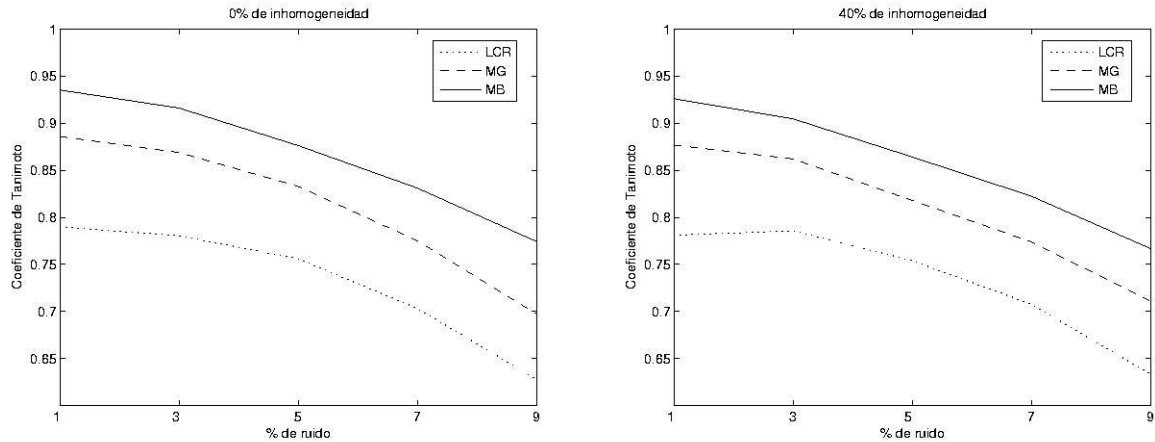


Figura 6.26: Índices de similitud de cinco volúmenes simulados. La gráfica de la izquierda no tiene inhomogeneidad en intensidad y muestra el índice de similitud para líquido cefalorraquídeo (LCR), materia gris (MG) y materia blanca (MB), contra el porcentaje de ruido en el volumen. La gráfica de la derecha muestra la misma información, sólo que ahora todos los volúmenes tienen un 40 % de inhomogeneidad.

Los índices de similitud, expresados como los coeficientes de Tanimoto, calculados para cinco volúmenes simulados con diferentes porcentajes de contaminación de ruido y manteniendo constante la inhomogeneidad de intensidad, se muestran en la Figura 6.26. En la gráfica de la izquierda de la Figura 6.26, todos los volúmenes tienen 0 % de inhomogeneidad en los niveles de intensidad de las rebanadas, mostrando la variación del índice de similitud contra el ruido. El valor en el eje de la abscisas, denota el porcentaje de ruido con el que esta contaminado cada volumen, de esta manera, el primer volumen contiene 1 % de ruido, el segundo 3 % y así sucesivamente. La gráfica muestra que el rendimiento del proceso de segmentación disminuye para los tres tipos de tejido que se segmentan: líquido cefalorraquídeo (LCR), materia gris (MG) y materia blanca (MB). En la gráfica de la derecha de la Figura 6.26, es igual a la de la izquierda, excepto que ahora todos los volúmenes tienen un 40 % de inhomogeneidad en intensidad. Comparando las dos gráficas, se observa que la variación del índice contra el ruido es similar, lo que permite inferir, que en este tipo de datos, el ruido presente en los datos tiene mayor efecto en el proceso de segmentación, que el porcentaje de inhomogeneidad.

Los coeficientes de Tanimoto no son la única métrica para evaluar la similitud entre imágenes. En el Apéndice F se describe otra métrica que se aplica a los mismos volúmenes de la Figura 6.26, y se muestran el mismo tipo de gráficas, de los resultados obtenidos con esta otra métrica, para la variación del índice de similitud.



## 6.3. Datos reales

El método de segmentación propuesto se aplicó también a datos reales, procesados de la misma manera que los datos sintéticos del cerebro simulado. Los datos reales están disponibles en el IBSR, y constan de 20 estudios de RM de cerebro del tipo T1 en orientación coronal, adquiridos con dos sistemas diferentes: un Siemens Magnetom y un General Electric, ambos de 1.5 teslas. La resolución de los estudios es de 3 mm en el eje coronal, 1 mm en el axial, y 1 mm en el sagital. Junto con los datos reales se proporcionan los resultados de la segmentación guiada manualmente por un experto. Aun cuando las segmentaciones manuales no pueden ser consideradas 100 % “verdad de tierra”, son una buena manera para comparar métodos de segmentación automatizados.

Los valores de los parámetros usados para procesar estos volúmenes se muestran en la Tabla 6.1, y se mantuvieron constantes para todas las rebanadas de los 20 estudios; debe notarse que son los mismos parámetros usados para procesar los datos del cerebro sintético.

### 6.3.1. Filtrado por CM

La determinación de los mapas de confianza y el filtrado de los datos son los primeros pasos en el proceso de segmentación. Los resultados de estos primeros pasos se muestran en la Figura 6.27, donde se muestran cinco rebanadas de uno de los volúmenes a segmentar, desplegadas en una ventana enmarcando sólo la masa cerebral, para mejorar la visualización. Las imágenes Org 1 a 5 son los datos originales, correspondiendo respectivamente a las rebanadas 7, 19, 25, 38 y 47 del volumen 112\_2. Cabe aclarar, que aun cuando sólo se muestra una ventana de cada rebanada, en el proceso de segmentación se utilizaron el total de los datos.

Lo que se pretende en el filtrado es conservar los bordes y homogeneizar los píxeles con variaciones dentro del ancho de banda de intensidad, para un vecindario dentro del ancho de banda espacial. Aplicando el algoritmo de CM, ponderado con los mapas de confianza, a los datos originales, se obtienen las imágenes filtradas Filt 1 a 5, en las cuales se observan regiones más homogéneas en el nivel de gris de las estructuras anatómicas que conforman al cerebro.

### 6.3.2. Análisis de adyacencias

Para refinar la delineación de regiones en el resultado del filtrado, se efectúa el análisis de adyacencia entre regiones. Como puede observarse en la Tabla 6.4, el número de

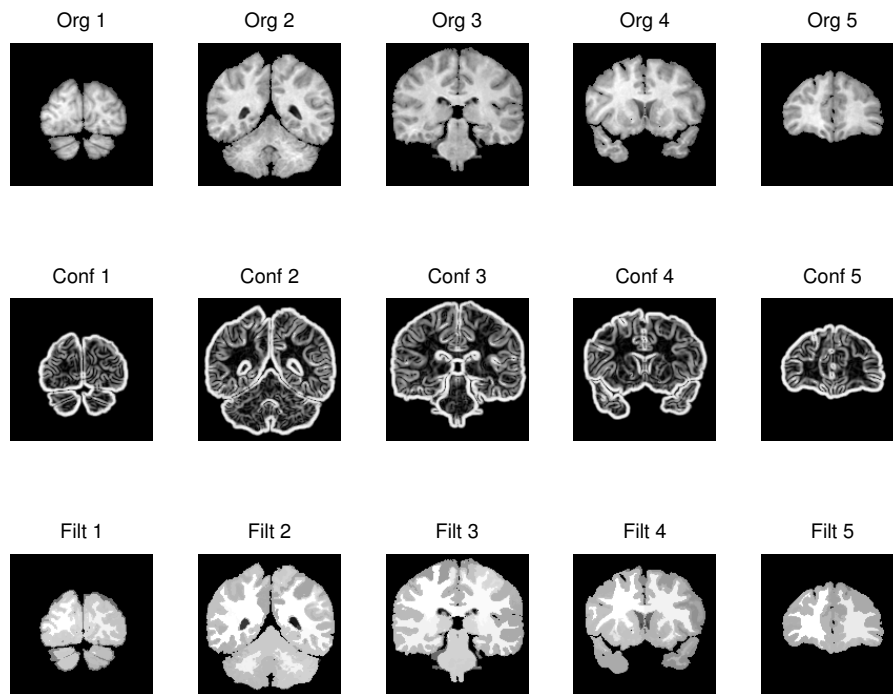


Figura 6.27: Filtrado de datos reales. El primer renglón corresponde a los datos originales, Org 1 a 5, seguido por sus respectivos mapas de confianza, Conf 1 a 5, y las imágenes filtradas, Filt 1 a 5, obtenidas a partir de las originales y los mapas.

Tabla 6.4: Número de regiones en datos reales para las rebanadas de la Figura 6.27.

Rebanada	Etiquetado	Fusión	Poda
7	1011	358	25
19	1522	575	29
25	1061	504	30
38	1385	482	31
47	452	172	10

regiones producido por la etapa previa de filtrado, reportado en la columna Etiquetado, es alto e impráctico para los fines que se pretenden en el método de segmentación investigado, que es identificar solamente las estructuras anatómicas del cerebro para LCR, MG y MB. En la misma tabla se aprecia la disminución del número de regiones, después de aplicar la fusión y la poda, alcanzándose el objetivo del análisis de adyacencias que es la disminución del número de regiones homogéneas.

El progreso en la fusión de regiones por el análisis de adyacencia se muestra en la Figura 6.28. En esta figura se muestra en el primer renglón, las rebanadas filtradas una vez que se les ha asignado una etiqueta a cada región para poder construir la MAR, sobre la que se realiza el análisis de adyacencia. El segundo renglón de la figura muestra las rebanadas CT 1 a 5, que son obtenidas después de aplicarles las operaciones de fusión, observándose una disminución importante en el número de regiones si se compara con las rebanadas etiquetadas. En las imágenes CT se aprecian, con cierta dificultad, algunas regiones pequeñas como si fueran puntos, las cuales son unidas con su región candidata por la etapa de podado como se muestra en las imágenes Poda 1 a 5. El número de regiones, en cada una de las rebanadas del volumen, se reduce sustancialmente con el análisis de adyacencia de las mismas.

### 6.3.3. Registro de datos y mapas de probabilidad

El registro de las imágenes de cerebros normales con una imagen plantilla se hizo de la misma manera que con el cerebro simulado. La Figura 6.29 muestra el resultado de esta normalización espacial, donde las tres imágenes de la izquierda corresponden a la plantilla, y las de la derecha a datos reales.

En la Figura 6.30 se muestra un ejemplo de los mapas probabilísticos obtenidos con el registro de datos reales. De estos mapas se obtiene la probabilidad de que un pixel en la imagen pertenezca a Fondo, LCR, MG y MB.

