



03063

5
26)

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Unidad Académica de Ciclos Profesionales y de Posgrado
del Colegio de Ciencias y Humanidades

ALGORITMO DE SEGMENTACIÓN BASADO EN
ESTADÍSTICA ROBUSTA Y MATRIZ DE COOCURRENCIA

T E S I S

Que para obtener el grado de:
MAESTRA EN CIENCIAS DE LA COMPUTACIÓN

Presenta:

MARÍA ELENA MARTÍNEZ PÉREZ

Director de Tesis: MCC. María Garza.

México, D.F.

Febrero 1996.

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Este trabajo fue realizado en el Departamento de Ciencias de la Computación del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, institución a la que agradezco su apoyo.

Con toda mi gratitud y aprecio a la M. en C. María Garza directora de este trabajo, por permitirme colaborar con ella, por sus enseñanzas y por su incondicional apoyo.

A Humberto Gutiérrez del Instituto de Fisiología Celular de la UNAM, por prestarme sus imágenes y por sus buenas ideas y comentarios.

A la Dirección General de Asuntos del Personal Académico (DGAPA) por apoyarme con una beca para realización de Estudios de Maestría.

A todas aquellas personas que de alguna manera hicieron posible la realización del presente trabajo.

Contenido

Introducción	4
1 Métodos Clásicos de Segmentación	6
1.1 Características basadas en umbrales o cúmulos	7
1.1.1 Segmentación por umbral	7
1.1.2 Segmentación por cúmulos	17
1.2 Detección de bordes	18
1.2.1 Técnicas Paralelas	19
1.2.2 Técnicas Secuenciales	20
1.3 Extracción de regiones	21
1.3.1 Búsqueda Regional de Vecinos	21
1.3.2 Heurísticas Multirregionales	22
1.3.3 Aproximación Funcional y Mezcla	23
1.3.4 Esquemas de División y Unión	23
1.3.5 Interpretaciones Regionales y Semánticas	24
2 Descripción del Algoritmo de Segmentación	26
2.1 Estimación de las μ_i	27
2.1.1 Estadística Robusta	27
2.1.2 Estimación de las μ_i . Implantación	30
2.2 Estimación de las σ_i	31
2.2.1 Matriz de Coocurrencia	31
2.2.2 Estimación de las σ_i . Implantación	33
2.3 Crecimiento de Regiones	34
2.3.1 Implantación	34
3 Resultados	37
3.1 Validación	37
3.2 Aplicación	43

4 Conclusiones	46
4.1 Tendencias al futuro	47
Bibliografía	48
Apéndice	52

Introducción

El procesamiento digital de imágenes (PDI) se encarga del estudio de las metodologías para modificar y analizar imágenes por medio de una computadora digital. En los últimos tiempos la potencialidad del PDI se ha vuelto ilimitada debido a la gran variedad de aplicaciones que han surgido, por lo que el desarrollo de nuevas técnicas de procesamiento ha crecido a una gran velocidad y continúa haciéndolo día con día.

Una de las primeras aplicaciones del PDI fué sobre imágenes de satélite o de percepción remota, imágenes de planetas, estrellas o satélites naturales como la luna. Otra aplicación de suma importancia es en la medicina, como por ejemplo imágenes de tomografía que son utilizadas para detectar, entre otros problemas, tumores cerebrales. Imágenes de tomografía por emisión de positrones que se utilizan para medir cambios metabólicos, por ejemplo del consumo de glucosa en el cerebro. También se analizan imágenes de rayos-X con las cuales es necesario resaltar algunas partes de interés. Y así podríamos mencionar muchos otros campos de aplicación como la electrónica, la robótica, la biología, la física, la agricultura, la geografía, la antropología, la restauración de obras de arte pictóricas, entre muchos otros más.

El PDI se clasifica generalmente en cuatro áreas según el tipo de procesamiento a realizar [27]. Estas son **realce**, **restauración**, **codificación**, y **análisis**. El *realce* consiste en mejorar la apariencia de la imagen para su visualización o bien en resaltar cierta información de interés para su posterior análisis; los métodos y los objetivos varían según la aplicación. La *restauración* sirve para eliminar o minimizar degradaciones conocidas en una imagen. La *codificación* se refiere a la representación de una imagen con los menos bits posibles de manera que no se pierda el nivel de calidad y de legibilidad de la imagen, para fines de transmisión o almacenaje. Finalmente el *análisis* o interpretación tiene como objetivo representar simbólicamente el contenido de una imagen para poder realizar medidas cuantitativas de los elementos que la componen; las principales aplicaciones del análisis son la visión por computadora, la robótica y la identificación de objetos. El *análisis* de imágenes difiere de las otras tres áreas en un aspecto muy importante. En el *realce*, *restauración* y *codificación*, las entradas y salidas de los procesos son imágenes, mientras que en el *análisis* la entrada es una imagen y la salida es una representación simbólica del contenido de la imagen de entrada, como por ejemplo medidas del tamaño y la forma de los objetos que contiene.

La motivación inicial del presente trabajo surgió con la necesidad de detectar la región de la corteza en imágenes histológicas de cortes cerebrales de rata. Las imágenes con las que se cuenta son de dos tipos: imágenes control cuya corteza está en buen estado, e imágenes experimentales cuya corteza contiene visiblemente una lesión sobre el tejido. Es necesidad del investigador detectar estas lesiones y hacer una medición de densidad óptica sobre ellas para compararlas respecto a las imágenes control. Debido a que estas imágenes están tomadas con muy poco contraste y en malas condiciones de luz no es posible segmentarlas utilizando los métodos convencionales. De aquí que surgió la necesidad de desarrollar una técnica *ad hoc* que permitiera resolver el problema.

El trabajo que se presenta tiene como objetivo el diseño, la implementación y la validación de un algoritmo robusto de segmentación de imágenes. Con la segmentación se detectan los objetos o regiones en una imagen por lo que esta técnica entra dentro del área de *realce* y está en la frontera con el área de *análisis* ya que el resultado de la segmentación es la información de entrada para cualquier proceso de análisis. El alcance de este trabajo se enfocó exclusivamente al desarrollo de los algoritmos de segmentación, y no así en la parte de análisis que sería la medición de la densidad óptica de las regiones encontradas.

Debido a la gran variedad de aplicaciones, no es de extrañarse que existan muchas técnicas de segmentación desarrolladas desde hace mucho tiempo para aplicaciones particulares; estas técnicas se han categorizado según la información que proporcionan y el tipo de proceso que se aplica. En el capítulo 1 haremos una breve revisión sobre las técnicas de segmentación existentes de manera que podamos ubicar dentro de esta categorización el método de segmentación propuesto. En el capítulo 2 desarrollaremos la descripción del algoritmo; en el capítulo 3 presentaremos los resultados de la validación del algoritmo así como ejemplos de aplicación con las imágenes histológicas y un ejemplo de imagen médica; finalmente en el capítulo 4 concluiremos el resultado del trabajo y se discutirán las tendencias hacia el futuro.

Capítulo 1

Métodos Clásicos de Segmentación

Como se mencionó en la introducción, las técnicas de segmentación están contenidas dentro del área de *realce*. Estas técnicas son de las más importantes para el posterior análisis de la información contenida en la imagen. El buen resultado del proceso de segmentación conlleva a una correcta interpretación de esta información.

La segmentación es la división de una imagen en regiones que no se traslapan y cuya unión es la imagen completa. Una región es un área de la imagen cuyos puntos tienen alguna propiedad en común.

Lo anterior puede definirse formalmente de la siguiente manera [20]. Sea X una rejilla muestra de puntos de una imagen, es decir, pares de puntos

$$(i, j), \quad i = 1, 2, \dots, N \quad j = 1, 2, \dots, M$$

donde N y M son el número de píxeles en las direcciones x y y respectivamente. Sea Y un subconjunto no vacío de X consistente en puntos contiguos de la imagen. Se define como un *predicado uniforme* $P(Y)$ aquel en el que Y toma el valor de *falso* o *verdadero* dependiendo sólo de las propiedades relacionadas con los tonos de la matriz $f(i, j)$, para los puntos de Y . Además P tiene la propiedad de que si Z es un subconjunto no vacío de Y , entonces $P(Y) = \text{verdadero}$ implica que siempre $P(Z) = \text{verdadero}$.

La segmentación se define como una partición de X en subconjuntos disjuntos X_1, X_2, \dots, X_K de tal manera que:

1. $\bigcup_{i=1}^K X_i = X$
2. X_i , para $i = 1, 2, \dots, K$ está conectado.
3. $P(X_i) = \text{VERDADERO}$ para $i = 1, 2, \dots, K$.
4. $P(X_i \cup X_j) = \text{FALSO}$ para $i \neq j$, donde X_i y X_j son adyacentes.

La primera condición implica que la segmentación debe ser completa, es decir, que todos los puntos de la imagen deben pertenecer a una región. La segunda condición implica que todas las regiones deben estar conectadas, es decir, compuestas por puntos de rejillas contiguas. La tercera condición determina qué tipo de propiedades (predicados lógicos) deben tener las regiones segmentadas, por ejemplo, niveles de gris uniformes. Y la cuarta condición expresa la disyunción entre las regiones segmentadas, es decir la ausencia de superposición.

Dado que las técnicas de segmentación propuestas hasta ahora son *ad hoc* a la aplicación y al tipo de imagen, no existen algoritmos generales. Otra de las razones por las que no existen este tipo de algoritmos es que una imagen bidimensional puede representar potencialmente un número infinito de posibilidades. De aquí que las técnicas de segmentación se agrupen normalmente en tres clases [14]:

- Características basadas en umbrales o cúmulos ("clustering").
- Detección de bordes.
- Extracción de regiones.

En este capítulo se hará una breve descripción de las técnicas mencionadas como introducción a los algoritmos que se presentan en este trabajo.

1.1 Características basadas en umbrales o cúmulos

1.1.1 Segmentación por umbral

La segmentación de una imagen por umbral es un proceso aplicable a escenas que contienen objetos sólidos en un fondo de contraste. Una imagen con estas características tiene un histograma de forma *bimodal*, con un valle como se ilustra en la Figura 1.1(a). La técnica consiste en encontrar el valor del umbral T en el cual se encuentra el valle, esto es:

$$S(x, y) = \begin{cases} C_0 & \text{para } 0 \leq f(x, y) < T \\ C_1 & \text{para } T \leq f(x, y) < f_{max} \end{cases}$$

donde $f(x, y)$ es la imagen de entrada, $S(x, y)$ es la imagen de salida, T es el valor del umbral y C_0 y C_1 son las etiquetas que toman los pixeles correspondientes al objeto y al fondo de la imagen.

Lo anterior nos limita a imágenes donde sólo existen dos clases, sin embargo esta idea puede generalizarse para imágenes que tienen varias clases diferentes, es decir imágenes cuyo histograma es *multimodal* como el de la Figura 1.1(b). De manera general la segmentación por umbral se define como:

$$S(x, y) = k \quad \text{si } T_{k-1} \leq f(x, y) < T_k \quad \text{para } k = 0, 1, 2, \dots, m$$

donde (x, y) son las coordenadas x y y de un pixel; $S(x, y)$, $f(x, y)$ son la imagen de salida y la imagen de entrada respectivamente; T_0, \dots, T_m son los valores de los umbrales con T_0 igual al mínimo y T_m igual al máximo; m es el número de etiquetas diferentes asignadas a la imagen segmentada.

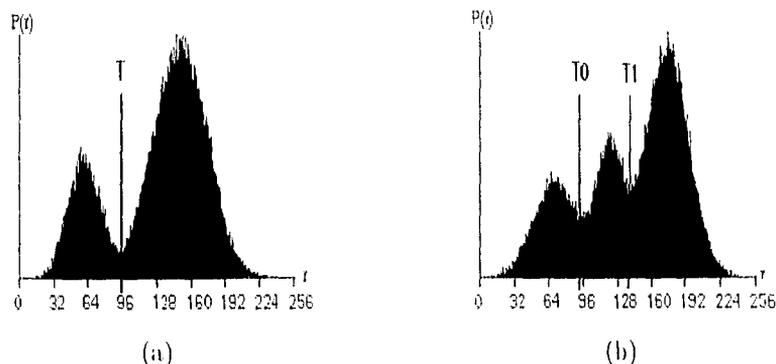


Figura 1.1: (a) Histograma bimodal. (b) Histograma multimodal.

En general, un método de umbral es aquel que determina el valor del umbral T basado en un cierto criterio. Si T se determina solamente en base al valor de gris de cada pixel entonces el método de umbral es **dependiente de un punto**. Si T se determina de una propiedad local (por ejemplo de una distribución local de niveles de gris) en la vecindad de un pixel entonces el método de umbral es **dependiente de una región**. Una técnica de umbral **global** es aquella que segmenta a una imagen completa con un sólo valor de umbral, mientras que una técnica de umbral **local** es aquella que parte a una imagen en subimágenes y determina un valor de umbral para cada una de ellas. De lo anterior se define la siguiente clasificación de métodos de umbral [44]:

- Técnicas globales dependientes de un punto.
- Técnicas globales dependientes de una región.
- Técnicas locales.
- Métodos Multiumbral.