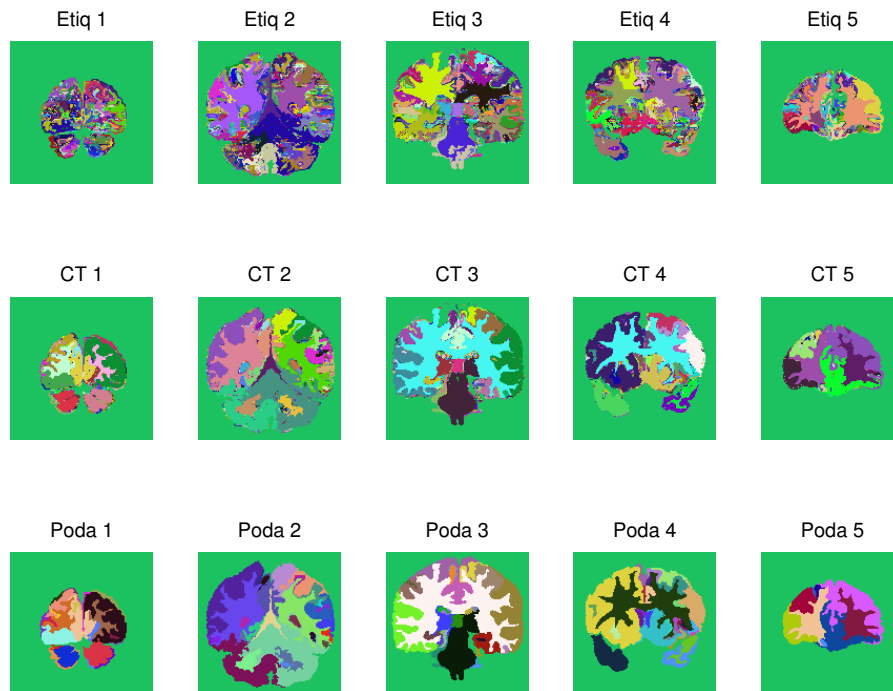


Figura 6.28: Fusión de regiones en datos reales. Las imágenes Etiqu 1 a 5 son el resultado de etiquetar las imágenes filtradas. Las imágenes CT1 a 5 representan las regiones obtenidas por la etapa de fusión, y Poda 1 a 5 son las imágenes podadas.

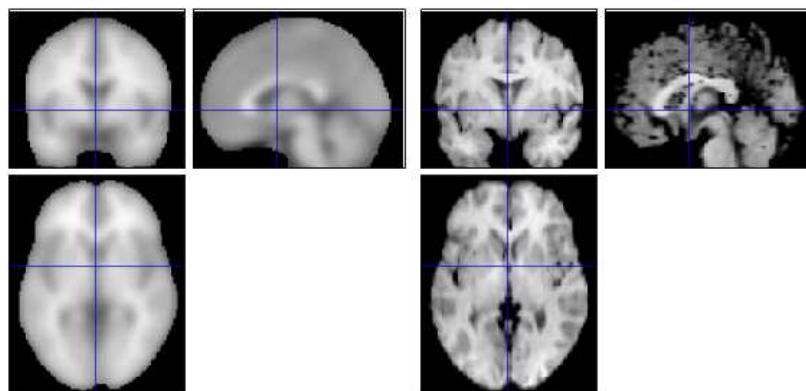


Figura 6.29: Normalización espacial entre datos reales y la plantilla.

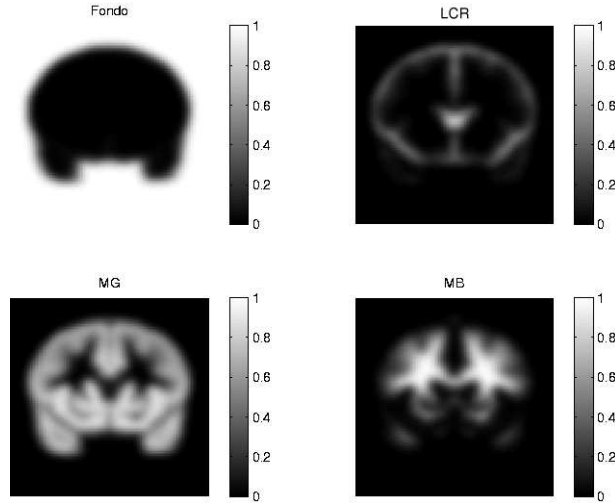


Figura 6.30: Mapas de probabilidad para una rebanada de datos reales. La imagen Fondo representa la probabilidad de la clase para fondo, LCR para líquido cefalorraquídeo, MG para materia gris, y MB para materia blanca.

#### 6.3.4. Clasificación de regiones

El resultado final del proceso de segmentación se muestra en la Figura 6.31; la primera columna presenta 5 rebanadas que corresponden a los datos originales que se procesaron, seguida por la segmentación obtenida con la metodología propuesta en la segunda columna, que se puede comparar con la segmentación manual realizada por el experto, mostrada en la tercer columna.

De los 20 volúmenes cerebrales de RM proporcionados por el IBSR, también están disponibles sus respectivas segmentaciones manuales para propósitos de comparación. El IBSR usa los coeficientes de Tanimoto para medir la similitud entre la segmentación manual y los resultados de métodos automáticos; en el mismo sitio están reportados varios métodos de segmentación y los índices de similitud que se obtienen al segmentar los 20 volúmenes de los archivos de datos que contienen sólo información cerebral. Estos volúmenes fueron elegidos porque ya han sido usados en estudios publicados [51][66] y tienen varios niveles de dificultad; los peores tienen bajo contraste y gradientes de intensidad relativamente grandes.

Las gráficas de las Figuras 6.32 y 6.33 muestran los resultados de similitud para MB y MG de cada uno de los estudios de cerebro reales. En estas gráficas se han incorporado los resultados del algoritmo de CM, para mostrar su ubicación al lado de los métodos reportados en el IBSR. Los estudios cerebrales están graficados en orden de dificultad para segmentarlos, siendo el más difícil el número 1 y el menos difícil el número 20.

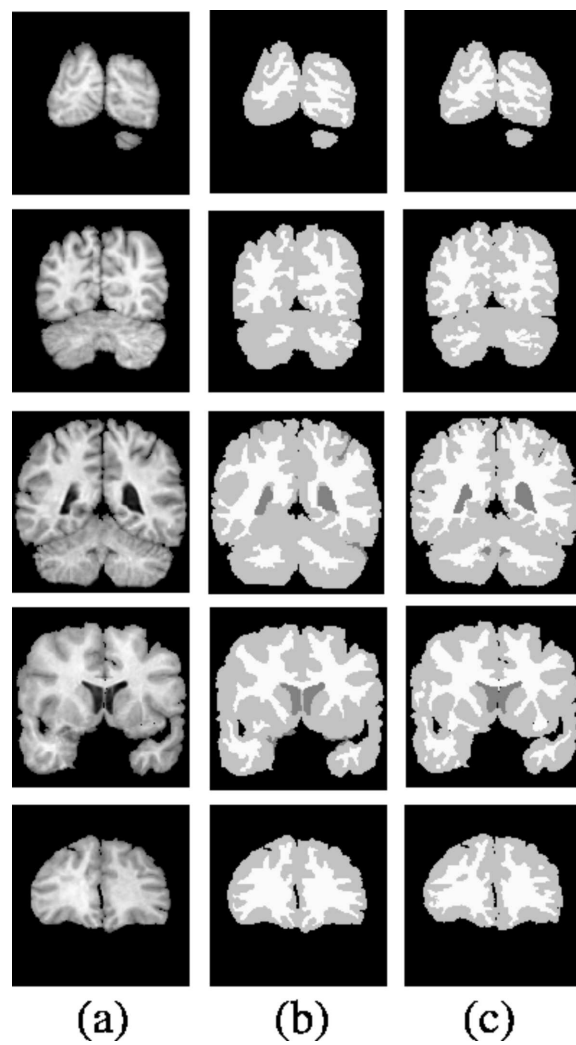


Figura 6.31: Ejemplos de segmentación de imágenes reales. Rebanadas de diferentes volúmenes mostrando en la columna (a) las imágenes originales, en (b) los resultados de la segmentación por CM, y en (c) las clasificaciones manuales por el experto.

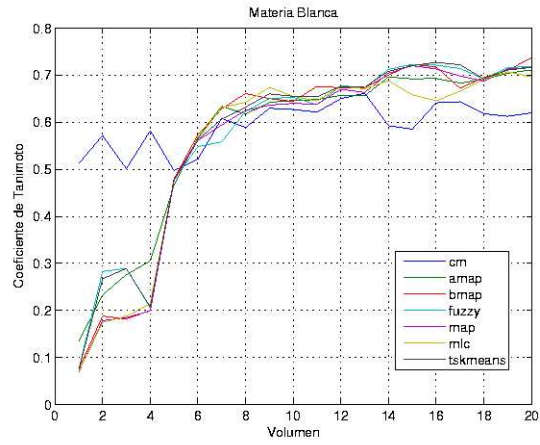


Figura 6.32: Indices de similitud para materia blanca para cada uno de los volúmenes.

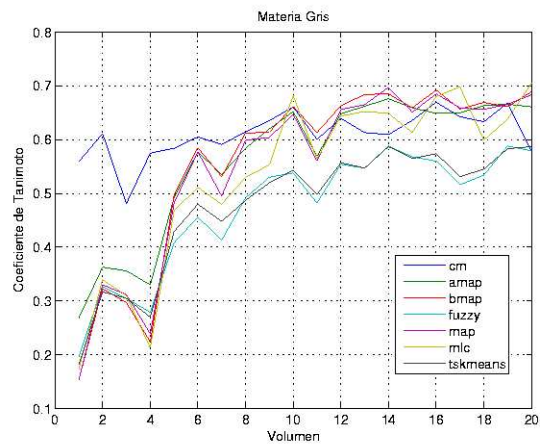


Figura 6.33: Indices de similitud para materia gris para cada uno de los volúmenes.

Tabla 6.5: Índice de similitud expresado como coeficientes de Tanimoto entre las imágenes segmentadas y la segmentación manual. Adaptada de [39]

Método	Fondo	LCR	MG	MB
Adaptive MAP	0.999	0.069	0.564	0.567
Biased MAP	0.999	0.071	0.558	0.562
Fuzzy c-means	0.999	0.048	0.473	0.567
MAP	0.999	0.071	0.550	0.554
Maximum-Likelihood	0.999	0.062	0.535	0.551
Tree-structure k-means	0.999	0.049	0.477	0.571
<b>Corrimiento de Media</b>	<b>0.997</b>	<b>0.318</b>	<b>0.610</b>	<b>0.594</b>

Tabla 6.6: Valores estadísticos para la clasificación de los 20 volúmenes reales

Estadística	Fondo	LCR	MG	MB
Promedio	0.997	0.318	0.610	0.594
Desv. estándar	0.004	0.107	0.044	0.050
Varianza	0	0.012	0.002	0.003

Los resultados del CM están por arriba de los otros métodos en el caso de los datos más difíciles de segmentar, indicando que es menos sensible a las inhomogeneidades y su rendimiento se mantiene relativamente constante a través de todos los volúmenes, aunque el rendimiento es similar a los otros métodos para MG y disminuye para MB, en el caso de los datos menos difíciles.

El rendimiento promedio, expresado como medida de similitud, obtenido por la aproximación propuesta, en la segmentación de los 20 volúmenes, y los resultados reportados en el IBSR para seis aproximaciones aplicadas a la segmentación de los mismos datos, se muestran en la Tabla 6.5. El último renglón presenta la cuantificación de los resultados obtenidos con el CM y puede observarse que tiene un rendimiento superior al de los otros métodos para las tres estructuras de interés segmentadas.

Adicionalmente a los promedios de la medida de similitud para todo el volumen, presentados en la Tabla 6.5, también se muestran los valores de desviación estándar y varianza en la Tabla 6.6 para el método de clasificación aquí presentado.

El algoritmo se ha implementado en una PC de escritorio con procesador Pentium III a 1 GHz, utilizando Matlab como herramienta de programación; con esta plataforma, la segmentación de una imagen de  $256 \times 256$  se tarda aproximadamente 1.1 minutos.



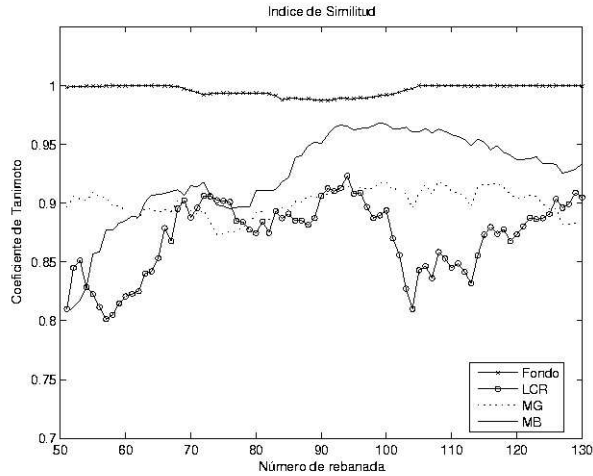


Figura 6.34: Índices de similitud de las rebanadas del cerebro sintético procesadas de manera volumétrica. La comparación se hace entre el cerebro sintético y el cerebro de referencia correspondiente.

## 6.4. Segmentación volumétrica

Todos los resultados presentados previamente, tanto sintéticos como reales, fueron procesados rebanada por rebanada, usualmente esta es la manera en que un experto realiza la segmentación. La delineación de las regiones se efectúa en el dominio conjunto espacio-intensidad, donde se emplean las coordenadas  $XY$  para ubicar el pixel en la cuadrícula de la imagen y la intensidad del pixel, dando como resultado un dominio conjunto de tres dimensiones. La metodología fue presentada para variables  $d$ -dimensionales, por lo que llevar el proceso de los datos a un dominio conjunto de mayor dimensión es directo e inmediato y no debe presentar problema alguno.

Si se requiere procesar los datos de manera volumétrica, sólo se debe aumentar una dimensión en el dominio espacial, para ubicar el voxel en el conjunto de datos y seguir considerando únicamente la característica de intensidad, teniendo ahora un dominio conjunto de cuatro dimensiones.

Considerando el dominio conjunto de tres dimensiones espaciales:  $X$ ,  $Y$  y  $Z$ ; y una dimensión de rango: la intensidad, se implementó todo el procedimiento de segmentación en este nuevo dominio de cuatro dimensiones, para procesar tanto los datos sintéticos como reales del cerebro. A continuación se presentan los resultados que se obtuvieron al segmentar de manera volumétrica.

Al proceso de los datos considerando una dimensión espacial adicional, le llamaremos proceso en 3D por referencia a las dimensiones espaciales. En el proceso 3D de

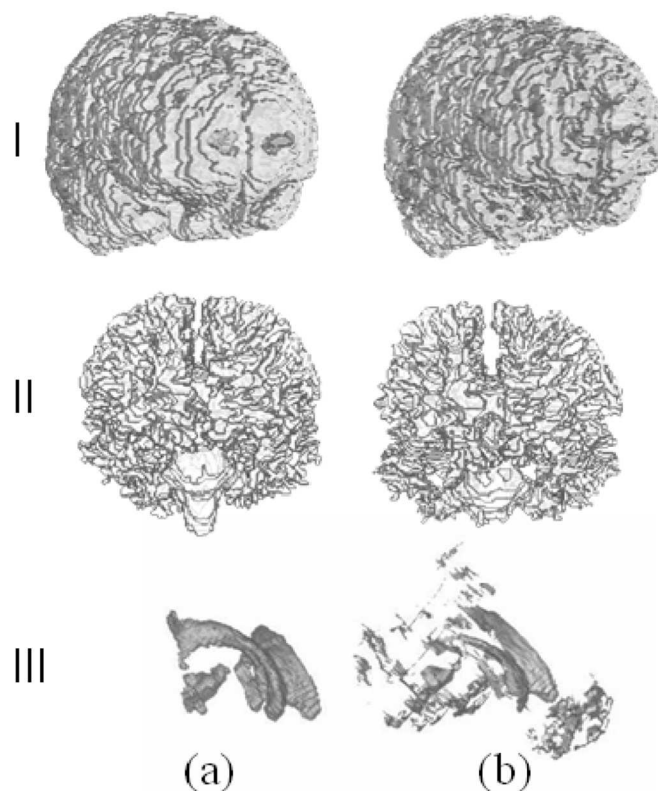


Figura 6.35: Reconstrucciones volumétricas de materia gris (I), materia blanca (II) y líquido cefalorraquídeo (III). La columna (a) representa el volumen de referencia y la (b) el volumen segmentado automáticamente.

los datos del cerebro simulado se empleó el mismo conjunto de parámetros que en el caso 2D, excepto en el parámetro de tamaño mínimo de región que se tuvo que incrementar de 30 a 90, debido a que el número de voxels que conforman ahora una región volumétrica es mayor. La Figura 6.34, muestra la variación del índice de similitud a lo largo de las rebanadas del conjunto de datos del cerebro sintético, para las cuatro clases de interés que se segmentaron. Comparando con la misma gráfica para el caso 2D (Figura 6.25), se observa que la caída del índice en dos rebanadas de los extremos ha desaparecido, indicando que la información adicional de la otra dimensión espacial ayuda a resolver errores eventuales en la clasificación, produciendo una delineación más precisa del volumen.

En cuanto a los datos reales, éstos también fueron procesados en 3D. La Figura 6.35 muestra la reconstrucción en 3D de uno de los 20 volúmenes, tanto de la segmentación manual del experto, columna (a), como de los resultados obtenidos, columna (b), al aplicar la metodología propuesta basada en el algoritmo del CM; las estructuras segmentadas son la materia gris (I), la materia blanca (II) y el líquido cefalorraquídeo

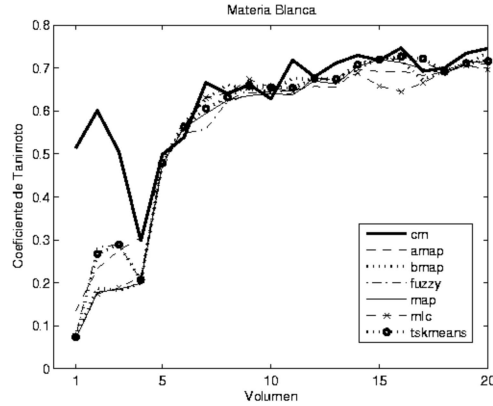


Figura 6.36: Índice de similitud para materia blanca en cada uno de los 20 volúmenes procesados de manera volumétrica. Se muestran los resultados de los seis métodos reportados en IBSR y el procedimiento de CM propuesto en la línea “cm”.

Tabla 6.7: Índice de similitud entre las imágenes segmentadas y la segmentación de referencia para 2D y 3D.

Datos / Tipo de proceso	Fondo	LCR	MG	MB
Cerebro simulado / 2D	0.993	0.622	0.737	0.805
Cerebro simulado / 3D	0.996	0.871	0.900	0.924
20 estudios reales / 2D	0.997	0.318	0.610	0.594
20 estudios reales / 3D	0.996	0.210	0.594	0.628

(III).

Para medir el rendimiento del proceso en 3D se calculó el coeficiente de Tanimoto como medida de similitud promedio en los 20 volúmenes. Para calcular esta medida ahora se consideraron volúmenes en lugar de regiones, así para cada conjunto de datos correspondiente a un estudio, se reconstruyeron los volúmenes para MG, MB y LCR, tanto del volumen segmentado por el método, como de la segmentación manual y se calculó el coeficiente para cada volumen. Los resultados de este cálculo se muestran en la Figura 6.36 para materia blanca y en la Figura 6.37 para materia gris. En ambos casos se observa una mejora del índice en 3D comparado con el 2D.

Los resultados para las imágenes cerebrales tanto sintéticas como reales se resumen en la Tabla 6.7, para el procesamiento en 2D y 3D. Para los datos sintéticos se observa un incremento en los índices de LCR, MG y MB, sin embargo, para los datos reales no se observa una mejora global en los índices, ya que el LCR y MG disminuyeron y MB aumentó.

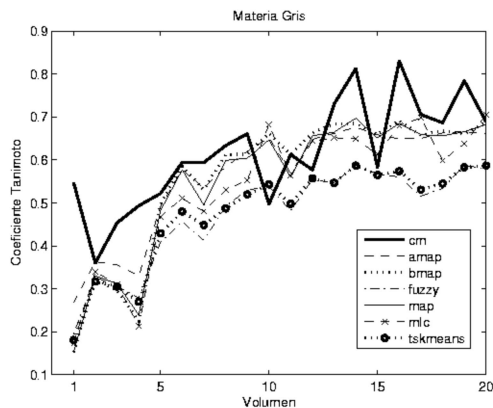


Figura 6.37: Índice de similitud para materia gris en cada uno de los 20 volúmenes procesados de manera volumétrica. Se muestran los resultados de los seis métodos reportados en IBSR y el procedimiento de CM puesto en la línea “cm”.

El parámetro que tiene un efecto directo en el tiempo de ejecución del algoritmo es el radio espacial, al aumentar una dimensión espacial se aumenta el vecindario de análisis del proceso, lo que es equivalente a tener un radio espacial mayor en 2D; por esta razón el tiempo de proceso se multiplicó por 13, esto es, procesar una rebanada en 2D toma aproximadamente 1 minuto, mientras que en 3D toma 13 minutos. Lo que se gana en 3D tal vez no es suficiente para compensar el aumento en el tiempo de proceso.

Las Figuras 6.31 y 6.35 proporcionan una comparación cualitativa entre las segmentaciones obtenidas con el método propuesto y el estudio de referencia, en vistas de 2D y 3D, respectivamente. Las similitudes entre estructuras segmentadas se pueden apreciar claramente para materia blanca y gris. En el caso de las reconstrucciones 3D de LCR, se ve que el algoritmo de CM es capaz de detectar tanto el LCR intra como extra ventricular, que no incluyen los estudios de referencia. Mientras que en una primera impresión esta condición parece como ruido de la segmentación, es en verdad una descripción más completa de la clase correspondiente. Por lo que se refiere a los resultados cuantitativos, el rendimiento del método investigado se aprecia claramente para las simulaciones de cerebro, donde los índices de Tanimoto de 0.87 y mayores fueron obtenidos para todas las estructuras. Los resultados para datos reales muestran un mejor rendimiento comparado con los métodos reportados en [39] (ver Tabla 6.5), mientras que otros autores [47, 66] han obtenido índices más altos para los mismos casos de imágenes. Sin embargo, la falta de sensibilidad de la medición de rendimiento para describir cambios locales hace difícil argumentar acerca de la superioridad total de un método con respecto a otro. En esta perspectiva, las comparaciones de método son más relevantes en el dominio de

las suposiciones del modelo, preprocesamiento requerido, o carga computacional.