Técnicas globales dependientes de un punto

1. *Método p -tilc*. En este método [11] se asume que una imagen consiste en objetos oscuros sobre un fondo claro. Se asume también que se conoce el porcentaje del área que ocupan los objetos. El umbral se define como el nivel de gris máximo que mapea por lo menos $(100 - p)\%$ de los pixeles que caen dentro de los objetos. Este método no funciona cuando no se conoce el porcentaje del área de los objetos.

2. *Método de Moda.* Para imágenes con objetos nítidos sobre un fondo, el histograma de los niveles de gris es bimodal. En este caso el umbral puede ser el nivel de gris que corresponde al valle del histograma [37]. Aunque este método es muy simple, no puede ser aplicado en imágenes cuyos picos en su histograma sean extremadamente diferentes, o aquellas cuyos valles sean anchos o planos.
3. *Método de Otsu.* Uno de los métodos clásicos de segmentación por umbral en la literatura, el cual utilizaremos en este trabajo para comparar nuestro método, es el método de OTSU [33]. El objetivo de este método es encontrar el valor de un umbral T , para el cual la varianza $\sigma_B^2(T)$ entre las regiones C_0 y C_1 a segmentar (si hablamos sólo de dos regiones) sea máxima. T es el punto en donde se encuentran más separadas.

Considérese una imagen con L niveles de gris. Normalizando su histograma para manejarlo como una distribución de probabilidades tenemos:

$$p_i = \frac{n_i}{N}, \quad p_i > 0, \quad \sum_{i=1}^L p_i = 1,$$

donde N es el número total de píxeles de la imagen, n_i el número de píxeles con tono de gris i , y p_i la probabilidad de ocurrencia del tono de gris i , para $i=1, 2, \dots, L$. La probabilidad de ocurrencia de cada región es:

$$\omega_0 = P_r(C_0) = \sum_{i=1}^T p_i = \omega(T)$$

$$\omega_1 = P_r(C_1) = \sum_{i=T+1}^L p_i = 1 - \omega(T)$$

y su nivel promedio:

$$\mu_0 = \sum_{i=1}^T \frac{ip_i}{\omega_0} = \frac{\mu(T)}{\omega(T)}$$

$$\mu_1 = \sum_{i=T+1}^L \frac{ip_i}{\omega_1} = \frac{\mu_T - \mu(T)}{1 - \omega(T)}$$

donde μ_T es el gris promedio de la imagen completa. Se trata entonces de maximizar la varianza entre regiones:

$$\sigma_B^2 = (\mu_1 - \mu_0)^2$$

Expresando σ_B^2 en términos de T tenemos:

$$\sigma_B^2(T) = \frac{[\mu_T \omega(T) - \mu(T)]^2}{\omega(T)[1 - \omega(T)]}$$

Con esta expresión se evalúan los posibles valores de T , hasta encontrar el valor para el cual $\sigma_B^2(T)$ sea máxima.

Este método también puede generalizarse a m número de umbrales, sin embargo mientras más grande sea este número, la expresión σ_B^2 y el procedimiento de maximización serán también cada vez más complicados.

4. *Método de análisis de concavidad del histograma.* Para imágenes en donde no es posible encontrar el valle en el histograma, se puede definir un buen umbral autilizando la concavidad del mismo [42].

Sea HS un histograma definido sobre el conjunto de niveles de gris g_0, g_1, \dots, g_{L-1} . Y sea el peso de los niveles de gris del histograma $h(g_0), h(g_1), \dots, h(g_{L-1})$, donde $h(g_i) \neq 0$ para toda i . Entonces, HS puede ser visto como una región bidimensional.

Para determinar las concavidades de HS , se construye el polígono convexo más pequeño \overline{HS} que contiene HS . Las concavidades de HS se determinan de la diferencia $HS - \overline{HS}$. Sea $\bar{h}(g_i)$ el peso de \overline{HS} en el nivel de gris g_i . Los posibles valores de umbral son aquellos niveles de gris en donde $\bar{h}(g_i) - h(g_i)$ tenga un máximo local.

5. *Métodos de Entropía.* En estos métodos el umbral óptimo se obtiene aplicando teoría de la información.

(i) *Método de Pun* [38]. Este método utiliza la entropía del histograma de los niveles de gris de la imagen como una medida de homogeneidad entre las clases. La *información* asociada con cada nivel de gris i del histograma se define como

$$I_i = -\log_e p_i$$

donde p_i es la probabilidad de ocurrencia de cada nivel de gris i . En base a esto la entropía del histograma se define como:

$$H = E[I_i] = -\sum_{i=0}^{L-1} p_i \log_e p_i$$

Sea T el valor del umbral, se definen las siguientes cuatro entropías parciales:

$$\begin{aligned} H_b &= -\sum_{i=0}^T p_i \log_e p_i & H'_b &= -\sum_{i=0}^T p_i \log_e p_i \\ H_n &= -\sum_{i=T+1}^{L-1} p_i \log_e p_i & H'_n &= -\sum_{i=T+1}^{L-1} p_i \log_e p_i \end{aligned}$$

Las dos entropías de la columna izquierda son una medida objetiva de la cantidad de información *a priori* asociada con los puntos blancos y negros relacionados con

el nivel T del umbral (en este contexto, *a priori* significa que los píxeles aún no han sido transformados a los niveles blanco y negro). Las dos entropías de la columna derecha H'_b y H'_n son las medidas de la cantidad de información *a posteriori* asociada con los puntos blancos y negros después de la operación de segmentación.

El problema es encontrar un umbral que nos de un resultado con la mayor cantidad de semántica aplicable. Primero, uno podría pensar en realizar esto maximizando la entropía *a posteriori* H' ; pero no tendría sentido maximizar medidas *a posteriori* debido a las propiedades estadísticas del histograma. Por lo que el autor propone que conociendo la entropía *a priori* del histograma H , se determine el umbral óptimo maximizando el límite superior de la entropía determinada *a posteriori* H' como una función paramétrica del nivel de umbral T

$$H' = H'_b + H'_n$$

6. *Método de Conservación del Momento.* En este método [46], el valor del umbral se calcula determinísticamente, de manera que los momentos de la imagen a ser segmentada se conservan en la imagen de salida (imagen binaria). El i -ésimo momento m_i se calcula como

$$m_i = \frac{1}{n} \sum_{g=0}^{L-1} g^i h(g), \quad i = 1, 2, 3,$$

donde n es el número total de píxeles en la imagen. El valor del umbral óptimo T^* , se obtiene del histograma de niveles de gris de la imagen eligiendo T^* como en el método p_0 -tile, donde p_0 esta dado por

$$p_0 = \frac{z - m_1}{(c_1^2 - 4c_0)^{1/2}}$$

y

$$c_0 = \frac{m_1 m_3 - m_2^2}{m_2 - m_1^2}, \quad c_1 = \frac{m_1 m_2 - m_3}{m_2 - m_1^2},$$

$$z = \frac{1}{2} \left\{ (c_1^2 - 4c_0)^{1/2} - c_1 \right\}$$

Específicamente, antes de segmentar, se calculan los momentos de los niveles de gris de la imagen de entrada. Los umbrales son entonces seleccionados de manera que los momentos de la imagen de salida se mantengan sin cambio. Este método puede ser considerado como una transformación de conservación del momento que recupera una imagen ideal binaria a partir de una versión borrosa. Este método puede seleccionar automática y determinísticamente múltiples umbrales sin necesidad de iteraciones o búsquedas. Además, se puede obtener también un nivel de gris significativo para cada clase.

7. *Método del Mínimo Error.* La idea principal de este método [25] es optimizar una función criterio relacionada con el promedio de error de la clasificación de los píxeles. En este método el histograma de niveles de gris, es una estimación de la función de densidad de probabilidad $p(g)$ de la mezcla de los píxeles de las poblaciones correspondientes a los objetos y al fondo. Generalmente se asume que cada una de estas dos componentes $p(g|i)$ de la mezcla están normalmente distribuidas con promedio μ_i , desviación estándar σ_i y una probabilidad *a priori* P_i . Este método propone que el valor del umbral puede ser seleccionado minimizando la función criterio $J(T)$ que está dada por:

$$J(T) = -1 + 2 \{P_1(T) \log_e \sigma_1(T) + P_2(T) \log_e \sigma_2(T)\} \\ - 2 \{P_1(T) \log_e P_1(T) + P_2(T) \log_e P_2(T)\}$$

Para un valor de umbral dado T , esta función criterio refleja la cantidad de traslape entre los modelos gaussianos de las poblaciones correspondientes a los objetos y al fondo. Mientras el valor de T varia, los modelos de la poblaciones cambian. El mejor ajuste entre los datos y los modelos conlleva a un menor traslape entre las funciones de densidad y por lo tanto a un menor error de clasificación.

Técnicas globales dependientes de una región

1. *Métodos de Transformación del Histograma.* El objetivo de estos métodos es transformar el histograma de niveles de gris en uno que contenga valles más profundos y picos más agudos de manera que pueda ser aplicado el método de moda descrito anteriormente. El nuevo histograma se obtiene asignándole peso a los píxeles de acuerdo a alguna propiedad local.

Mason [30] propuso el uso de un *operador borde* (por ejemplo el Laplaciano o el Roberts) para asignar los pesos. Los valores del *operador borde* son pequeños para píxeles dentro de regiones homogéneas, a estos píxeles se les asigna mayor peso. Los valores del *operador borde* son grandes para píxeles en la vecindad de un borde, a éstos se les asigna menor peso.

Otro método de transformación del histograma es el conocido como *método de quadtree* [48]. Este método se basa en que la desviación estándar de una región homogénea es pequeña, mientras que en regiones no homogéneas es grande. Iniciando con la imagen original, el método divide la imagen en cuadrantes, si las desviaciones estándar de cada subregión sobrepasan un cierto valor predefinido entonces éstas vuelven a subdividirse en cuadrantes. Esto conduce a una división de la imagen en bloques cuyas desviaciones estándar son pequeñas. Debido a la homogeneidad de cada bloque el histograma de niveles de gris resultante tendrá picos más agudos y valles más profundos.

2. *Métodos basados en Estadística de niveles de gris de Segundo Orden.* Una de las desventajas de los métodos de umbral dependientes de un punto, es que dependen solamente de la estadística de los niveles de gris de primer orden (el histograma). Los métodos que se describen a continuación, dependen de la estadística de los niveles de gris de segundo orden.

(i) *Método de la Matriz de Coocurrencia.* La matriz de coocurrencia es introducida por Haralick para el análisis de texturas [17]. De manera general la matriz de coocurrencia es aquella cuyos elementos son las frecuencias relativas de ocurrencia de dos píxeles vecinos con niveles de gris i y j , separados a una distancia d y con una orientación ϕ , este método lo describiremos con mayor detalle en el capítulo 2. En el trabajo de Ahuja y Rosenfeld [1], la matriz de coocurrencia de niveles de gris se define como:

$$M = M_{(1,0)} + M_{(1,\pi/2)} + M_{(1,\pi)} + M_{(1,3\pi/2)}$$

esto es, el elemento (i, j) de M es la frecuencia de ocurrencia de un nivel de gris i como vecino en conectividad 4 de un nivel de gris j .

Debido a su homogeneidad, los píxeles en el interior de los objetos o del fondo contribuyen principalmente en los elementos cercanos a la diagonal principal de M . Los píxeles cerca de algún borde contribuyen principalmente fuera de los elementos de la diagonal de M . Por lo tanto la matriz M puede ser usada para definir dos nuevos histogramas: (a) el histograma basado en los elementos cercanos a la diagonal principal, que tendrá valles profundos entre los niveles de gris de los objetos y del fondo, y (b) el histograma basado en los elementos fuera de la diagonal principal, que tendrá picos agudos entre los niveles de gris de los objetos y del fondo. El umbral de la imagen puede ser elegido dentro del rango de niveles de gris en donde el valle de (a) se traslape con el pico de (b).

(ii) *Método de la Gráfica Dispersa (nivel de gris, nivel de gris promedio local).* Este método es similar al descrito anteriormente. Se construye una gráfica dispersa *nivel de gris* contra *nivel de gris promedio local* tomada sobre una ventana. Se define el origen de la gráfica sobre la esquina superior izquierda de la ventana, el *nivel de gris* incrementa de izquierda a derecha, y el *nivel de gris promedio local* incrementa de arriba hacia abajo. La intensidad de un punto en esta gráfica es proporcional a la frecuencia de ocurrencia de un par correspondiente (*nivel de gris, nivel de gris promedio local*). En el trabajo de Kirby y Rosenfeld [24], los promedios locales son tomados de ventanas de 3×3 . Una vez construida la gráfica se tienen dos histogramas como se describen en los inciso (a) y (b) del método anterior. El umbral se establece de la misma manera.

3. *Método de Deravi y Pol.* En este método [10] se utilizan dos matrices de transición, que son similares a la matriz de coocurrencia, para definir dos "medidas de interac-

ción” para la selección del umbral. Sea M_{ij} una matriz con elementos (i, j) y T el umbral que divide al conjunto de niveles de gris en dos clases: $C_0 = \{0, 1, \dots, T\}$ y $C_1 = \{T+1, T+2, \dots, L-1\}$. Entonces M_{ij} puede ser dividida en cuatro regiones definidas por los siguientes parámetros:

$$a = \sum_{i=0}^T \sum_{j=0}^T M_{ij}, \quad b = \sum_{i=T+1}^{L-1} \sum_{j=T+1}^{L-1} M_{ij},$$

$$c = \sum_{i=0}^T \sum_{j=T+1}^{L-1} M_{ij}, \quad d = \sum_{i=T+1}^{L-1} \sum_{j=0}^T M_{ij},$$

donde a , b , c y d representan el número total de transiciones dentro de C_0 , dentro de C_1 , desde C_0 a C_1 y desde C_1 a C_0 , respectivamente. La probabilidad conjunta y condicional de las transiciones entre C_0 y C_1 puede ser estimada como

$$P_j(T) = \frac{c + d}{a + b + c + d}$$

$$P_c(T) = \frac{1}{2} \left\{ \frac{c}{a + c} + \frac{d}{b + d} \right\}$$

Los autores llaman a $P_j(T)$ y a $P_c(T)$ medidas de interacción. El umbral óptimo se obtiene de maximizar cualquiera de estas dos medidas de interacción.

4. *Métodos de Relajación.* La idea de la relajación fue introducida por Southwell [45] para mejorar la convergencia de las soluciones recursivas de sistemas de ecuaciones lineales. En segmentación de imágenes la relajación se aplica como sigue [41]: Los píxeles de una imagen primero se clasifican probabilísticamente en clases “claras” y “oscuras” basadas en sus niveles de gris. Luego la probabilidad de cada píxel es ajustada de acuerdo a las probabilidades de sus píxeles vecinos. Este proceso de ajuste es iterativo de manera que las “probabilidades claras” (respectivamente, “probabilidades oscuras”) se vuelven muy altas para píxeles que pertenecen a las regiones claras (respectivamente regiones oscuras). Y así, se forman histogramas con valles profundos.
5. *Métodos de Relajación de Gradiente.* En la relajación de gradiente, el esquema óptimo de etiquetación se determina maximizando una función criterio con optimización de gradiente. Sea λ_1 y λ_2 etiquetas de las clases de los píxeles claros y oscuros respectivamente, sea $\{[p_i(\lambda_1), p_i(\lambda_2)]^T, i = 0, 1, \dots, L-1\}$ el conjunto de vectores de probabilidad asociados con los niveles de gris, y sea $\{[q_i(\lambda_1), q_i(\lambda_2)]^T, i = 0, 1, \dots, L-1\}$ el conjunto de vectores de compatibilidad, donde $q_i(\lambda_k) = \frac{1}{8} \sum_{x_j \in N_i} p_j(\lambda_k)$, N_i es el ocurrencia de vecinos con conectividad 8 del píxel x_i . Baham y Feingers [2] sugirieron que el esquema óptimo de etiquetación para imágenes

con distribución de gris unimodal puede ser determinado maximizando

$$C(p) := \sum_{i=0}^{L-1} p_i^T(\lambda) q_i(\lambda)$$

sujeto a

$$p_i(\lambda) \in K := \left\{ p(\lambda) = (p(\lambda_1), p(\lambda_2)) \mid p(\lambda_j) \geq 0, \sum_{j=1}^2 p(\lambda_j) = 1 \right\}$$

Todos los *métodos globales* descritos hasta aquí calculan un umbral T único para segmentar toda la imagen, lo que produce resultados deficientes en las fronteras de los objetos. Con el método robusto que nosotros proponemos tratamos de resolver precisamente este problema.

Técnicas locales

En la segmentación por umbral local, la imagen original es dividida en subimágenes pequeñas y se determina un umbral para cada subimagen. Esto da como resultado una imagen segmentada con discontinuidades de los niveles de gris en las fronteras de dos subimágenes diferentes. El umbral de una región puede ser determinado por algún método dependiente de un punto o bien por un método dependiente de una región. Se aplica después una técnica de suavizamiento para eliminar las discontinuidades.

Por ejemplo, Chow y Kaneko [8] dividieron la imagen original en subimágenes de 7×7 y calcularon un umbral para cada una de ellas. Los umbrales para subimágenes unimodales se interpolaron de subimágenes vecinas. Los umbrales para subimágenes bimodales se calcularon como sigue: primero el histograma de los niveles de gris de la subimagen se aproxima a la suma de dos distribuciones gaussianas, luego se obtiene el umbral minimizando el error de clasificación respecto al valor del umbral.

Métodos Multimbral

Los métodos de umbral globales como el de Otsu [33], el de Pun [38], el de conservación del momento [46], el del mínimo error [25], pueden ser generalizados para el caso de multimbral. En adelante discutiremos otros métodos multimbral que no han sido incluidos en esta sección.

1. *Método de Segmentación por Amplitud.* Este método propuesto por Boukharouba [3], utiliza las propiedades intrínsecas de la función de distribución acumulada de la imagen a ser segmentada. Se examina la curvatura de la función de distribución

para obtener los valores de los umbrales. La función de distribución $F(k)$ en el punto k está dada por

$$F(k) = \frac{\sum_{g=0}^k h(g)}{\sum_{g=0}^{L-1} h(g)}$$

La curvatura de F se define como

$$C(x) = F''(x) [1 + (F'(x))^2]^{-3/2}$$

donde F' y F'' son la primera y la segunda derivada de F , respectivamente. $C(x)$ resulta ruidosa y debe ser suavizada y aproximada para ser utilizada en el cálculo de los umbrales. Los ceros de la curvatura determinan los umbrales así como los niveles de gris para ser asignados a cada clase.

2. *Método de Wang y Haralick.* Esta es una técnica recursiva [47]; los píxeles se clasifican primero como píxeles borde o no borde. Luego los píxeles borde son clasificados, en base a sus vecinos, como relativamente oscuros o relativamente claros. Se obtiene el histograma de los niveles de gris para aquellos píxeles que son píxeles borde y relativamente oscuros, y se obtiene un segundo histograma para aquellos píxeles que son píxeles borde y relativamente claros. Se selecciona un umbral basado en la intensidad del nivel de gris correspondiente al pico mayor para cada histograma. Para obtener múltiples umbrales, este procedimiento se aplica recursivamente utilizando sólo aquellos píxeles cuyas intensidades son menores al primer umbral inicial, y luego utilizando aquellos píxeles cuyas intensidades son mayores que el último umbral obtenido en cada pasada.
3. *Método de Contraste Uniforme.* Esta es otra técnica recursiva propuesta por Kohler [26]. El umbral óptimo para la segmentación de una imagen es aquel que detecta más bordes con altos contrastes y menos bordes con bajos que cualquier otro umbral. En este método se construye un histograma del contraste promedio $\mu(T)$ para cada posible umbral T , y el pico más alto en el histograma corresponde a el umbral óptimo. El contraste promedio $\mu(T)$ se calcula de la relación

$$\mu(T) = \frac{C(T)}{N(T)}$$

con $\mu(T) = 0$ si $N(T) = 0$. Donde $C(T)$ es el total de contraste detectado por el umbral T , y $N(T)$ es el número de bordes detectados por T . Para encontrar múltiples umbrales se selecciona cualquier umbral inicial y se recalcula un nuevo histograma de $\mu(T)$ quitando la contribución de los bordes ya detectados por el umbral

inicial. Este procedimiento continua hasta que el máximo contraste promedio para cualquier umbral, cae por debajo de algún criterio de contraste promedio mínimo $\theta > 1$.

1.1.2 Segmentación por cúmulos

El agrupamiento o acumulación de rasgos característicos aplicados a la segmentación de imágenes es una extensión multidimensional del concepto de segmentación por umbral. Se utilizan típicamente dos o más rasgos característicos dispuestos en un plano cartesiano, es decir un eje por rasgo, y cada una de las clases o regiones se asume que formarán un grupo o cúmulo de puntos distinto en el espacio. Sobre este espacio de rasgos podrán encontrarse los umbrales o límites de decisión que separan a los cúmulos, como se muestra en el ejemplo de la Figura 1.2 [14]. Estos cúmulos son después mapeados al dominio espacial original para producir la segmentación de la imagen.

Los rasgos característicos que son comúnmente utilizados para la segmentación de imágenes por cúmulos no sólo incluyen los valores de gris de cada pixel sino también pueden incluir cualquier otro tipo de rasgo que se crea útil para la segmentación como por ejemplo medidas de textura definidas en una vecindad [5].

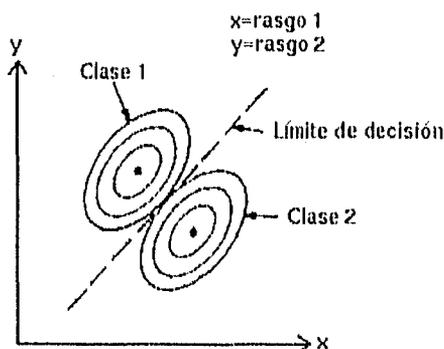


Figura 1.2: Ejemplo de Segmentación por Cúmulos. Espacio bidimensional de rasgos.

Lo difícil en esta técnica es definir los rasgos característicos que nos den información sobre las regiones.

Panda y Rosenfeld [34] sugirieron un método para segmentar glóbulos blancos sobre un fondo oscuro utilizando segmentación por cúmulos.

Considérese el histograma de niveles de gris para todos los pixeles que tienen gradientes pequeños. Si un pixel tiene gradiente pequeño, seguramente no es candidato a ser un borde. Si no es un borde, entonces será un pixel oscuro perteneciente al fondo, o un pixel claro perteneciente a un glóbulo blanco. Por lo tanto el histograma de todos los

pixeles que tienen gradientes pequeños será bimodal, y para los pixeles con gradientes grandes, el valle entre las dos modas será un buen punto de umbral.

Ahora considérese el histograma de niveles de gris para todos los pixeles con gradientes grandes. Si un pixel tiene gradiente grande, seguramente será un borde. Si es un borde, entonces separa el fondo oscuro de un glóbulo blanco claro, y si la frontera que los separa no es intensa pero de alguna manera es difusa, entonces el histograma será unimodal.

El objetivo es encontrar un buen umbral que separe los pixeles oscuros del fondo de los pixeles claros de los glóbulos blancos. Por lo tanto los autores sugirieron determinar dos umbrales: uno para pixeles de gradientes pequeños y otro para pixeles de gradientes grandes. Para esto desarrollaron un método de segmentación por cúmulos basado en un espacio bidimensional, en donde en un eje se encuentran las intensidades de los niveles de gris y en el otro eje el valor del gradiente. La forma del límite de decisión en este espacio bidimensional se muestra en la Figura 1.3.

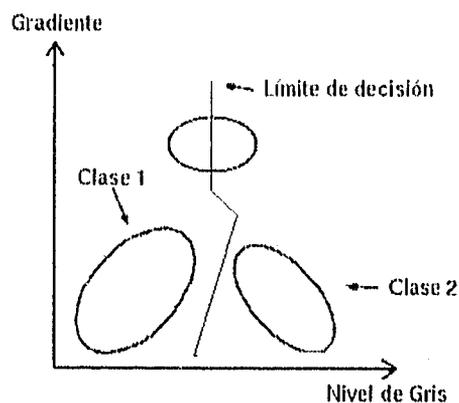


Figura 1.3: La clase 1 corresponde al fondo (tono oscuro y gradiente pequeño), la clase 2 corresponde a los glóbulos blancos (tono claro y gradiente pequeño), y el último óvalo corresponde a los bordes difusos (gradiente grande).

1.2 Detección de bordes

La detección de bordes es una técnica de segmentación basada en la detección de discontinuidades. Un borde o límite es la zona donde la imagen tiene cambios abruptos en el nivel de gris. Algunos de los factores motivantes para el estudio de estas técnicas son: 1) la mayoría de la información en una imagen cae entre los bordes de las diferentes regiones, y 2) los sistemas biológicos visuales aparentemente hacen uso de la detección de bordes, no así de la de umbrales. Se han categorizado las técnicas de detección de bordes de la siguiente manera [9]:

- **Técnicas Paralelas**

1. *Extracción del elemento borde.*
2. *Combinación de elementos borde.*

- **Técnicas Secuenciales**

1.2.1 Técnicas Paralelas

Por técnicas paralelas de detección de bordes se entiende que la decisión de si un conjunto de puntos pertenece o no a un borde se hace en base a el nivel de gris del conjunto y al de algún conjunto de sus vecinos; no depende en principio si otros conjuntos de puntos no vecinos caen o no en el borde. El operador de detección de bordes puede ser aplicado simultáneamente en cualquier lugar de la imagen.