# Capítulo 7

## Conclusiones

Muchas de las técnicas de segmentación de cerebro requieren apearse a diversas suposiciones, tales como una distribución de densidad de probabilidad específica [61, 47, 8, 66, 60], homogeneidad de nivel de gris lineal por tramos [67], o un número dado de clases de tejido presente en la imagen [35, 61, 47, 60]. Además, usualmente hay requerimientos para procedimientos de inicialización [8, 66, 60, 67], o preprocesamiento de imagen para corrección de inhomogeneidad [35, 62, 45, 66, 67], todo realizado antes de la segmentación real de la imagen. Por otra parte, el análisis del espacio característico para la segmentación de imagen, tal como se propone en este trabajo, está exclusivamente guiado por los datos cuando se aplica el algoritmo de corrimiento de media, evitando los puntos débiles mencionados anteriormente (suposiciones estadísticas, inicialización o preprocesamiento). La segmentación por corrimiento de media ha encontrado su aplicación en reconocimiento y seguimiento de objetos, y en segmentación de imagen de color y compresión. Hasta donde sabemos, el planteamiento propuesto en este trabajo es la primera aplicación de tan poderosa técnica para la segmentación de imágenes de cerebro.

El algoritmo de CM ponderado por el mapa de confianza de borde ha mostrado un buen desempeño en la búsqueda y localización de las modas de una *fdp* desconocida, permitiendo la delineación de regiones homogéneas de una IRM cerebral. Adicionalmente, la estrategia de clasificación propuesta permite la identificación del tipo de tejido a que corresponde cada una de las regiones. La combinación de mapas de confianza de borde y segmentación de regiones homogéneas, mostró un buen desempeño en la presencia de ruido e inhomogeneidad de intensidad, conservando las fronteras entre regiones, y probando ser robusto en la separación de clases de la imagen. En la estrategia propuesta, la identificación de estructuras neuroanatómicas depende del número de mapas probabilísticos disponibles.

El CM es un estimador no paramétrico sencillo, y dado que es un procedimiento manejado por los datos, extrae las características de una imagen sin necesidad de suposiciones acerca de la distribución de los datos, y sin requerir la intervención de un usuario para dar información previa o de inicialización.

Aun cuando el procedimiento no requiere de una etapa de inicialización, los parámetros de proceso pudieran tener que ajustarse antes de la segmentación. Una vez ajustados, permanecen constantes a lo largo de todo el proceso de segmentación. Esta tarea fue realizada estimando el conjunto de parámetros que mejor segmentaron una imagen de prueba de cerebro sintético, considerando el índice de Tanimoto como el criterio de evaluación. Estos valores fueron también usados con éxito para las imágenes reales, indicando de esta manera que se puede determinar un conjunto de parámetros por defecto.

Los parámetros que tienen mayor influencia en el desempeño del algoritmo son los radios del kernel. La resolución espacial,  $h_e$ , tiene un efecto mayormente en el tiempo de ejecución: a mayor radio espacial, mayor el tiempo de procesamiento, debido al mayor vecindario que se tiene que considerar para el análisis de un pixel. Por otro lado, la calidad de la segmentación depende considerablemente de la resolución en intensidad: un  $h_i$  pequeño produce muchas regiones, sobresegmentando así la imagen, mientras que un radio de intensidad mayor puede producir imágenes subsegmentadas (pocas regiones). Solamente las características de imagen con soporte espacial grande son representadas en la imagen filtrada cuando  $h_e$  aumenta, y sólo las estructuras con valores de intensidad muy diferentes permanecen independientes cuando  $h_i$  es grande.

Un valor de umbral de borde cercano a cero o uno produce, respectivamente, sobre o subsegmentación; mientras que un factor de mezcla pequeño (relevancia de la confianza de borde mayor que la de magnitud de gradiente) conserva los bordes débiles y resuelve los efectos de inhomogeneidad. El umbral de fuerza de borde se calcula del mapa de confianza, e influye en el paso de fusión de región del análisis de adyacencia, sobresegmentando cuando su valor es cercano a cero, resaltando los bordes. Por otro lado, cuando tiende a uno, produce subsegmentación porque considera como débiles todos los bordes.

La técnica de segmentación propuesta, extrae los bordes que son significativos de las imágenes de cerebro de RM, utilizando para ello la medida de confianza. Esta medida es indicativa de si realmente se está ante la presencia de un borde. La aproximación propuesta, basada en la combinación de mapas de confianza de bordes y la segmentación de regiones homogéneas, al aplicarse a datos tanto sintéticos como reales manifestó un buen desempeño, aun en la presencia de ruido e inhomogeneidades de intensidad,

preservando las fronteras entre regiones y probando ser robusta en la separación de las clases de los modelos de imagen. Lo anterior se evidenció, al separar las clases del fantoma esférico, en aquellos datos donde las intensidades están totalmente traslapadas; es importante también resaltar que las diferencias más notorias en los índices de similitud para los datos reales, fueron manifestadas en los conjuntos de datos más difíciles, esto es, donde los otros métodos tienen un rendimiento bajo, el algoritmo de CM los supera de manera notable.

En resumen, los parámetros que tienen una mayor influencia en el desempeño del algoritmo son los radios del kernel ( $h_i, h_e$ ) y el factor de mezcla ( $\beta$ ). Hasta este momento, la implementación presentada no corre en un modo totalmente automático, debido a que el paso de calibración inicial puede ser necesario para el ajuste de parámetros, si no se desea usar el conjunto por defecto. En trabajo futuro, se puede proponer un esquema de radios adaptables.

La identificación de las estructuras neuroanatómicas presentes en los datos reales, está en función de los mapas probabilísticos disponibles, y hasta este momento el procedimiento es capaz de clasificar en cuatro las regiones de una imagen de RM de cerebro. Comparando la segmentación alcanzada con la aproximación propuesta, con otras reportadas en el sitio del IBSR, se observó un rendimiento superior en tres de las cuatro estructuras neuroanatómicas.

El uso de un atlas digital, permite que la técnica no paramétrica propuesta, realice la etapa de clasificación final sin la necesidad de intervención del usuario, siendo así una herramienta alternativa para el radiólogo, ahorrándole tiempo y proporcionándole información que depende sólo de los datos de la imagen.

La velocidad de ejecución del proceso de segmentación completo es de aproximadamente un minuto en 2D y de 13 minutos en 3D por rebanada, siendo el algoritmo del CM, la parte más demandante en tiempo de ejecución. La implementación del proceso de segmentación está hecha totalmente en Matlab, el cual siendo un lenguaje interpretado, tiene la desventaja de ser lento en la ejecución. La alternativa para mejorar la velocidad de ejecución sería tener una implementación compilada, por ejemplo, en lenguaje C o aun en el mismo Matlab.

Esta técnica de segmentación se puede continuar investigando, por ejemplo para obtener de manera también guiada por los datos, el ancho de banda de la estimación por kernel, de manera tal que el ancho de banda se vaya ajustando automáticamente con la variación de los datos, buscando que sea mayor cuando los datos estén muy dispersos, y menor cuando sean densos. En el dominio de la imagen, el ancho de banda debe ajustarse de acuerdo a la variación del espacio de características de interés. La



selección de un ancho de banda manejado por los datos y para datos multivariados es un problema complejo, siendo un tema abierto y aun no respondido por las técnicas actuales.

Se debe seguir investigando en busca de la metodología más adecuada para realizar la clasificación usando el conocimiento a priori, se debe elevar la medida de similitud buscando aproximarla lo más posible a los índices de referencia considerados como “verdad de tierra”. Tal vez se pueda mejorar agregando una etapa de regularización utilizando, por ejemplo, un método de relajación probabilística.

Para concluir, las contribuciones principales de la metodología propuesta en este trabajo son: (i) la aplicación es totalmente guiada por los datos contenidos en la imagen, usando un conjunto de parámetros fijo que no está ajustado individualmente para cada caso, (ii) la estimación por CM es robusta aun en la presencia de inhomogeneidad de intensidad severa, sin la necesidad de una estimación de inhomogeneidad previa o simultánea y procedimiento de corrección, (iii) el uso de un mapa de confianza de borde conserva adecuadamente las fronteras de tejido, aun para bordes débiles debido a los efectos de volumen parcial, (iv) la unión de regiones, que cumplen criterios de magnitud de borde y similitud en intensidad, se realiza por la reducción de un grafo, (v) se aplica y valora el CM en la segmentación de imágenes cerebrales, (vi) la técnica de segmentación es extensible a más de una característica del espacio de la imagen. Más allá de la selección inicial de parámetros, el planteamiento es totalmente automático siendo por lo tanto una herramienta alternativa para los radiólogos, ahorrando tiempo y proporcionando información que depende solamente de los datos de la imagen.

# Apéndice A

## Trabajos publicados derivados de la investigación doctoral

1. J. R. Jiménez, V. Medina, O. Yáñez; “Nonparametric Density Gradient Estimation for Segmentation of Cerebral MRI”, *Second Joint EMBS/BMES Conference*, Houston, 2002, pp. 1076-1077.
2. Juan R. Jiménez, Verónica Medina, Oscar Yáñez; “Nonparametric MRI segmentation using mean shift and edge confidence maps”, *Medical Imaging 2003: Image Processing*, Milan Sonka, J. Michael Fitzpatrick, Editors, Proc. of SPIE, San Diego, 2003, vol. 5032, pp. 1433-1441
3. R. Valdes-Cristerna, J.R. Jimenez, O. Yanez-Suarez, J.F. Lerallut, V. Medina; “Texture-based Echocardiographic Segmentation Using a Non-parametric Estimator and an Active Contour Model”, *Proc. 26th. Annu. Int. Conf. IEEE EMBS*, San Francisco, 2004, 1806-1809.
4. J. R. Jiménez-Alaniz, V. Medina-Bañuelos, and O. Yáñez-Suárez; “Data-Driven Brain MRI Segmentation Supported on Edge Confidence and *A Priori* Tissue Information”, *IEEE Trans. Medical Imag.*, vol. 25, núm. 1, pp. 74-83, January, 2006.

# Apéndice B

## Mapa de confianza

Cálculo del mapa de confianza para la presencia de un borde en los datos, procesados con el operador gradiente, determinada como una combinación lineal de la magnitud del gradiente y la medida de confianza.

***MapaConfianza*** = `mapaconf (Imagen, m, FactorMezcla, UmbralBorde)`

**Entradas:**

- *Imagen*: Datos a procesar
- *m*: Tamaño de la ventana del filtro para derivar la imagen en X e Y
- *FactorMezcla* - ( $\beta$ ): Valor entre 0 y 1 para controlar la mezcla del gradiente y la medida de confianza en el mapa
- *UmbralBorde* - ( $\delta$ ): Umbral de la magnitud del gradiente, con valor entre 0 y 1, para que exista un borde

**Salida:**

- *MapaConfianza* - ( $\varphi$ ): Mapa de confianza de la imagen con valores entre 0 y 1

1. Calcular filtro  $h(i; r, p)$  para  $m = 2$

a) Secuencia de suavizado  $s(k) = h(k; 0, 0) = \omega(k)$

b) Secuencia de diferenciación  $d(k) = h(k; 1, 1) = (2k/m)\omega(k)$

2. Calcular gradiente de la imagen por diferenciación en una dimensión, seguida por suavizado en la otra dimensión

- a) Derivando renglón:  $derivada(i, j) = Imagen(i, j + k) * d(k)$  para  $k = -m, \dots, m$
- b) Suavizando columna:  $gradX(i, j) = derivada(i + k, j) * s(k)$  para  $k = -m, \dots, m$
- c) Derivando columna:  $derivada(i, j) = Imagen(i + k, j) * d(k)$  para  $k = -m, \dots, m$
- d) Suavizando renglón:  $gradY(i, j) = derivada(i, j+k) * s(k)$  para  $k = -m, \dots, m$

3. Determinar la función de distribución acumulativa empírica con los rangos normalizados de la magnitud del gradiente

- a) Calcular la magnitud del gradiente:  $fuerza = \sqrt{gradX^2 + gradY^2}$
- b) Ordenar de menor a mayor las magnitudes del gradiente
- c) Determina el número de valores distintos de magnitud del gradiente, es decir, los rangos
- d) Calcula los percentiles de la distribución acumulativa, dividiendo cada rango por el número total de rangos

4. Calcula la medida de confianza del borde

- a) Determina la estimación de la orientación del borde en grados.

$$orientacion_{ij} = -\tan^{-1}(gradX/gradY) * 180/\pi$$

- a) Determina la plantilla de borde ideal para cada ángulo entre  $-180^\circ$  y  $180^\circ$  con resolución de un grado
- b) Calcula la medida de confianza para cada dato
- c) Si el rango normalizado del dato  $(\rho_{ij}) > UmbraI\Bor\de$  entonces calcula la medida de confianza
  - 1) Extrae una ventana de datos centrada en el dato  $\mathbf{A}_{ij} = imagen(i-m : i+m, j-m : j+m)$ , la cual se manipula para que quede con media cero y norma de Frobenius unitaria
  - 2) Se extrae la plantilla de borde ideal  $\mathbf{t}_{ij} = plantilla(orientacion_{ij})$

3) Se calcula la medida de confianza  $\eta_{ij} = |\mathbf{t}_{ij}^\top \mathbf{A}_{ij}|$

d) En otro caso la medida de confianza es cero  $\eta_{ij} = 0$

5. Calcula el mapa de confianza como una combinación lineal, controlada por el factor de mezcla, del gradiente y la medida de confianza,

$$\text{MapaConfianza} = \text{FactorMezcla} * \rho + (1 - \text{FactorMezcla}) * \eta$$

# Apéndice C

## Filtrado por corrimiento de media

Aplicación del algoritmo de Corrimiento de Media (CM) a los datos de la imagen en el dominio conjunto espacial-rango. Se incorpora la información de los mapas de confianza como medida de peso que afecta al algoritmo de CM. La finalidad del CM es encontrar las modas a las cuales convergen los datos, aplicando el algoritmo iterativamente hasta que el dato alcance su moda o se cumpla un número fijo de iteraciones.

***CorrimientoMedia = corrmed (Imagen, RadioEspacial, RadioIntensidad, MapaConfianza)***

**Entrada:**

- *Imagen*: Datos de la imagen a procesar
- *RadioEspacial* - ( $h_e$ ): Ancho de la ventana del kernel en el dominio espacial
- *RadioIntensidad* - ( $h_i$ ): Ancho de la ventana del kernel en el dominio del rango
- *MapaConfianza* - ( $\varphi$ ): Mapa de confianza de la imagen con valores entre 0 y 1

**Salida:**

- *CorrimientoMedia* -  $M_h(x)$ : Modas a las que convergieron cada uno de los datos de la imagen

1. Establecer condiciones iniciales

- a) Número de iteraciones máximo, *Iteraciones* = 100
- b) Umbral de convergencia, si el CM no mueve el dato por más de este umbral, se alcanzó la convergencia, *Umbral* = 0.01
- c) Inicializa puntos de convergencia, *CorrimientoMedia*<sub>*ij*</sub> = 0

2. Para cada uno de los datos de la imagen

a) Obtener dato a procesar,  $Dato = Imagen_{ij}$

b) Mientras no se alcance el número de *Iteraciones* y el *Umbral* de convergencia

1) Determinar el vecindario de *Dato* en el dominio espacial, esto es, todos aquellos pixeles cuya distancia Euclidiana a partir de dato no sea mayor a *RadioEspacial*,  $Vecino_{mn}$

2) Para cada uno de los  $Vecino_{mn}$

3) Si diferencia de intensidad entre Dato y el vecino es menor al radio de intensidad

$$Vecino_{mn} < RadioIntensidad$$

a' Suma el vecino ponderado por el mapa de confianza,

$$Sumatoria+ = (1 - MapaConfianza_{mn})Vecino_{mn}$$

4) Calcula el CM,  $M_h(Dato) = Sumatoria - Dato$

c) Almacena el punto de convergencia,  $CorrimientoMedia_{ij} = M_h(Dato)$

# Apéndice D

## Análisis de adyacencia de regiones

Se aplica a una Matriz de Adyacencia de Regiones (MAR), actualizando las etiquetas, modas y número de puntos de cada moda para reflejar el nuevo conjunto de regiones fusionadas.

### D.1. Etiquetar imagen filtrada

Etiquetar las regiones homogéneas de la imagen filtrada, obtenidas al aplicar el procedimiento de Corrimiento de Media (CM). El valor de las etiquetas es un entero que va desde 0 hasta el número de regiones contenidas en la imagen filtrada. El etiquetado se realiza por medio de un llenado de conectividad 8.

$[ImagenEtiquetada, NumeroEtiquetas, Modas, DatosModa] = etiqueta(Filtrada)$

**Entrada:**

- *Filtrada*: Imagen filtrada por CM

**Salida:**

- *ImagenEtiquetada*: Imagen filtrada con regiones etiquetadas
- *NumeroEtiquetas*: Número de etiquetas usadas en el etiquetado de la imagen
- *Modas*: Vector con las modas correspondientes a cada una de las regiones de la imagen
- *DatosModa*: Vector con el número de datos que convergieron a cada moda

1. Inicializa *ImagenEtiquetada* con -1 para indicar dato no etiquetado



2. Inicializa  $NumeroEtiquetas = 0$

3. Para cada uno de los datos

a) Si el dato no esta etiquetado,  $ImagenEtiquetada_{ij} < 0$ , etiquetarlo

1) Incrementa contador de etiquetas,

$$NumeroEtiquetas+ = 1$$

2) Asigna etiqueta al dato,

$$ImageEtiquetada_{ij} = NumeroEtiquetas$$

3) Almacena la moda correspondiente a la región,

$$Modas(NumeroEtiquetas) = Filtrada_{ij}$$

4) Actualiza la cuenta del número de datos que tienen la misma moda, es decir, cuenta el número de datos que pertenecen a la misma región,  $DatosModa(NumeroEtiquetas)+ = 1$

5) Etiqueta todos los datos que pertenecen a la misma región, analizando sus 8 vecinos. Si el vecino tiene el mismo valor de intensidad que el dato, forma parte de la misma región y se le asigna la misma etiqueta y se incrementa la cuenta del número de datos de la región,  $DatosModa(NumeroEtiquetas)+ = 1$

## D.2. Fusión de regiones por contracción de un grafo de adyacencias

Para procesar el grafo, primero se debe hacer una representación del mismo construyendo una MAR. Las dimensiones de la MAR están dadas por el número de etiquetas de la imagen filtrada, esto es la MAR es de tamaño  $(NumeroEtiquetas \times NumeroEtiquetas)$ . La MAR se construye con valores booleanos en la que  $MAR(m, n) = 1$  si existe una arista desde el vértice  $m$  al  $n$ , es decir, si la región  $m$  es vecina de la región  $n$ , y es cero en caso contrario. La diagonal de la MAR se deja en cero, significando que una región no tiene adyacencia consigo misma.

La fusión de regiones se realiza aplicando criterios de similitud en modas de intensidad y bordes débiles. La reducción del grafo se logra empleando un algoritmo de encontrar y unir, las modas de las regiones fusionadas se recalculan y finalmente, se etiquetan las regiones fusionadas.

**[ImagenEtiquetada, NumeroEtiquetas, Modas, DatosModa] =  
 fusiona (ImagenEtiquetada, NumeroEtiquetas, FuerzaBorde, MapaConfianza,  
 Modas, DatosModa, RadioIntensidad)**

**Entrada:**

- *ImagenEtiquetada*: Imagen filtrada con las regiones etiquetadas
- *NumeroEtiquetas*: Número de etiquetas usadas en el etiquetado de la imagen
- *FuerzaBorde* - ( $\xi$ ): Umbral de fuerza promedio entre los bordes de una región con sus regiones vecinas
- *MapaConfianza* - ( $\varphi$ ): Mapa de confianza de la imagen con valores entre 0 y 1
- *Modas*: Vector con las modas correspondientes a cada una de las regiones de la imagen
- *DatosModa*: Vector con el número de datos que convergieron a cada moda
- *RadioIntensidad* - ( $h_i$ ): Radio de intensidad del kernel utilizado en el cálculo del CM

**Salida:**

- *ImagenEtiquetada*: Imagen con las regiones fusionadas
- *NumeroEtiquetas*: Número de etiquetas actualizado de acuerdo a las imágenes fusionadas
- *Modas*: Vector con las modas actualizadas para cada una de las regiones fusionadas
- *DatosModa*: Vector con el número de datos actualizado que conforman cada región fusionada

1. Construir la  $MAR(m,n)$  a partir de *ImagenEtiquetada*, observando los vecinos a la derecha y abajo de un dato dado, si sus etiquetas son diferentes, entonces son adyacentes.