Los métodos de detección paralelos pueden dividirse en dos pasos:

1. *Extracción del elemento borde.*

- *Filtros en el dominio de altas frecuencias.* Como las frecuencias altas están asociadas con cambios agudos en la intensidad, uno puede resaltar o extraer los bordes aplicando un filtro pasa-altas a una imagen, por ejemplo la transformada de Fourier. Se aplica la transformada de Fourier directa a la imagen, estando en el dominio de las frecuencias se filtran las altas frecuencias, al resultado del filtro se le aplica la transformada de Fourier inversa para regresar al dominio original; el resultado de lo anterior es una imagen con bordes realzados. El problema aquí reside en el diseño del filtro.
- *Operadores gradiente.* Un buen número de técnicas de detección de bordes propuestas en la literatura están basadas en la versión digital del operador gradiente las cuales producen una magnitud alta cuando existen cambios abruptos en el nivel de gris y magnitudes bajas cuando existen pequeños cambios en el nivel de gris, por lo que sólo bordes muy agudos con altos contrastes entre los objetos son detectados por este método. Existen otros operadores como el Kirsch, Sobel, Prewitts que están basados en vecindades de 3×3 ; la principal diferencia entre estos operadores es el peso que se asigna a cada elemento de la plantilla de 3×3 .
- *Aproximaciones funcionales.* La detección de bordes puede ser considerada como un problema de aproximaciones [21]. El objetivo de este método es aproximar de la mejor manera un elemento borde empírico cuyos valores de gris están en $f(x, y)$ a un elemento borde ideal caracterizado por una función F , por ejemplo una función escalón. Esta técnica es empleada también para la detección de líneas [22]. Este método funciona mejor que los anteriores especialmente en imágenes con bordes de bajo contraste.

2. *Combinación de elementos borde.*

La combinación de elementos borde consiste en eliminar los falsos elementos bordes y mezclar elementos borde en segmentos borde mayores llamados rayas o líneas.

- *Búsqueda heurística.* La búsqueda heurística es una técnica que utiliza métodos de búsqueda en el espacio de estados donde la información heurística es usada para delimitar el espacio de búsqueda. Se ha formulado [28] [29] el problema de la detección de bordes como una búsqueda heurística de la ruta más corta en una gráfica. Los nodos de la gráfica (o estados) son elementos borde definidos por dos píxeles vecinos, por ejemplo los puntos $A = (i, j)$, $B = (i, j + 1)$ definiendo el elemento borde dirigido AB . Después se establece que un borde es una secuencia de elementos borde adyacentes, por lo que un borde es una ruta en una gráfica que representa el espacio de estados y el problema de encontrar el mejor borde en una imagen se reduce a encontrar la ruta óptima en la gráfica.
- *Relajación.* Se han implementado técnicas de relajación para conectar elementos bordes [40]. Esta técnica es un proceso iterativo en donde la probabilidad de que un elemento borde candidato sea un verdadero elemento borde es re-estimada en cada iteración. Algunas de las ventajas de esta técnica son que es un proceso paralelo y utiliza información espacial. Algunas desventajas son que la construcción de la función de compatibilidad que es la que actualiza las probabilidades de los elementos borde no es trivial y la rapidez de convergencia es generalmente lenta.
- *Ajuste de líneas y curvas.* Otra técnica para conectar elementos borde es ajustándolos a líneas o a curvas. Supóngase que tenemos un conjunto de n puntos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ que representan elementos borde, lo que queremos es encontrar líneas rectas que se ajusten a ellos; por ejemplo utilizando la transformada de Hough [12].

Existen otras técnicas [39] de detección de bordes basadas en ajustes como el ajuste a plantillas, que pueden ser aplicadas también en otras áreas como por ejemplo la detección de objetos.

1.2.2 Técnicas Secuenciales

Por técnicas secuenciales de detección de bordes se entiende que el resultado sobre un punto es dependiente de los resultados de aplicar el operador en puntos previamente examinados. Los componentes principales en el procedimiento de detección de bordes secuencial son:

1. La elección del punto inicial correcto: la ejecución del procedimiento en su totalidad dependerá en la elección de un buen punto de inicio.

2. La dependencia de estructura: cómo afectan los resultados obtenidos en puntos previamente examinados tanto para la elección del siguiente punto a ser examinado como para su resultado.
3. El criterio de paro: debe haber una manera para que el procedimiento determine que se ha llegado al final.

Hay un gran número de técnicas secuenciales que también utilizan búsqueda heurística y programación dinámica [7] [23].

En general, las *técnicas de detección de bordes* son poco populares en la segmentación de imágenes, porque poder llegar a un conjunto de bordes significativos resulta a veces una labor muy complicada, y además existe el problema después de pegar bordes.

1.3 Extracción de regiones

Las técnicas de extracción de regiones se conocen también como técnicas de crecimiento de regiones y se refieren al proceso de segmentación secuencial de una imagen en donde los píxeles son sucesivamente añadidos a regiones incompletas o inician nuevas regiones cuando no es apropiado hacerlos parte de una región incompleta existente. Zucker [49] a clasificado las técnicas de crecimiento de regiones de la siguiente manera:

- **Búsqueda Regional de Vecinos.**
- **Heurísticas Multirregionales.**
- **Aproximación Funcional y Mezcla.**
- **Esquemas de División y Unión.**
- **Interpretaciones Regionales y Semánticas.**

1.3.1 Búsqueda Regional de Vecinos

Mierle y Allen [32] definieron una región como una porción de una imagen compleja en la cual la estadística de la distribución de los niveles de gris es razonablemente uniforme. Su método puede ser descrito en términos de los siguientes pasos:

- (i) Se segmenta la imagen completa en celdas de por ejemplo de 2×2 , 4×4 , o 8×8 píxeles.
- (ii) Se calculan las medidas estadísticas de las intensidades de cada celda, por ejemplo se estima la distribución de probabilidades de los niveles de gris.

- (iii) Iniciando por la primera celda en la esquina superior izquierda de la imagen, se comparan sus estadísticas con las de sus celdas vecinas y se determina si son similares. El criterio utilizado como medida de similitud fue la prueba de Kolmogorov-Smirnov y una diferencia absoluta promedio máxima entre las distribuciones (áreas). Si es exitosa se mezclan las celdas para formar un fragmento. Si no, se etiquetan las celdas distintas como rechazadas. Se recalcula la función de densidad de probabilidad para el fragmento estimado.
- (iv) Continúa el crecimiento del fragmento examinando todos los vecinos hasta que no queden vecinos que puedan ser unidos. Etiquetando el fragmento como una región completa.
- (v) Se posiciona en la siguiente celda incompleta, y se repiten los pasos anteriores hasta que todas las celdas estén etiquetadas.

En este método sólo se utiliza la información de los niveles de gris para realizar el proceso de crecimiento.

1.3.2 Heurísticas Multirregionales

El establecer los predicados que controlan el proceso de crecimiento puede involucrar consideraciones locales y globales. Los vecinos son unidos por un proceso regional o local, pero lo que realmente se quiere al final del algoritmo es un resultado global satisfactorio. Por lo que el problema consiste en utilizar la mayor información global posible para poder tomar decisiones locales.

La propuesta de Brice y Fennema [4] fue obtener una aproximación más global desarrollando heurísticas que evalúen parámetros dependientes de más de una región, con el objetivo de delinear fronteras. Comenzaron dividiendo la imagen en segmentos iniciales con píxeles que tienen intensidades idénticas llamadas regiones "atómicas" (que podrían llamarse píxeles semillas), insertando segmentos elementales de frontera (vectores) entre los píxeles de diferente intensidad. Las regiones atómicas en esta primera fase son combinadas con la aplicación sucesiva de dos heurísticas. Estas heurísticas utilizan información birregional contenida en las fronteras de pares de regiones adyacentes.

La primera heurística mezcla dos regiones adyacentes si la frontera entre ellas es débil, por lo que la nueva región resultante tiene una frontera menor que las dos regiones previas. La segunda heurística une dos regiones si la parte débil su frontera en común es igual a algún porcentaje determinado del total de la frontera que comparten. Comparando estas dos heurísticas, la primera incluye la idea natural de borrar fronteras débiles mientras que la segunda evalúa estas acciones según sus efectos sobre ciertos criterios de forma.

Como paso final los vectores frontera elementales que sobreviven se conectan y se suavizan. Esto se realiza uniendo con líneas rectas las cadenas de segmentos.

1.3.3 Aproximación Funcional y Mezcla

La aproximación funcional utiliza un conjunto dado de funciones conocidas y fáciles de manipular para aproximar otro conjunto de funciones que son, o bien complejas o que sólo se conocen en algunos puntos discretos.

Sean F y Φ dos familias de funciones conocidas en un dominio D . Las funciones en F son aproximadas por aquellas de Φ de manera que la diferencia (error) entre ellas caiga dentro de algún límite tolerable. Para poder realizar esto, D se divide en M regiones D_i , $i = 1, 2, \dots, M$ de manera que para ϕ_i definida en D_i ,

$$E(f, \phi_i) \leq \epsilon, \quad \phi_i \in \Phi$$

donde E es un error adecuadamente establecido. Una vez que las funciones de aproximación Φ son seleccionadas, por ejemplo una constante o funciones lineales, el problema se traduce en encontrar las regiones locales apropiadas sobre las cuales una sola función de aproximación sea suficientemente buena. Esto se realiza iniciando con muchas regiones pequeñas, cada una aproximada por una función de la misma forma, y mezclando aquellas regiones cuyos coeficientes de aproximación sean lo suficientemente cercanos.

Pavlidis [35] utilizó esta técnica sobre imágenes bidimensionales. El método rebana la imagen original en tiras delgadas de manera que cada tira pueda ser considerada como una función de una sola variable, por ejemplo $f(x, y_0)$. Ahora estas tiras pueden ser aproximadas encontrando la mejor partición de cada tira en J segmentos. Cada segmento debe estar cercano a una norma de error adecuada.

Estos segmentos aproximados se mezclan en regiones. Lo anterior se puede expresar en una gráfica en donde los segmentos representan los vértices, y los bordes representan las diferencias entre de los coeficientes de los segmentos con fronteras traslapadas. El algoritmo recorre esta gráfica, etiquetando los nodos idénticos si están conectados por un borde cuyo peso está por debajo de un cierto umbral. De esta manera, las regiones se forman de los segmentos originales, si sus coeficientes de aproximación son lo suficientemente iguales.

1.3.4 Esquemas de División y Unión

El principio del método de “división y unión”, desarrollado por Pavlidis y Horowitz [20], es unir regiones adyacentes que tengan aproximaciones similares y dividir aquellas regiones que tengan grandes normas de error.

El algoritmo inicia realizando una primera segmentación llevando a cabo una división de la imagen original en bloques rectangulares de un tamaño específico. La imagen es representada en un *árbol de segmentación*, el cual está basado en una estructura de datos de *quadtree* (un árbol cuyos nodos tienen cuatro hijos). La imagen completa se representa por el nodo raíz. Los hijos de la raíz son regiones obtenidas de dividir la raíz en cuatro rectángulos iguales, y así sucesivamente. Una segmentación se define como un *conjunto*

de corte, es decir un conjunto mínimo de nodos que separan la raíz de todas las hojas. En la estructura de árbol el proceso de unión consiste en quitar cuatro nodos del *conjunto de corte* y reemplazarlos por su padre. La división consiste en quitar un nodo del *conjunto de corte* y reemplazarlo por sus cuatro hijos. Los dos procesos son mutuamente excluyentes: todas las operaciones de unión están seguidas por operaciones de división. Los criterios de unión o división se basan en la intensidad de gris promedio de cada rectángulo, si caen dentro de un valor promedio de tolerancia se unen sino se dividen.

La división y unión en la estructura de árbol al final del proceso, es seguida por un procedimiento de agrupamiento que puede unir bloques adyacentes no relacionados encontrados en el *conjunto de corte* final. El agrupamiento se determina en base a los niveles de gris máximos y mínimos para cada región.

1.3.5 Interpretaciones Regionales y Semánticas

Los métodos de crecimiento de regiones presentados hasta aquí, pueden ser considerados de alguna manera como “sintácticos”, es decir, se basan en propiedades de la imagen como sus niveles de gris, la intensidad de sus bordes, o la forma de sus regiones. En los métodos semánticos se considera la información semántica sobre las clases que son procesadas. Esto se realiza interpretando las regiones formadas y permitiendo que estas interpretaciones tengan influencia en el criterio de crecimiento.