2. Calcular la fuerza promedio (*FuerzaPromedio*) de los bordes entre regiones adyacentes. Para esto, se suman los valores de *MapaConfianza* para todos los datos que forman parte de un borde y se cuenta el número de pixeles que lo constituyen, para poder promediar. El resultado se almacena en una estructura igual a la MAR, cuyo contenido es el valor de la fuerza promedio de las adyacencias.
3. Generar los representantes canónicos de cada región, esto es la etiqueta original que le corresponde a cada región, *Etiquetas = 1* hasta *NumeroEtiquetas*
4. Intenta unir las regiones vecinas  $R_m$  y  $R_n$  para todo  $m \neq n$ . Para cada una de las regiones en la MAR, desde  $m = 1$  hasta *NumeroEtiquetas*
  - a) Obtener el número de vecinos adyacentes a  $R_m$ , *NumeroVecinos*, utilizando la MAR
  - b) Para cada uno de los vecinos, desde  $n = 1$  hasta *NumeroVecinos*
    - 1) Si la diferencia de intensidad normalizada respecto al radio de intensidad es menor a 0.5 ( $|(Modas(m) - Modas(n))/h_i| < 0,5$ ) y *FuerzaPromedio* con su vecino es menor al umbral *FuerzaBorde*,  $R_m$  y su vecino  $R_n$  pertenecen al mismo conjunto, por lo tanto, unirlos
      - a' Ejecutar operación *encontrar* sobre el representante canónico de  $R_m$ , *Canonico = m*, y si el representante cambio, obtener el nuevo representante que le fue asignado al fusionarla con otra región, *Canonico = Etiquetas(m)*
      - b' Ejecutar operación *encontrar*, igual que arriba, para su vecino  $R_n$ , *CanonicoVecino = n*
      - c' Si los elementos canónicos son diferentes, entonces ejecutar operación *unir*, haciendo que el representante canónico con la etiqueta menor sea el padre de la otra región, esto es si  $Canonico < CanonicoVecino$ , entonces  $Etiquetas(CanonicoVecino) = Canonico$ , en otro caso,  $Etiquetas(Canonico) = CanonicoVecino$
5. Equilibrar peso de árboles formados por los elementos canónicos

6. Acumular las modas y recalcular el número de datos usando los elementos canónicos. Para cada  $R_m$ , obtener su (tal vez, nuevo) elemento canónico y el número de datos de las regiones que se fusionaron,  $m = 1$  hasta  $NumeroEtiquetas$

a)  $Canonico = Etiquetas(m)$

b)  $ModaTemporal(Canonico) += DatosModa(m)*Modas(m)$

c)  $DatosTemporal(Canonico) += DatosModa(m)$

7. Calcular las modas de las regiones nuevas y su número de datos.

a) Los elementos distintos de cero de  $DatosTemporal$  son el número de datos de las nuevas regiones, y los índices de estos elementos representan las nuevas etiquetas,  $Indice = (DatosTemporal \neq 0)$ . Inicializar contador de etiquetas nuevas,  $EtiquetaNueva = 0$

b) Para cada una de las regiones resultantes de la fusión,  $m = 1$  hasta  $tamaño(Indice)$

1) Obtener elemento canónico de la región,  $Canonico = Indice(m)$

2) Encontrar regiones que se fusionaron al elemento canónico de la región,  $Fusionadas = (Etiquetas == Canonico)$

3) Si hubo regiones fusionadas

a' Incrementar contador de etiquetas nuevas,  $EtiquetaNueva += 1$

b' Asignar una etiqueta nueva a las regiones fusionadas,  $EtiquetasTemporal(Fusionadas) = EtiquetaNueva$

c' Recalcular la moda de las regiones fusionadas,  $ModaNueva(m) = ModaTemporal(Canonico) / DatosTemporal(Canonico)$

d' Asignar el nuevo número de datos de la región,  $DatosNuevos(m) = DatosTemporal(Canonico)$

8. Reetiquetar la imagen con las nuevas etiquetas,  $ImagenEtiquetada(i, j) = EtiquetasTemporal(Etiquetas(ImagenEtiquetada(i, j)))$

9. Actualizar el número de etiquetas, correspondiendo al nuevo número de regiones de la imagen fusionada,  $NumeroEtiquetas = EtiquetaNueva$

10. Actualizar modas de región,  $Modas = ModaNueva$  y número de datos de región,  $DatosModa = DatosNuevos$

La fusión de regiones se logra aplicando iterativamente la función `fusiona` hasta que el número de regiones, entre iteraciones, permanezca sin cambios, o bien se alcance un número dado de iteraciones, como se muestra a continuación

```
anterior = NumeroEtiquetas
contador = 0
diferenciaRegiones = anterior
mientras diferenciaRegiones >0 y contador <10
{
    ■ [ImagenEtiquetada, NumeroEtiquetas, Modas, DatosModa] = fusiona
      (ImagenEtiquetada, NumeroEtiquetas, FuerzaBorde, MapaConfianza,
       Modas, DatosModa, RadioIntensidad)
    ■ diferenciaRegiones = anterior - NumeroEtiquetas
    ■ incrementa contador
    ■ anterior = NumeroEtiquetas
}
```

### D.3. Podado de la imagen fusionada

Podar las regiones de la imagen fusionada que tengan un número de elementos inferior al tamaño mínimo de región dado por un umbral, removiendo las regiones cuya área sea inferior al umbral.

`[ImagenEtiquetada, NumeroEtiquetas, Modas, DatosModa] = podado (ImagenEtiquetada, NumeroEtiquetas, Modas, DatosModa, TamañoRegion, RadioIntensidad)`

**Entrada:**

- *ImagenEtiquetada*: Imagen fusionada con las regiones etiquetadas
- *NumeroEtiquetas*: Número de etiquetas de la imagen fusionada, equivale al número de regiones

- *Modas*: Vector con las modas correspondientes a cada una de las regiones de la imagen
- *DatosModa*: Vector con el número de datos que convergieron a cada moda
- *TamanoRegion* - ( $\mu$ ): Número mínimo de píxeles que debe tener una región
- *RadioIntensidad* - ( $h_i$ ): Radio de intensidad del kernel utilizado en el cálculo del CM

**Salida:**

- *ImagenEtiquetada*: Imagen en la que las regiones menores al umbral se han podado
  - *NumeroEtiquetas*: Número de etiquetas actualizado de acuerdo a las regiones podadas
  - *Modas*: Vector con las modas actualizadas para cada una de las regiones resultantes del podado
  - *DatosModa*: Vector con el número de datos actualizado para cada región resultante del podado
1. Construir la  $MAR(m, n)$  a partir de *ImagenEtiquetada*, observando los vecinos a la derecha y abajo de un dato dado, si sus etiquetas son diferentes, entonces son adyacentes.
  2. Generar los representantes canónicos de cada región, esto es la etiqueta original que le corresponde a cada región,  $Etiquetas = 1$  hasta *NumeroEtiquetas*
  3. Recorrer la MAR uniendo regiones cuya área sea menor a *TamanoRegion* con su región candidata
    - a) Para cada una de la regiones desde  $m = 1$  hasta *NumeroEtiquetas*
      - 1) Si el tamaño de la región es menor al umbral,  $DatosModa_m < TamanoRegion$ , únela a su región candidata
        - a' Determinar las regiones vecinas a  $R_m$ , *NumeroVecinos*, utilizando la MAR y considerarlas como regiones candidatas

*b'* De las regiones vecinas desde  $n = 1$  hasta *NumeroVecinos*, determinar la región candidata como aquella que tiene una diferencia mínima en moda con  $R_m$ ,

$$Candidata = \underset{n}{\text{mín}}([Modas(m) - Modas(n)]/RadioIntensidad)$$

*c'* Unir la región  $R_m$  con la región *Candidata*, haciendo al elemento canónico con etiqueta menor, el padre de la otra región, esto es, si

$Etiquetas(R_m) < Etiquetas(Candidata)$  entonces

$Etiquetas(Candidata) = Etiquetas(R_m)$ , en otro caso

$Etiquetas(R_m) = Etiquetas(Candidata)$

4. Equilibrar peso de árboles formados por los elementos canónicos

5. Acumular las modas y recalcular el número de datos usando los elementos canónicos. Para cada  $R_m$ , obtener su (tal vez, nuevo) elemento canónico y el número de datos de las regiones que se fusionaron,  $m = 1$  hasta *NumeroEtiquetas*

*a)*  $Canonico = Etiquetas(m)$

*b)*  $ModaTemporal(Canonico) += DatosModa(m)*Modas(m)$

*c)*  $DatosTemporal(Canonico) += DatosModa(m)$

6. Calcular las modas de las regiones nuevas y su número de datos.

*a)* Los elementos distintos de cero de *DatosTemporal* son el número de datos de las nuevas regiones, y los índices de estos elementos representan las nuevas etiquetas,  $Indice = (DatosTemporal \neq 0)$ . Inicializar contador de etiquetas nuevas,  $EtiquetaNueva = 0$

*b)* Para cada una de las regiones resultantes de la fusión,  $m = 1$  hasta *tamaño(Indice)*

1) Obtener elemento canónico de la región,  $Canonico = Indice(m)$

2) Encontrar regiones que se fusionaron al elemento canónico de la región,  $Fusionadas = (Etiquetas == Canonico)$

3) Si hubo regiones fusionadas

- a' Incrementar contador de etiquetas nuevas,  $EtiquetaNueva += 1$
  - b' Asignar una etiqueta nueva a las regiones fusionadas,  $EtiquetasTemporal(Fusionadas) = EtiquetaNueva$
  - c' Recalcular la moda de las regiones fusionadas,  $ModaNueva(m) = ModaTemporal(Canonico) / DatosTemporal(Canonico)$
  - d' Asignar el nuevo número de datos de la región,  $DatosNuevos(m) = DatosTemporal(Canonico)$
7. Reetiquetar la imagen con las nuevas etiquetas,  $ImagenEtiquetada(i, j) = EtiquetasTemporal(Etiquetas(ImageEtiquetada(i, j)))$
  8. Actualizar el número de etiquetas, correspondiendo al nuevo número de regiones de la imagen fusionada,  $NumeroEtiquetas = EtiquetaNueva$
  9. Actualizar modas de región,  $Modas = ModaNueva$  y número de datos de región,  $DatosModa = DatosNuevos$

El proceso de podado se aplica iterativamente hasta que el número de regiones, entre iteraciones, permanezca sin cambios, o bien se alcance un número dado de iteraciones, como se muestra a continuación

anterior = NumeroEtiquetas

contador = 0

diferenciaRegiones = anterior

mientras diferenciaRegiones >0 y contador <10

{

- $[ImagenEtiquetada, NumeroEtiquetas, Modas, DatosModa] = \text{podado}(ImagenEtiquetada, NumeroEtiquetas, Modas, DatosModa, Tama\tilde{n}oRegion, RadioIntensidad)$

- $diferenciaRegiones = anterior - NumeroEtiquetas$

- incrementa contador

- anterior = NumeroEtiquetas

}



# Apéndice E

## Clasificación de regiones

### E.1. Determinar probabilidad de pertenencia de cada región

Utilizando los mapas de probabilidad que vienen con el SPM99, determinar la probabilidad de cada una de la regiones de la imagen de pertenecer a algún tipo de tejido, ya se Materia Gris, Materia Blanca, Líquido Cefalorraquídeo o ninguno de estos tres, y asignar pertenencia a Fondo.