Feldman y Yakimovsky [13], realizaron un algoritmo de crecimiento de regiones basado en la teoría de decisiones. El principal objetivo del sistema es maximizar la probabilidad de que la imagen esté segmentada correctamente. Esta probabilidad está condicionada sobre el conocimiento contextual previo disponible en la imagen, y sobre las medidas tomadas de una región particular que esté siendo procesada. Más formalmente el objetivo es maximizar:

$$\begin{aligned}
 & P \{ \text{interpretación global} \mid \text{contexto, medidas} \} \\
 &= \prod_i P \{ R(i) \text{ en } INT(i) \mid \text{valores de las medidas sobre } R(i) \} \\
 & \times \prod_{i,j} P \{ B(i,j) \text{ entre } INT(i) \text{ y } INT(j) \mid \text{medidas sobre } B(i,j) \}
 \end{aligned}$$

El primer término de esta ecuación es el producto de las probabilidades de que cada región $R(i)$ tenga una interpretación $INT(i)$, dados los valores de las medidas sobre $R(i)$. Se asume que la correcta interpretación de cada región requiere medidas sólo de esa región. Esta suposición de independencia de medidas permite que la probabilidad conjunta sobre todas las regiones sea escrita como un producto de probabilidades para cada región.

Las regiones, sin embargo, no son completamente independientes. Las fronteras entre ellas deben tener una interpretación que sea consistente con las interpretaciones regionales. Por ejemplo, dadas dos regiones adyacentes con interpretaciones “cielo” y “tierra”, la

interpretación de la frontera debe reflejar el conocimiento de que el cielo normalmente esta sobre la tierra. Estas interpretaciones birregionales de frontera están expresadas en el segundo término de la ecuación anterior. Se asume también que las fronteras son independientes de influencias no locales a ellas, y por lo tanto, el segundo término también puede ser escrito como un producto.

Las técnicas estándar para encontrar un máximo real de la ecuación anterior deben involucrar una examinación exhaustiva de todas las posibilidades. Por lo tanto los autores desarrollaron una aproximación heurística para encontrar un pseudomáximo válido el cual debe ser cercano al máximo real. La técnica utiliza información semántica para determinar cuáles regiones deben ser unidas, así como determinar cuando debe parar el algoritmo de crecimiento. La complejidad de programar este sistema refleja la dificultad de utilizar una buena interacción entre técnicas sintácticas y semánticas.

Esta es la manera general como se categorizan los métodos de segmentación, seguramente existen muchos mas que aquí no se mencionan, o bien un enfoque diferente de organizarlos, pero el objetivo de este capítulo es dar una idea general de lo que existe.

Para la solución de un problema en particular no siempre se utiliza una sola técnica sino que suelen hacerse combinaciones de varias. El desarrollo del algoritmo propuesto en este trabajo no es la excepción; en el siguiente capítulo se hará la descripción detallada del algoritmo y podremos ubicarlo dentro de esta categorización.

Capítulo 2

Descripción del Algoritmo de Segmentación

Se propone un método para segmentar específicamente las imágenes de cortes histológicos de cerebro de rata, en donde queremos encontrar la corteza en las imágenes control, y el tejido lesionado en las imágenes experimentales. La Figura 2.1 muestra un ejemplo de este tipo de imágenes.



Figura 2.1: (a) Corte control. (b) Corte lesionado.

Se desecharon los métodos de umbral único porque producen errores en las fronteras de los diferentes objetos, inclusive los métodos multiumbrales.

La motivación principal de este trabajo fue diseñar un método que adapte localmente los umbrales entre las clases tomando en cuenta las condiciones de vecindad entre píxeles.

Los métodos desarrollados en este trabajo se basan principalmente en estadística robusta, que no puede ubicarse dentro de la categorización de métodos descritos anteriormente, debido a que es un área relativamente nueva; y también se basan en la utilización de una técnica de crecimiento de regiones que se puede ubicar dentro de las técnicas de

búsqueda regional de vecinos, específicamente conocida también como *agregación* cuyos predicados lógicos de agregación tienen una característica peculiar que describiremos más adelante.

Se decidió explorar los resultados de utilizar técnicas de estadística robusta principalmente por dos razones, 1) existe poca literatura al respecto en aplicaciones en segmentación de imágenes de histología, por lo que se trató de explorar y diseñar un nuevo método; y 2) es posible que algunos de los métodos descritos anteriormente funcionen; sin embargo en general, estos métodos hacen el cálculo de un sólo umbral y detectan dos clases; en algunos métodos, resulta complicado implantar la generalización cuando se requieren más de dos clases. En nuestro caso necesitamos encontrar tres clases en las imágenes control y cuatro clases en las imágenes experimentales. Con el método robusto que proponemos es posible encontrar un número variable de clases.

El algoritmo de segmentación propuesto consiste de dos pasos: (i) la estimación de los parámetros μ_i y σ_i de cada clase i , donde μ_i es el estimador de localización, σ_i es el estimador del ruido, con $i = 1, 2, 3, \dots$; y (ii) el crecimiento de regiones por agregación, el cual se basa en los valores de μ_i y σ_i para definir las semillas del crecimiento.

Específicamente el método consiste en:

1. Encontrar las μ_i de cada clase i , a partir de la información global contenida en el histograma de la imagen, por medio de un análisis robusto de cúmulos (**análisis espectral**).
2. Estimar las σ_i de cada clase i , a partir de la matriz de coocurrencia, utilizando también análisis robusto de cúmulos (**análisis espacial**).
3. Aplicar un método de crecimiento de regiones que utiliza las μ_i y σ_i encontradas de manera iterativa, usando técnicas de relajación, hasta que todos los píxeles de la imagen son clasificados. Lo interesante de este método de crecimiento de regiones por agregación, es que resuelve los conflictos de pertenencia de clase y relaja las condiciones en sus predicados de agregación.

A continuación se describirán las bases y los métodos implantados para cada uno de los incisos anteriores; en particular limitamos la explicación de los procesos para un número fijo de **tres clases**.

2.1 Estimación de las μ_i

2.1.1 Estadística Robusta

El análisis de regresión es una herramienta estadística importante y se aplica rutinariamente en muchas ciencias. Una de las técnicas de regresión más comúnmente utilizada

es el método de los mínimos cuadrados (least squares **LS**) que consiste, para el caso de una dimensión en donde el modelo es con una sola variable $y_i = \theta + e_i$, en usar como estimador de localización $\hat{\theta}$ a la *media*:

$$\hat{\theta}_{LS} = \frac{\sum_{i=1}^n y_i}{n}$$

que minimiza $\sum_{i=1}^n r_i^2$ donde $r_i = y_i - \hat{\theta}$ son los residuales, y como estimador del ruido a la desviación estándar

$$\hat{s}_{LS} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{\theta}_{LS})^2}{n-1}}$$

Este método de regresión generalmente fracasa cuando la muestra contiene datos aberrantes (“outliers”), cosa que en muestras reales ocurre frecuentemente, llevándonos a soluciones erróneas.

Para mostrar lo anterior, considérense las siguientes 10 lecturas de datos que corresponden, por ejemplo, a niveles de gris de los píxeles de una imagen:

90, 93, 86, 92, 95, 83, 75, 40, 88, 80.

En la Figura 2.2 se representan los datos en una recta numérica, en donde $\hat{\theta}_{LS} = 82.2$ con $\hat{s}_{LS} = 16.06$, y la mediana es 87.0. Nótese la influencia del dato 40 sobre el valor de la media que ocasiona que se aleje de la mayoría de los datos, por lo que estas medidas de tendencia central no son confiables cuando tenemos este tipo de datos.

En el caso de imágenes digitales, los píxeles en las fronteras de dos objetos tienden a contribuir como datos aberrantes.

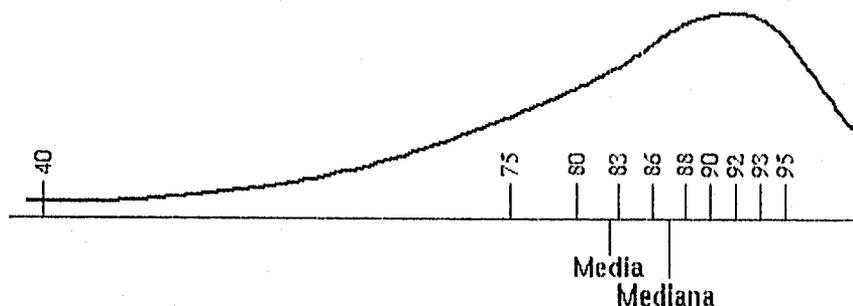


Figura 2.2: Representación en la recta numérica del LS.

Para solucionar este problema, se han desarrollado nuevas técnicas estadísticas que no son fácilmente afectadas por datos aberrantes. Estas técnicas se conocen como **métodos robustos** [43], cuyos resultados se consideran suficientemente confiables a pesar de que una cierta cantidad de datos en la muestra estén contaminados.

Un estimador robusto de localización es la *moda*, que corresponde al valor más probable de la función de distribución de densidad de probabilidad. La moda puede ser estimada a partir del LMS (least median of squares) [43], que está definido de la siguiente manera, $\hat{\theta}_{LMS}$ que:

$$\text{Minimiza } \text{med}_i r_i^2$$

El método consiste en encontrar el centro del intervalo más corto que contenga el 50% de los datos. Dicho intervalo corresponde a la región más densa de la distribución. El centro de dicho intervalo corresponde a la moda, $\hat{\theta}_{LMS}$.

Para ilustrar el método numéricamente considerense las mismas 10 lecturas que en ejemplo anterior, pero esta vez ordenadas, mostradas en la Figura 2.3. Con $n = 10$ y

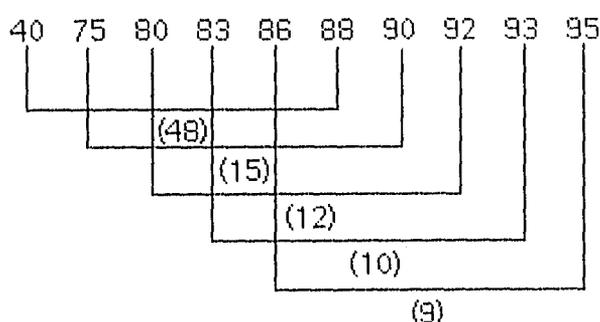


Figura 2.3: Cálculo del LMS.

$h = [10/2] + 1 = 6$; las líneas horizontales indican los intervalos que contienen la mitad de la muestra (el 50% de los datos, h), se calculan las diferencias que corresponden a los valores 48, 15, 12, 10 y 9 (tamaño del intervalo). La distancia más pequeña de ellas corresponde al intervalo entre 86 y 95. Finalmente el punto medio de este intervalo corresponde a 90.5 que es el $\hat{\theta}_{LMS}$ estimado.

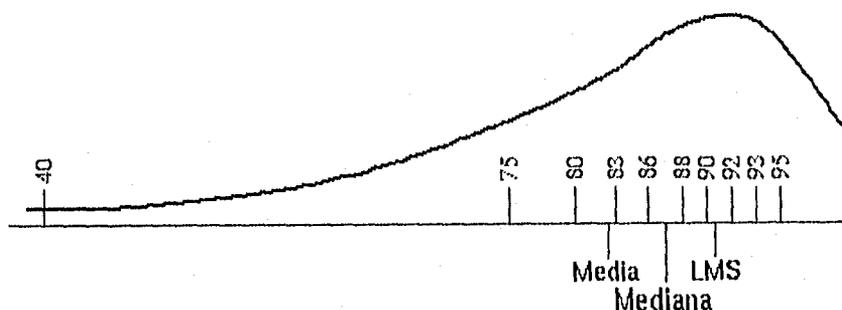


Figura 2.4: Representación en la recta numérica del LMS.

Nótese en la Figura 2.4 que la presencia del dato 40 no afecta al resultado del $\hat{\theta}_{LMS}$ a pesar de estar considerado en la estimación. La curva dibujada es una distribución hipotética a la que podrían pertenecer los datos del ejemplo. El lector se puede referir a [43] para una explicación más amplia de métodos de estimación para espacios de dimensión mayor.

2.1.2 Estimación de las μ_i . Implantación

Para la estimación de las μ_i de cada clase, nos basamos únicamente en la información disponible en el histograma de la imagen, encontrando los intervalos que las contienen. Se escogió utilizar el histograma ya que se trata de una función muy fácil y rápida de calcular.