***PertenenciaRegion*** = probabilidades (*ImagenEtiquetada*, *NumeroEtiquetas*, *MapaFondo*, *MapaCSF*, *MapaGris*, *MapaBlanca*)

**Entrada:**

- *ImagenEtiquetada*: Imagen resultante del podado con las regiones a clasificar
- *NumeroEtiquetas*: Número de etiquetas (regiones) diferentes en la *ImagenEtiquetada*
- *MapaFondo*: Mapa probabilístico de la clase que no contiene tejido cerebral (es decir, CSF, MG o MB)
- *MapaCSF*: Mapa probabilístico de la clase que pertenece a Líquido Cefalorraquídeo
- *MapaGris*: Mapa probabilístico de la clase que pertenece a Materia Gris
- *MapaBlanca*: Mapa probabilístico de la clase que pertenece a Materia Blanca

**Salida:**

- *PertenenciaRegion*: Probabilidad de cada región de pertenecer a Fondo, CSF, MG o MB

1. Calcular la probabilidad a priori  $P(\omega_i)$  para cada clase

$$\omega_i = \{Fondo, CSF, MG, MB\}$$

utilizando la interpretación frecuentista de la probabilidad:

$$Apriori(\omega_i) = \frac{\sum Mapa_i}{\sum_{j=1}^4 (Mapa_j)}$$

2. Para cada una de la regiones de *ImagenEtiquetada*, calcular la probabilidad de pertenencia, para  $m = 1$  hasta *NumeroEtiquetas*

- Obtener los pixeles que forman la región  $R_m$  de *ImagenEtiquetada*,  $Pixeles = \{x_m \in ImagenEtiquetada\}$  tal que  $x_m \in R_m$
- Calcular la probabilidad que aportan los datos observados,  $p(Pixeles | \omega_i)$  de la región

$$ProbDatos(Pixeles | \omega_i) = \frac{\sum Mapa_i(Pixeles)}{\sum_{j=1}^4 Mapa_j(Pixeles)}$$

- Calcular la probabilidad de pertenencia a la clase  $\omega_i$  dados los datos (*Pixeles*) de la región

$$PertenenciaRegion(Pixeles | \omega_i) = \frac{ProbDatos(Pixeles | \omega_i) Apriori(\omega_i)}{\sum_{j=1}^4 ProbDatos(Pixeles | \omega_j) Apriori(\omega_j)}$$

## E.2. Etiquetado final de las regiones

Asignación de las regiones de la imagen a un tipo de tejido cerebral, basada en una medida de proximidad a la región semilla para cada tejido. La región semilla se selecciona considerando en la probabilidad de pertenencia de la región a cada una de las clases de interés: Fondo, CSF, MG y MB.

***ImagenSegmentada = segmenta (ImagenEtiquetada, NumeroEtiquetas, PertenenciaRegion, Modas, DatosModa, RadioIntensidad, MapaConfianza)***

## Entrada:

- *ImagenEtiquetada*: Imagen con las regiones resultantes del podado
- *NumeroEtiquetas*: Número de etiquetas actualizado de acuerdo a las regiones podadas
- *PertenenciaRegion*: Probabilidad de cada región de pertenecer a Fondo, CSF, MG o MB
- *Modas*: Vector con las modas actualizadas para cada una de las regiones resultantes del podado
- *DatosModa*: Vector con el número de datos actualizado para cada región resultante del podado

## Salida:

- *ImagenSegmentada*: Imagen con las regiones asignadas a algún tipo de tejido cerebral
1. Determinar las regiones semilla para cada una de las clases, buscando la región con mayor número de píxeles y con mayor probabilidad en cada clase, obteniéndose así, la moda característica de la región semilla
    - a) Inicializa las regiones semilla,  $Semilla_{ij} = 0$  para  $i = 1$  a 4 y  $j = 1$  a 3, donde  $i$  representa las clases, y  $j$  representa el número de píxeles de la región, la moda de la región y número de etiqueta de la región
    - b) Para cada una de las regiones  $R_m$  de la imagen,  $m = 1$  hasta *NumeroEtiquetas*
      - 1) Obtener el índice de la probabilidad máxima de la región, para  $\omega_i = \{Fondo, CSF, MG, MB\}$ 
$$I = \underset{i}{\text{máx}}\{PertenenciaRegion_m(\omega_i)\}$$
      - 2) Verificar que sea la región más grande. Si el número de píxeles de  $R_m$  es mayor que el máximo de la clase,  $DatosModa(m) > Semilla_{I1}$ , actualizar la región semilla

$$a' \text{ Semilla}_{I1} = \text{DatosModa}(m)$$

$$b' \text{ Semilla}_{I2} = \text{Modas}(m)$$

$$c' \text{ Semilla}_{I3} = m$$

2. Para cada una de las regiones  $R_m$ ,  $m = 1$  hasta  $\text{NumeroEtiquetas}$

a) Calcular medida de distancia entre intensidad de región y semilla,  $\text{Distancia}_i = |\text{Modas}(m) - \text{Semilla}_{i2}|$ , para cada  $\omega_i$

b) Calcular probabilidad de cercanía de la región a cada una de las  $\omega_i$  clases,  $\text{Cercania}(m)_i = 1 - \frac{\text{Distancia}_i}{\sum_{k=1}^4 \text{Distancia}_k}$

3. La imagen de regiones clasificadas se inicializa como  $\text{Fondo}$ ,  $\text{ImagenSegmentada} = 0$

4. Para cada una de las regiones  $R_m$ ,  $m = 1$  hasta  $\text{NumeroEtiquetas}$

a) Obtener los pixeles que forman la región  $R_m$  de  $\text{ImagenEtiquetada}$ ,  $\text{Pixeles} = \{x_m \in \text{ImagenEtiquetada}\}$  tal que  $x_m \in R_m$

b) Obtener la clase máxima de la región,  $\text{Clase} = \max_i \{\text{Cercania}(m)_i\}$

c) Segmentar  $\text{ImagenEtiquetada}$  de acuerdo a la clase máxima de  $R_m$

1) Si  $\text{Clase}$  es igual a 'LCR', asignar etiqueta de LCR a la región,

$$\text{ImagenSegmentada}(\text{Pixeles}) = \text{LCR}$$

2) Si  $\text{Clase}$  es igual a 'MG', asignar etiqueta de MG a la región,

$$\text{ImagenSegmentada}(\text{Pixeles}) = \text{MG}$$

3) Si  $\text{Clase}$  es igual a 'MB'; asignar etiqueta de MB a la región,

$$\text{ImagenSegmentada}(\text{Pixeles}) = \text{MB}$$

# Apéndice F

## Índice de Dice

Para comparar los resultados del proceso de segmentación investigado con el volumen de referencia, se puede también calcular la métrica utilizada en [45], donde se emplea la siguiente expresión para evaluar la similitud entre imágenes:

$$ID_{XY}(k) = \frac{2n_{X \cap Y}(k)}{n_X(k) + n_Y(k)}$$

donde  $n_{X \cap Y}(k)$ , denota el número de voxeles que se asignan a la clase  $k$ , tanto por el proceso de segmentación automática como por el volumen de referencia, considerado como verdad de tierra. Similarmente,  $n_X(k)$  y  $n_Y(k)$  denotan el número de voxeles asignados a la clase  $k$  por la segmentación automática y la referencia, respectivamente, dando un valor de 1 si ambas imágenes son iguales y 0 en caso contrario. Llamaremos a  $ID_{XY}(k)$  el índice de Dice [25] por ser precisamente este autor el primero en describirlo.

La Figura F.1, muestra la variación del índice de similitud contra el ruido, para cinco volúmenes de cerebro simulados, calculando la similitud con el índice de Dice. De las gráficas se observa que básicamente, tanto la medida de similitud calculada con los coeficientes de Tanimoto y el índice de Dice, tienen un comportamiento muy similar.

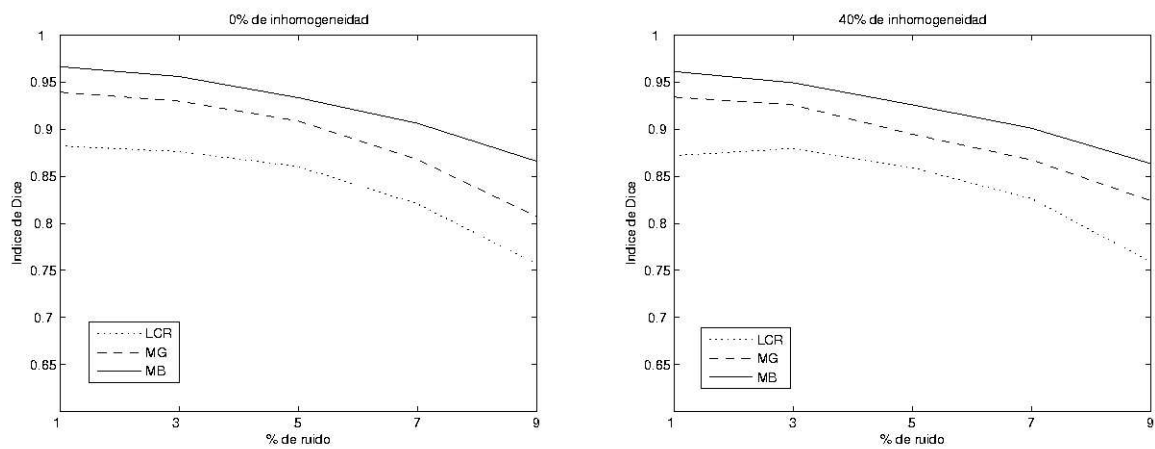


Figura F.1: Índices de similitud de Dice de cinco volúmenes simulados. Gráfica izquierda sin inhomogeneidad en intensidad y mostrando el índice de similitud para líquido cefalorraquídeo (LCR), materia gris (MG) y materia blanca (MB), contra porcentaje de ruido en el volumen. Gráfica derecha muestra la misma información, sólo que ahora todos los volúmenes tienen un 40 % de inhomogeneidad.

# Bibliografía

- [1] *McConnell Brain Imaging Centre*. <http://www.bic.mni.mcgill.ca/brainweb/>.
- [2] J. Ashburner. *Computational Neuroanatomy*. PhD thesis, University College London, 2000. [www.fil.ion.ucl.ac.uk/spm/doc/theses/john](http://www.fil.ion.ucl.ac.uk/spm/doc/theses/john).
- [3] J. Ashburner, K. Friston, A. Holmes, and J. B. Poline. Statistical Parametric Mapping - SPM99b. Technical report, The Wellcome Department Cognitive Neurology, Univ. College London, London, U. K., 1999. <http://www.fil.ion.ucl.ac.uk/spm/>.
- [4] J. Ashburner and K. J. Friston. Multimodal image coregistration and partitioning - a unified framework. Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK, 1997.
- [5] J. Ashburner and K. J. Friston. Spatial transformation of images. In R. S. J. Frackowiak, K. J. Friston, C. Frith, R. Dolan, and J. C. Mazziotta, editors, *Human Brain Function*, pages 43–58. Academic Press USA, 1997. <http://www.fil.ion.ucl.ac.uk/spm/doc/books/hbf1/Ch2.pdf>.
- [6] J. Ashburner and K. J. Friston. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7(4):254–266, 1999.
- [7] J. Ashburner, P. Neelin, D. L. Collins, A. Evans, and K. Friston. Incorporating prior knowledge into image registration. *Neuroimage*, 6:344–352, 1997. Article No. NI970299.
- [8] C. Baillard, P. Hellier, and C. Barillot. Segmentation of brain 3D MR images using level sets and dense registration. *Med. Image Anal.*, 5:185–194, 2001.
- [9] 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, April 1993.
- [10] P. Brodatz. *Textures: a photographic album for artists and designers*. Dover, New York, 1966.

- [11] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. and Machine Intell.*, 8(6):679–698, November 1986.
- [12] A. Chakraborty, L. H. Staib, and J. S. Duncan. Deformable boundary finding in medical images by integrating gradient and region information. *IEEE Trans. Med. Imag.*, 15(6):859–870, December 1996.
- [13] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. and Machine Intell.*, 17(8):790–799, August 1995.
- [14] E. Choi and P. Hall. Data sharpening as a prelude to density estimation. *Biometrika*, 86(4):941–947, December 1999.
- [15] C. M. Christoudias, B. Georgescu, and P. Meer. Synergism in low level vision. In *16th International Conference on Pattern Recognition*, pages 150–155, Quebec City, Canada, August 2002. vol. IV.
- [16] L. P. Clarke, R. P. Velthuizen, M. A. Camacho, J. J. Heine, M. Vaidyanathan, L. O. Hall, R. W. Thatcher, and M. L. Silbiger. MRI segmentation: Methods and applications. *Magn. Reson. Imag.*, 13(3):343–368, 1995.
- [17] L. P. Clarke, R. P. Velthuizen, S. Phuphanich, J. D. Schellenberg, J. A. Arrington, and M. Silbiger. MRI: Stability of three supervised segmentation techniques. *Magn. Reson. Imaging*, 11(1):95–106, 1993.
- [18] D. Comaniciu. An algorithm for data-driven bandwidth selection. *IEEE Trans. Pattern Anal. and Machine Intell.*, 25(2):281–288, 2003.
- [19] D. Comaniciu and P. Meer. Distribution free decomposition of multivariate data. *Patt. Anal. Appl.*, 2:22–30, 1999.
- [20] D. Comaniciu and P. Meer. Mean shift analysis and applications. In *IEEE Int. Conf. Computer Vision (ICCV'99)*, pages 1197–1203, Kerkyra, Greece, 1999.
- [21] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. and Machine Intell.*, 24(5):603–619, May 2002.
- [22] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. In *IEEE Conf. Comp. Vis. Patt. Recogn.*, volume 2, pages 142–149, Hilton Head, Shout Carolina, 2000.



- [23] D. Comaniciu, V. Ramesh, and P. Meer. The variable bandwidth mean shift and data-driven scale selection. In *IEEE Int. Conf. Computer Vision (ICCV'01)*, volume 1, pages 438–445, Vancouver, Canada, 2001.
- [24] D. Demigny and T. Kamle. A discrete expression of Canny’s criteria for step edge detector performances evaluation. *IEEE Trans. Pattern Anal. Machine Intell.*, 19(11):1199–1211, 1997.
- [25] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
- [26] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley & Sons, Inc., 605 Third avenue, New York, NY, second edition, 2001.
- [27] J. S. Duncan and N. Ayache. Medical image analysis: Progress over two decades and the challenges ahead. *IEEE Trans. Pattern Anal. and Machine Intell.*, 22(1):85–106, January 2000.
- [28] V. A. Epanechnikov. Nonparametric estimation of a multivariate probability density. *Theory Prob. Appl.*, 14:153–158, (URSS) 1969.
- [29] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters. 3D statistical neuroanatomical models from 305 MRI volumes. In *Proc. IEEE-nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817, 1993.
- [30] A. C. Evans, M. Kamber, D. L. Collins, and D. Macdonald. An MRI-based probabilistic atlas of neuroanatomy. In S. Shorvon, D. Fish, F. Andermann, G. Bydder, and H. Stefan, editors, *Magnetic resonance scanning and epilepsy*, volume 264 of *NATO ASI Series A, Life Sciences*, pages 263–274. Plenum Press, 1994.
- [31] C. Ferreira and E. Mainar. Estudio asintótico de polinomios ortogonales en la tabla de askey. *Rev. Real Academia de Ciencias Zaragoza*, 57:147–188, 2002.
- [32] K. J. Friston, J. Ashburner, C. D. Frith, J. B. Poline, J. D. Heather, and R. S. J. Frackowiak. Spatial registration and normalization of images. *Human Brain Mapping*, 2:165–189, 1995.
- [33] K. Fukunaga and L. D. Hostetler. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Information Theory*, 21(1):32–40, January 1975.

- [34] B. Georgescu, I. Shimshoni, and P. Meer. Mean shift based clustering in high dimensions: A texture classification example. In *Int. Conf. Computer Vision*, pages 456–463, Nice, France, 2003.
- [35] M. A. González Ballester, A. Zisserman, and M. Brady. Segmentation and measurement of brain structures in MRI including confidence bounds. *Med. Image Anal.*, 4:189–200, 2000.
- [36] R. Guillemaud and M. Brady. Estimating the bias field of MR images. *IEEE Trans. Med. Imag.*, 16(3):238–251, June 1997.
- [37] M. D. Heath, S. Sarkar, T. Sanocki, and K. W. Bowyer. A robust visual method for assessing the relative performance of edge-detection algorithms. *IEEE Trans. Pattern Anal. and Machine Intell.*, 19(12):1338–1359, December 1997.
- [38] P. Hellier, C. Barillot, E. Memin, and P. Perez. An energy-based framework for dense 3D registration of volumetric brain images. In *IEEE Conf. Comp. Vis. Patt. Recogn.*, pages 270–275, Hilton Head, South Carolina, June 2000.
- [39] IBSR. *Internet Brain Segmentation Repository*. <http://www.cma.mgh.harvard.edu/ibsr>.
- [40] A. K. Jain, R. P. W. Duin, and J. Mao. Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal. and Machine Intell.*, 22(1):4–37, January 2000.
- [41] J. R. Jiménez, V. Medina, and O. Yáñez. Nonparametric MRI segmentation using mean shift and edge confidence maps. *Progress in Biomedical Optics and Imaging*, 4(23):1433–1441, February 2003.
- [42] N. M. John, M. R. Kabuka, and M. O. Ibrahim. Multivariate statistical model for 3D image segmentation with application to medical images. *Journal of Digital Imaging*, 16(4):365–377, December 2003.
- [43] R. Koekoek and R. F. Swarttouw. The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue. Technical Report 98-17, Delft University of Technology, Faculty of Information Technology and Systems, Department of Technical Mathematics and Informatics, 1998. <http://aw.twi.tudelft.nl/koekoek/askey.html>.
- [44] K. V. Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based bias field correction of MR images of the brain. *IEEE Trans. Med. Imag.*, 18(10):885–896, October 1999.

- [45] K. V. Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Med. Imag.*, 18(10):897–908, October 1999.
- [46] Z. Liang, J. R. Macfall, and D. P. Harrington. Parameter estimation and tissue segmentation from multispectral MR images. *IEEE Trans. Med. Imag.*, 13(3):441–449, September 1994.
- [47] J. L. Marroquin, B. C. Vemuri, S. Botello, S. Calderon, and A. Fernandez-Bouzas. An accurate end efficient bayesian method for automatic segmentation of brain MRI. *IEEE Trans. Med. Imag.*, 21(8):934–945, 2002.
- [48] P. Meer and B. Georgescu. Edge detection with embedded confidence. *IEEE Trans. Pattern Anal. and Machine Intell.*, 23(12):1351–1365, December 2001.
- [49] P. Meer and I. Weiss. Smoothed differentiation filters for images. *J. Visual Comm. and Image Representation*, 3:58–72, 1992.
- [50] H. Park, P. H. Bland, and C. R. Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans. Med. Imag.*, 22(4):483–492, April 2003.
- [51] J. C. Rajapakse and F. Kruggel. Segmentation of MR images with intensity inhomogeneities. *Image and Vision Computing*, 16:165–180, 1998.
- [52] J. Rogowska. *Handbook of Medical Imaging Processing and Analysis*. Academic Press, 2000.
- [53] P. Schoroeter, J.-M. Vesin, T. Langenberger, and R. Meuli. Robust parameter estimation of intensity distributions for brain magnetic resonance images. *IEEE Trans. Med. Imag.*, 17(2):172–186, April 1998.
- [54] R. Sedgewick. *Algorithms in C*. Addison-Wesley, 1990.
- [55] S. Z. Selim and M. A. Ismail. K-means-type algorithms: a generalized convergence theorem and characterization of local optimality. *IEEE Trans. Pattern Anal. and Machine Intell.*, 6(1):81–87, 1984.
- [56] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Number 26 in Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, 1986.

- [57] J. S. Suri. Two-dimensional fast magnetic resonance brain segmentation. *IEEE Eng. in Medicine and Biology*, pages 84–95, July/August 2001.
- [58] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *IEEE Int. Conf. Computer Vision*, pages 839–846, Bombay, India, January 1998.
- [59] M. Vaidyanathan, L. P. Clarke, C. Heidtman, R. P. Velthuizen, and L. O. Hall. Normal brain volume measurements using multispectral MRI segmentation. *Magn. Reson. Imag.*, 15(1):87–97, 1997.
- [60] R. Valdés-Cristerna, V. Medina-Bañuelos, and O. Yáñez-Suárez. Coupling of radial-basis network and active contour model for multispectral brain MRI segmentation. *IEEE Trans. Biomed. Eng.*, 51(3):459–470, 2004.
- [61] Y. Wang, T. Adah, S.-Y. Kung, and Z. Szabo. Quantification and segmentation of brain tissues from MR images: A probabilistic neural network approach. *IEEE Trans. Image Processing*, 7(8):1165–1181, August 1998.
- [62] W. M. Wells III, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag.*, 15(4):429–442, August 1996.
- [63] C. Xu and J. L. Prince. Generalized gradient vector flow external forces for active contours. *Signal Processing*, 71:131–139, 1998.
- [64] C. Xu and J. L. Prince. Snakes, shapes and gradient vector flow. *IEEE Trans. Image Processing*, 7(3):359–369, March 1998.
- [65] Y. Xu and E. C. Uberbacher. 2D image segmentation using minimum spanning-trees. *Image and Vision Computing*, 15(1):47–57, January 1997.
- [66] X. Zeng, L. H. Staib, R. T. Schultz, and J. S. Duncan. Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation. *IEEE Trans. Med. Imag.*, 18(10):927–937, October 1999.
- [67] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.*, 20(1):45–57, January 2001.