Sea $h(j)$ el histograma de la imagen con $j = 0, 1, \dots, L$ y L el número de niveles de gris; sea $H = \sum_{j=0}^L h(j)$ la masa de la imagen, y sea $h_c(k)$ con $k = k_{c1}, \dots, k_{c2}$ un intervalo de h para los valores de gris dentro del rango $[k_{c1}, k_{c2}]$. Asumimos que las modas en h corresponden a los valores de gris más frecuentes de las diferentes regiones. Estas se identifican iterativamente de la siguiente manera:

- Paso 1. Se localiza el intervalo $h_c(k)$ más compacto en h que contenga qH datos [31]. La compacidad del intervalo $h_c(k)$ se define como:

$$\text{compacidad} = \frac{\sum_{k=k_{c1}}^{k_{c2}} h_c(k)}{k_{c2} - k_{c1}}$$

El intervalo óptimo es aquel para el cual con dos valores consecutivos de q , $q = 0.50, 0.45, 0.40, \dots, 0.15$, el centro del intervalo es el mismo, o bien si q alcanza el valor de 0.15.

- Paso 2. Se eliminan los datos del histograma que están contenidos en el intervalo óptimo $[k_{c1}, k_{c2}]$ encontrado en el Paso 1, $h(k) = 0$ para $k = k_{c1}, \dots, k_{c2}$, y se asume el centro de este intervalo como el estimador $\hat{\theta}_{LMS}$ del valor promedio θ_i para la región i .
- Paso 3. Se recalcula H y se vuelve al Paso 1 hasta que se encuentran las tres modas. Esto se puede generalizar para más modas.

De haber estimado las modas utilizando algún método de umbral podríamos llegar más fácilmente a errores, ya que cuando tenemos clases que espectralmente están muy cercanas, los métodos de umbrales fracasan aunque el histograma contenga modas significativas; por lo que aproximarnos encontrando intervalos compactos que contienen a la mayoría de los datos, y luego eliminándolos del histograma hace que el cálculo sea más exacto a las modas de clase, aunque éstas estén muy cercanas. Claro que hasta cierto

límite este método de estadística robusta puede también fallar, cuando dos clases están totalmente mezcladas en una sola curva gaussiana, es decir, cuando las modas no son lo suficientemente significativas.

2.2 Estimación de las σ_j

Debido a que la información disponible en el histograma de la imagen es muy pobre por ser únicamente espectral y por ser además una mezcla de las funciones de densidad de probabilidad de las clases, es necesario realizar un análisis espacial que nos de más información sobre la caracterización de cada región de manera que podamos obtener parámetros más confiables, en adición a las μ_i obtenidas. Lo que queremos hacer es aislar la función de densidad de probabilidad de cada clase sacándola de la imagen, de manera que podamos calcular las σ_i de las nuevas distribuciones independientes. Para ésto, utilizamos una variante de la matriz de coocurrencia.

2.2.1 Matriz de Coocurrencia

El cálculo de la matriz de coocurrencia está basado en las dependencias espaciales sobre los tonos de gris. La coocurrencia puede ser definida como la matriz de las frecuencias relativas P_{ij} con las que dos pixeles vecinos separados por una distancia d aparecen en una imagen, uno con nivel de gris i y el otro con nivel de gris j [17]. Dichas matrices de dependencia espacial de niveles de gris son simétricas y están en función de la relación angular entre pixeles vecinos así como en función de su distancia. Formalmente, para ángulos a intervalos de 45° , las frecuencias no normalizadas están definidas como sigue:

$$P(i, j, d, 0^\circ) = \#\{ [(k, l), (m, n)] \mid \\ k - m = 0, \quad |l - n| = d \\ I(k, l) = i, \quad I(m, n) = j \}$$

$$P(i, j, d, 45^\circ) = \#\{ [(k, l), (m, n)] \mid \\ (k - m = d, \quad l - n = -d) \\ \text{o} \quad (k - m = -d, \quad l - n = d) \\ I(k, l) = i, \quad I(m, n) = j \}$$

$$P(i, j, d, 90^\circ) = \#\{ [(k, l), (m, n)] \mid \\ |k - m| = d, \quad l - n = 0 \\ I(k, l) = i, \quad I(m, n) = j \}$$

$$P(i, j, d, 135^\circ) = \#\{ [(k, l), (m, n)] \mid \\ (k - m = d, \quad l - n = d) \\ \text{o} \quad (k - m = -d, \quad l - n = -d) \\ I(k, l) = i, \quad I(m, n) = j \}$$

donde # denota el número de elementos en el conjunto. Como ya se dijo estas matrices son simétricas: $P(i, j; d, a) = P(j, i; d, a)$.

Para ilustrar lo anterior con un ejemplo, considere que la Figura 2.5(a) representa una imagen de 4×4 píxeles con cuatro niveles de gris que van de 0 al 3. La Figura 2.5(b) muestra la forma general de la matriz de dependencia espacial de niveles de gris. Por ejemplo el elemento en la posición (2,1) de la matriz horizontal P_H a distancia $d = 1$ representa el total de veces que dos niveles de gris de valor 2 y 1 ocurren horizontalmente adyacentes uno del otro. Para determinar este número, se cuenta el número de pares de píxeles en la imagen cuyos píxeles 2 tengan un vecino horizontal 1 y viceversa. En la Figura 2.5(c) a (f) están calculadas las cuatro matrices de dependencia espacial de niveles de gris a distancia $d = 1$.

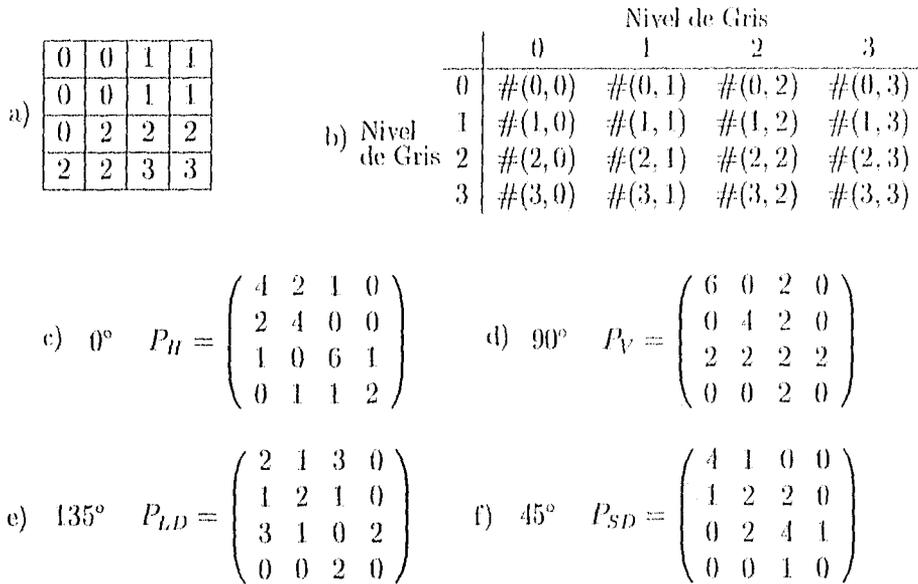


Figura 2.5: Ejemplo de Matriz de Coocurrencia.

La ventaja del cálculo de la coocurrencia de niveles de gris es que caracteriza las interrelaciones espaciales de los niveles de gris en un patrón de textura y además es invariante ante transformaciones monótonicas de los niveles de gris. En base a esta matriz pueden hacerse cálculos estadísticos que, según la aplicación, pueden dar información relevante sobre la imagen como son por ejemplo:

$$\text{Entropía} = - \sum_{i,j} P_{ij} \log P_{ij}$$

$$\begin{aligned}
\text{Contraste} & \sum_{i,j} |i-j|^k (P_{ij})^t \\
\text{Correlación} & \sum_{i,j} \frac{(i-\mu)(j-\mu)P_{ij}}{\sigma^2} \\
\text{donde } \mu & = \sum_{i,j} iP_{ij}
\end{aligned}$$

entre otras (para más detalle referirse a [17]).

2.2.2 Estimación de las σ_i . Implantación

Las σ_i de los valores de gris para cada región i no puede ser estimada, como lo hemos dicho anteriormente, directamente del histograma de la imagen debido a que éste representa una mezcla de funciones de densidad de probabilidad (FDP). Por lo tanto el algoritmo calcula una aproximación a la FDP de cada región i , utilizando una variante de la matriz de coocurrencia, contando la coocurrencia para cada pixel con valor gris promedio $\hat{\theta}_i$ y sus 8 vecinos, es decir, en este caso hacemos la suma algebraica de las cuatro matrices presentadas anteriormente en una sola matriz que representa las cuatro direcciones [6]. La variante radica en que la matriz de coocurrencia está definida como una matriz cuadrada de tamaño $L \times L$ donde L es el número de niveles de gris. En este caso nuestra matriz de coocurrencia es de tamaño $i \times L$ donde i es el número de clases y n es el número de niveles de gris.

Entonces, tomando las modas estimadas en el análisis espectral $\hat{\theta}_1$, $\hat{\theta}_2$, y $\hat{\theta}_3$ se hace el cálculo de la matriz de coocurrencia para los tonos de las modas:

	0	1	...	255
$\hat{\theta}_1$	$\#(\hat{\theta}_1, 0)$	$\#(\hat{\theta}_1, 1)$...	$\#(\hat{\theta}_1, 255)$
$\hat{\theta}_2$	$\#(\hat{\theta}_2, 0)$	$\#(\hat{\theta}_2, 1)$...	$\#(\hat{\theta}_2, 255)$
$\hat{\theta}_3$	$\#(\hat{\theta}_3, 0)$	$\#(\hat{\theta}_3, 1)$...	$\#(\hat{\theta}_3, 255)$

De esta manera se tiene una matriz de 3×256 , o bien tres arreglos de 256 datos que contienen la información espacial y de textura de cada $\hat{\theta}_i$, cada uno de estos arreglos corresponde a una sola FDP de los tonos vecinos a la moda en cuestión. El siguiente paso es obtener nuevamente el $\hat{\theta}_{LMS}$ para $q = 0.50$ en cada una de estas FDP, así encontramos una nueva moda $\hat{\theta}'_i$ que es la estimada como la que caracteriza a la clase i . Falta además calcular la desviación estándar \hat{s}_i que deberá ser también robusta y se define como [43]:

$$\hat{s}_i = 1.4826 \sqrt{\text{med}_i r_{ij}^2}$$

donde $r_{ij} = y_{ij} - \hat{\theta}_i$, y 1.4826 es un factor de corrección para ajustar de 0.50 al 0.68 de los datos, que es el porcentaje correspondiente al cálculo de la desviación estándar para una distribución gaussiana.

Hay que hacer notar que los dos parámetros encontrados en esta sección corresponden a la $\hat{\theta}_i$ y \hat{s}_i de las *distribuciones de los vecinos de las $\hat{\theta}_i$* de cada clase, no así la distribución de cada clase, hablando estrictamente.

Cuando tenemos distribuciones de clase que no se traslapan, este método permite encontrar con una gran exactitud los parámetros de las clases originales, como se demostrará en el siguiente capítulo donde mostraremos los resultados obtenidos. Pero en la vida real la mayoría de las clases están espectralmente mezcladas por lo que consideramos que este método nos da una buena aproximación, por eso la insistencia de que obtenemos las *distribuciones de los vecinos de las $\hat{\theta}_i$* .

2.3 Crecimiento de Regiones

Como su nombre lo indica y como lo mencionamos con anterioridad, el crecimiento de regiones es un procedimiento que agrupa píxeles o subregiones en regiones mayores. La manera más simple de este procedimiento es la *agregación de píxeles*, que comienza con un conjunto de puntos llamados “semillas” y en base a éstas crecen las regiones agregando a cada semilla aquellos puntos vecinos que tengan propiedades similares como por ejemplo el nivel de gris, la textura o el color. Siguiendo esta lógica nos encontramos con dos problemas básicos: 1) encontrar los puntos iniciales semillas apropiados, y 2) encontrar la regla de propiedades o predicados lógicos para llevar a cabo la agregación; así como establecer el criterio de paro que generalmente es cuando ya no existen píxeles que satisfagan las propiedades pedidas. Existen una gran variedad de criterios para el crecimiento de regiones en la literatura [16] [18]; nosotros proponemos en este trabajo un algoritmo novedoso de crecimiento de regiones que trabaja iterativamente de manera automática adaptando los umbrales según las condiciones locales de la imagen, y utiliza técnicas de relajación en el criterio de agregación. Describiremos en la siguiente subsección el algoritmo.

2.3.1 Implantación

Partimos en este momento de los parámetros $\hat{\theta}_i$ y \hat{s}_i encontrados en las dos secciones anteriores que consideramos como “semillas”. El algoritmo de crecimiento de regiones implantado fue pensado para resolver los casos cuando un valor de gris puede pertenecer a dos distribuciones diferentes a la vez, como se muestra en la Figura 2.6.

Los intervalos I_1 y I_2 , cuyo tamaño depende de la desviación estándar de cada distribución, se incrementan iterativamente. Para cada iteración, un valor de gris p que cae dentro de un intervalo es clasificado como perteneciente a dicha distribución, dependiendo del número de vecinos que ya han sido previamente clasificados como pertenecientes a dicha región. Enseguida se describe el algoritmo en detalle.

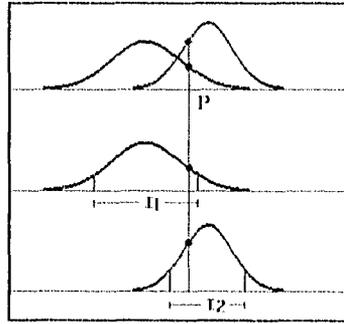


Figura 2.6: Confusión en la pertenencia de clase.

Inicialmente ningún pixel está clasificado, entonces:

- Paso 1. El algoritmo inicia plantando **semillas** para cada región i . Un pixel p es semilla de la región i si su valor de gris pertenece al intervalo $I_i = (\theta_i - a\hat{s}_i, \theta_i + a\hat{s}_i)$ para $a = 1.0$ y además por lo menos el 50% de sus vecinos también pertenecen al mismo intervalo. El resultado de este procedimiento nos asegura que serán clasificados como semillas de la región i sólo aquellos pixeles que caen dentro de la región i con una muy alta probabilidad, ya que están rodeados de pixeles que también pertenecen a la región i .

Una vez que las semillas de todas las regiones han sido plantadas:

- Paso 2. Las regiones crecen de la siguiente manera: un pixel con valor de gris p dentro del intervalo I_i es clasificado como perteneciente a la clase i si tiene por lo menos un vecino clasificado como de clase i . Un valor de gris que pertenece a dos distribuciones distintas a la vez se clasifica en la clase que tenga la mayor ocurrencia de vecinos ya clasificados; en caso de empate se clasifica en la clase cuya distancia de Mahalanobis respecto a la moda sea menor, es decir se clasifica como perteneciente a la clase más cercana. Este procedimiento se repite hasta que ya no existen pixeles que clasificar o agregar a ninguna clase.
- Paso 3. Las restricciones se relajan incrementando $a \leftarrow a + 0.5$ y el Paso 2 se repite para $a = 1.5, 2.0, \dots$ hasta que todos los pixeles de la imagen están clasificados en alguna clase.

El resultado de este algoritmo es un crecimiento de regiones por umbral adaptativo que depende de la información local, tanto espectral como espacial de la imagen. Esta idea se ilustra en la Figura 2.7. La posición del umbral T entre dos modas μ_1 y μ_2 depende del análisis local dentro de un intervalo I , hecho sobre un pixel respecto a sus vecinos clasificados; de manera que el umbral se adapta según sea el caso.

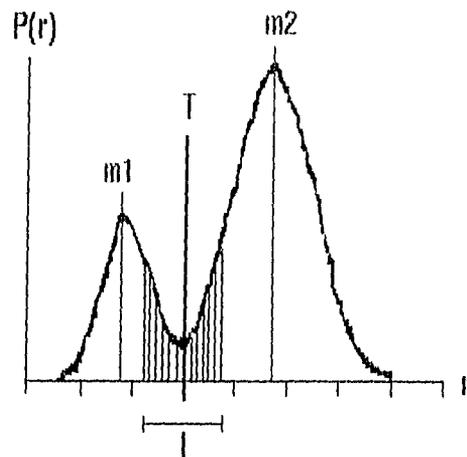


Figura 2.7: Umbral adaptativo.

Capítulo 3

Resultados

El método de segmentación desarrollado fue validado con imágenes sintéticas en donde conocemos con exactitud la estadística de cada clase; también fue probado como mencionamos al principio con las imágenes de histología de cerebro de rata y una imagen médica de resonancia magnética nuclear.

3.1 Validación

La primera validación que hicimos fue probar que efectivamente con el análisis espectral y espacial propuesto obteníamos los parámetros que caracterizan a las clases.

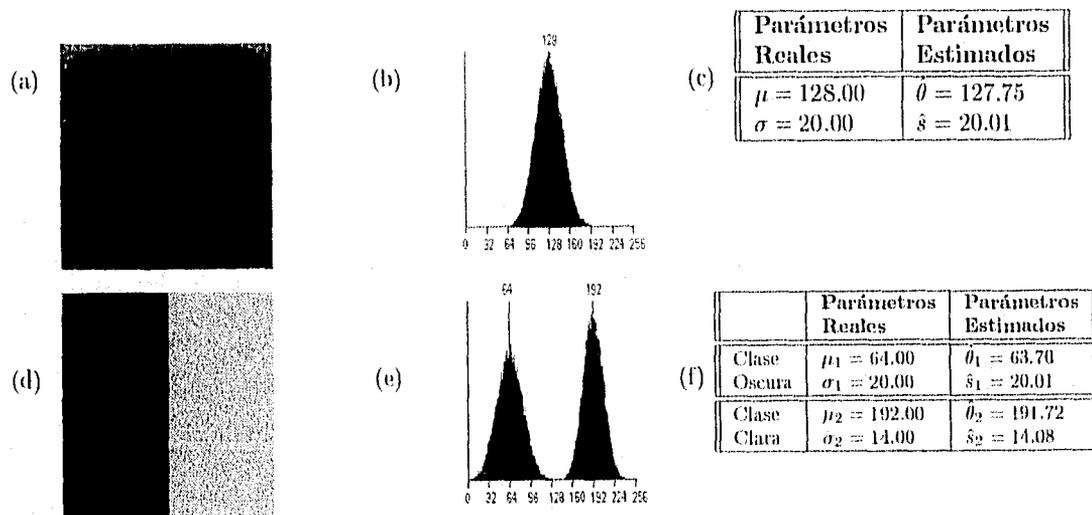


Figura 3.1: (a) Imagen Original; (b) Histograma de (a); (c) Tabla de Estadística; (d) Imagen Original; (e) Histograma de (d); (f) Tabla de Estadística.

En la Figura 3.1a y b tenemos la representación de una imagen sintética con una sola distribución y su respectivo histograma en donde se señala la posición de la moda. En la Figura 3.1c tenemos la tabla de estadística: en la primera columna se encuentran los *parámetros reales*, es decir los parámetros con los que se generó la distribución gaussiana que se muestra; en la segunda columna de la misma tabla se encuentran los *parámetros estimados*, es decir los calculados por el método en donde podemos ver que se recuperan los parámetros originales con gran exactitud. Igualmente se muestra en la Figura 3.1d, e y f, pero en este caso son dos distribuciones gaussianas que no se traslapan, de la misma manera vemos que los parámetros se recuperan perfectamente.

Se generaron otras dos imágenes sintéticas simulando objetos cuyas distribuciones espectrales están traslapadas de manera notable.

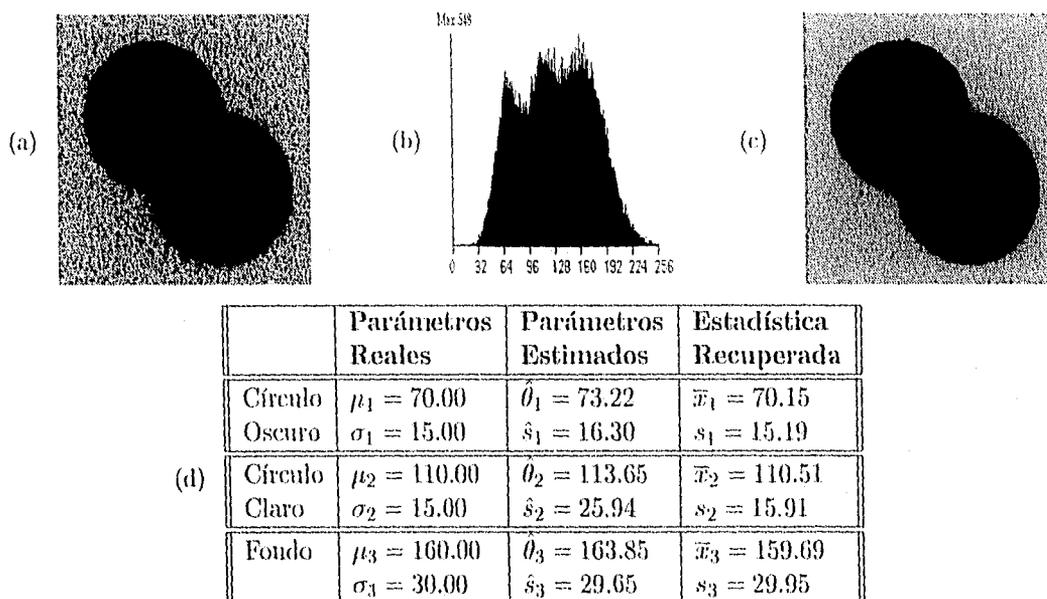


Figura 3.2: (a) Imagen Original; (b) Histograma de (a); (c) Imagen Clasificada; (d) Tabla de Estadísticas.

La Figura 3.2a representa una imagen sintética de 256×256 píxeles y 2^8 niveles de gris, que contiene tres diferentes regiones: el círculo oscuro, el círculo claro y el fondo. En la Figura 3.2b está representado su histograma y en la tabla 3.2d la estadística correspondiente. En la primera columna de la tabla 3.2d se encuentran los *parámetros reales* con los que se generó cada clase, el histograma es una mezcla de estas tres distribuciones gaussianas. En la segunda columna de la misma tabla se encuentran los *parámetros*

estimados por el método. En estos parámetros se basa el método para hacer el crecimiento de regiones que se describió en el capítulo anterior.

Aquí hay que hacer notar que la desviación estándar de cada clase está sobrestimada respecto a las desviaciones originales, principalmente en la clase central (círculo claro, de 15.0 a 25.94) a diferencia de los ejemplos mostrados en la Figura 3.1 en donde vemos una recuperación muy exacta. Esto es debido a que las curvas gaussianas de cada población están traslapadas, entonces ocurre que existen por ejemplo, píxeles con nivel de gris 114 en la clase del círculo oscuro que corresponde a su vez a la moda del círculo claro, por lo que al hacer un conteo por coocurrencia de los vecinos del tono 114 entran en esta distribución puntos que no están contenidos en la clase círculo claro y esto hace que la σ de esta población crezca. Es por esto que a estos datos, cuando tenemos distribuciones traslapadas, los llamamos *distribución de vecinos de la μ_i* , ya que de manera estricta no representan a la distribución de la clase i .

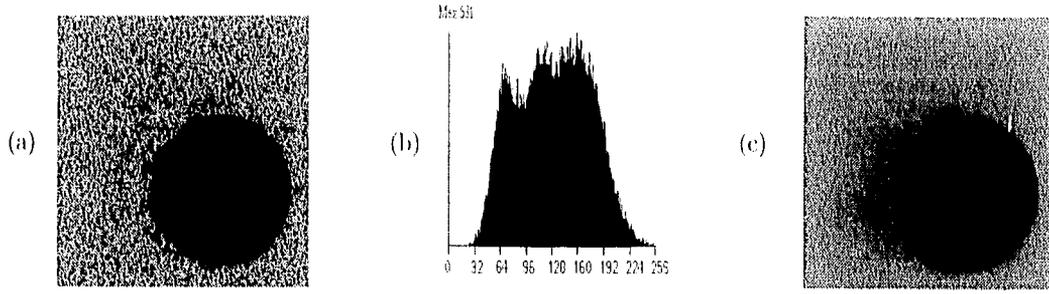
Sin embargo esta estadística es una buena aproximación a la estadística de la clase, el método de crecimiento de regiones se encarga de hacer un análisis espacial al momento de agregar píxeles, por lo que esta sobrestimación no afecta en la clasificación como puede verse en la Figura 3.2c que representa la imagen resultado de la segmentación en la que vemos que los contornos de los objetos están muy bien definidos. Por último en la tabla 3.2d tercera columna se muestran las estadísticas recuperadas de los valores de gris originales de aquellos píxeles clasificados como pertenecientes a alguna región, con los datos de esta columna evaluamos también la eficiencia del método de crecimiento de regiones ya que recuperamos la estadística original.

En la Figura 3.3a tenemos el segundo ejemplo de imagen sintética de 256×256 píxeles con 2^8 niveles de gris que fue generada con la misma estadística que el ejemplo de la Figura 3.2, como se muestra en la primera columna de la tabla 3.3d.

En esta imagen se ven claramente dos clases el círculo y el fondo, sin embargo el histograma de la Figura 3.3b muestra que estadísticamente es el mismo que en el ejemplo de la Figura 3.2b. Esto se debe que la clase central (círculo claro) fue distribuida espacialmente en forma aleatoria en el fondo teniendo el cuidado de no alterar la n de cada clase, con lo que tenemos un círculo y dos fondos. En la primera columna de la tabla 3.3d se muestra además el promedio ponderado de la media y desviación estándar de los dos fondos definida como [19]:

$$\mu_{2'} = \frac{n_2\mu_2 + n_3\mu_3}{n_2 + n_3}$$

$$\sigma_{2'} = \sqrt{\frac{n_2\sigma_2^2 + n_3\sigma_3^2}{n_2 + n_3}}$$



	Parámetros Reales		Parámetros Estimados	Estadística Recuperada
	Círculo		$\mu_1 = 70.00$ $\sigma_1 = 15.00$	$\hat{\theta}_1 = 72.02$ $\hat{s}_1 = 21.49$ $\bar{x}_1 = 70.21$ $s_1 = 15.40$
(d)	Fondo 1	$\mu_2 = 110.00$ $\sigma_2 = 15.00$	$\mu_{2'} = 147.45$ $\sigma_{2'} = 27.03$	$\hat{\theta}_{2'} = 145.20$ $\hat{s}_{2'} = 39.28$ $\bar{x}_{2'} = 147.15$ $s_{2'} = 34.66$
	Fondo 2	$\mu_3 = 160.00$ $\sigma_3 = 30.00$		

Figura 3.3: (a) Imagen Original; (b) Histograma de (a); (c) Imagen Clasificada; (d) Tabla de Estadística.

Partiendo de la estadística inicial, en principio obtenemos las mismas modas que en el ejemplo de la Figura 3.2 durante el análisis espectral ya que partimos del mismo histograma, pero al hacer el análisis espacial con la matriz de coocurrencia se deduce que los dos fondos son una misma clase. Estrictamente hablando los resultados en el análisis espacial dan para este ejemplo los siguientes datos:

$$\text{Fondo 1 } \hat{\theta}_2 = 145.20 \quad \hat{s}_2 = 39.28$$

$$\text{Fondo 2 } \hat{\theta}_3 = 141.56 \quad \hat{s}_3 = 39.28$$

El criterio para considerar estas dos clases como la misma población está basado en dos condiciones, la primera:

$$|\hat{\theta}_2 - \hat{\theta}_3| < 5.0,$$

sólo si se cumple esta condición entonces la segunda condición es llevar a cabo la prueba de *Kolmogorov-Smirnov* (K-S) [36] que consiste en calcular y comparar las curvas de las funciones de *distribución acumulada* de ambas distribuciones. Lo que la prueba K-S calcula es un parámetro D que corresponde al *valor máximo* de la diferencia absoluta entre las dos funciones de distribución acumulada.

Para comparar dos diferentes funciones de distribución acumulada $S_{N_1}(x)$ y $S_{N_2}(x)$, la expresión K-S es la siguiente:

$$D = \max_{-\infty < x < \infty} |S_{N_1}(x) - S_{N_2}(x)|$$

Lo que hace útil a la prueba K-S es que se puede calcular su distribución para el caso de la hipótesis nula (que los conjuntos de datos pertenecen a la misma distribución), con una muy buena aproximación, así nos da la significancia para cualquier valor de D diferente de cero. La función que realiza el cálculo de la significancia puede escribirse como la suma:

$$Q_{KS}(\lambda) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2 \lambda^2}$$

cuyos valores límites son:

$$Q_{KS}(0) = 1 \quad Q_{KS}(\infty) = 0.$$

En términos de esta función, el nivel de significancia para un valor observado de D (como una refutación de la hipótesis nula sobre distribuciones iguales) está dado aproximadamente por la fórmula:

$$\text{Probabilidad}(D > \text{observada}) = Q_{KS} \left(\left[\sqrt{N_e} + 0.12 + 0.11/\sqrt{N_e} \right] D \right)$$

donde N_e es el número efectivo de puntos,

$$N_e = \frac{N_1 N_2}{N_1 + N_2},$$

para el caso de dos distribuciones donde N_1 es el número de puntos en la primera distribución, y N_2 es el número de puntos en la segunda distribución (en este caso $N_1 = N_2 = 256$ puntos). Mientras más pequeño sea el valor de D la probabilidad de que los dos conjuntos de datos pertenezcan a la misma distribución es mayor; por lo que en la segunda condición establecimos que:

$$D < 0.1 \quad .$$

En el ejemplo anterior $D=0.0585$ y la Probabilidad= 0.7595 , por lo que las curvas Fondo 1 y Fondo 2 pertenecen a la misma distribución. De esta manera de tres clases se reduce a dos.

De lo anterior se desprende la importancia del análisis espacial que hicimos aplicando la variante de matriz de coocurrencia. No es suficiente partir solamente de la información disponible en el histograma para obtener buenos resultados, sobre todo cuando tenemos histogramas cuyas clases están muy traslapadas, o bien cuando tenemos casos como el ejemplo anterior de imagen sintética en los que el histograma no nos da ninguna información espacial sobre los dos fondos; sin el análisis espacial no es posible tomar la decisión de que los dos fondos son uno sólo.

En la segunda columna de la tabla 3.3d se muestran los parámetros estimados por el método, el resultado de la segmentación está representado en Figura 3.3c. Finalmente en

la tercera columna de la tabla 3.3d se muestran las estadísticas recuperadas de los valores de gris originales de aquellos píxeles clasificados como pertenecientes a alguna región. Nuevamente vemos que las desviaciones estimadas están por arriba de las originales, sin embargo al aplicar el algoritmo de crecimiento de regiones éstas finalmente se recuperan.

Para comparar nuestros resultados con un método clásico hemos elegido el método de OTSU [33], que ya describimos anteriormente. La razón por la cual elegimos este método, y no otro, fue básicamente porque requerimos hacer la comparación para tres y cuatro clases, por lo que se necesitó hacer la generalización, y consideramos que el método de OTSU es el menos complicado de generalizar.

En la Figura 3.4 se muestra la segmentación de estas dos últimas imágenes utilizando el **método de OTSU** generalizado para dos umbrales. Las Figuras 3.4a y d son las imágenes originales, las 3.4b y e corresponden respectivamente a sus histogramas donde se señala la posición de los umbrales calculados por OTSU, y en 3.4c y f se tiene el resultado de la segmentación; en este caso se etiquetaron las clases como negro el círculo oscuro, gris el círculo claro y blanco el fondo.

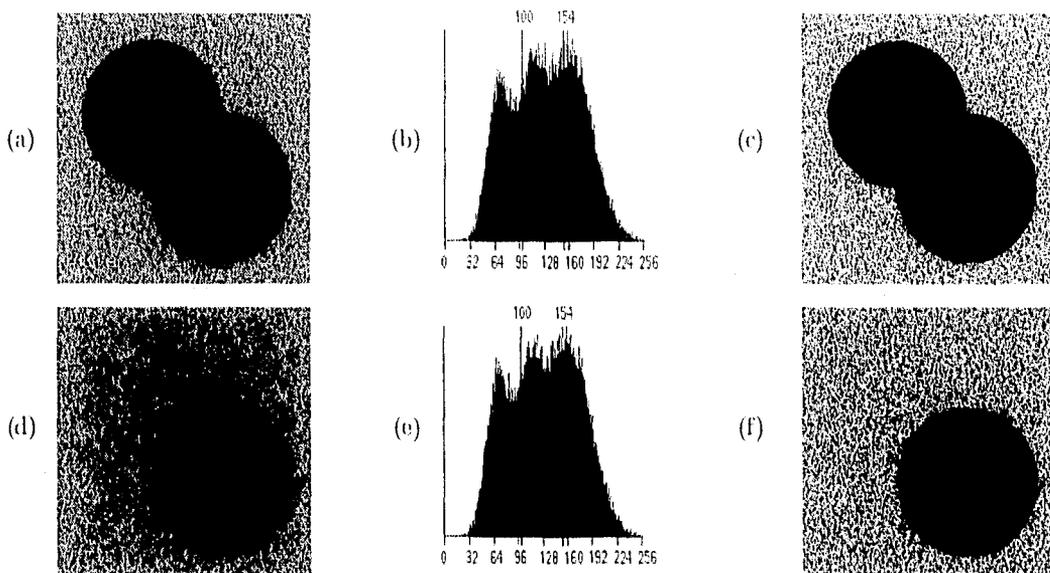


Figura 3.4: (a) Imagen Original Ejemplo 1; (b) Histograma de (a); (c) Imagen Clasificada de (a); (d) Imagen Original Ejemplo 2; (e) Histograma de (d); (f) Imagen Clasificada de (d).

Puede notarse en las Figuras 3.4c y f, que ninguna clase está bien definida, sólo el caso del círculo oscuro es el que tiene menos error. En este método no se hace ningún tipo de análisis espacial, es exclusivamente espectral, por lo que el resultado es muy pobre.

3.2 Aplicación

Como mencionamos en la introducción de este trabajo, uno de los problemas a resolver fue poder segmentar cortes histológicos de cerebro de rata. Hay básicamente dos tipos de imágenes de cerebro de rata, los cortes control y los cortes lesionados. En los primeros tenemos tres regiones: el fondo, la corteza y el caudado; en los segundos tenemos cuatro regiones: el fondo, la corteza, la lesión y el caudado. El objetivo es segmentar la corteza en los cortes control, y la corteza con la respectiva lesión en los cortes lesionados.

En la Figura 3.5a se ilustra una imagen de 256×256 píxeles y 2^8 niveles de gris de un corte histológico control de cerebro de rata, en la Figura 3.5b su histograma y en Figura 3.5c la imagen clasificada con nuestro método, donde aparecen tres regiones: la clase marcada con negro se llama el caudado, la clase marcada con gris se llama corteza cerebral y la clase blanca es el fondo.

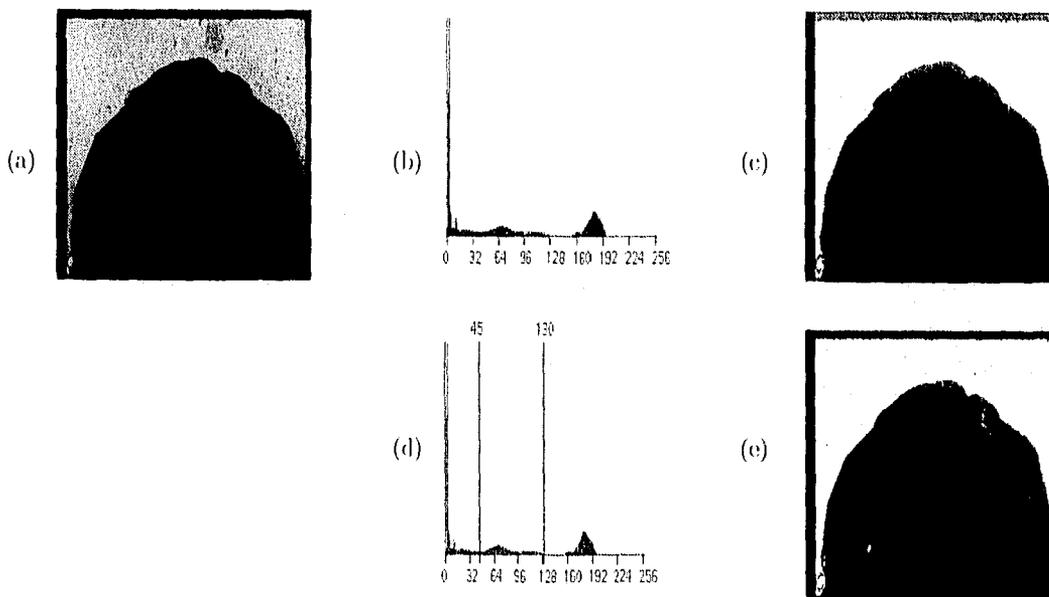


Figura 3.5: (a) Corte histológico control de cerebro de rata; (b) Histograma de (a); (c) Imagen Clasificada de (a) por el método propuesto; (d) Histograma de (a) con umbrales por OTSU; (e) Imagen Clasificada por OTSU.

Nótese que la corteza es una zona muy oscura en la imagen original, sin embargo con este método la frontera aparece delineada con mucha exactitud. También hay que notar que en el histograma de esta imagen a simple vista pueden encontrarse tres distribuciones, lo que quiere decir que el histograma contiene tres modas significativas que se encuentran

sin problemas durante el análisis espectral, con esto y con el método de crecimiento implantado se logra un buen resultado sobre todo en la parte oscura de la corteza.

En la Figura 3.5d se muestra el mismo histograma donde se señalan los umbrales calculados por OTSU, y en la Figura 3.5e el resultado de la segmentación por OTSU donde se pierde en gran parte del área de interés.

En la Figura 3.6a se ilustra una imagen de 256×256 píxeles y 2^8 niveles de gris de un corte histológico de cerebro de rata con lesión en la corteza cerebral, en la Figura 3.6b su histograma y en Figura 3.6c la imagen clasificada con nuestro método donde aparecen cuatro regiones: la negra es el caudado, la gris oscura es la corteza cerebral, la gris clara es la lesión y la blanca el fondo. En esta última imagen se observa que la segmentación no es del todo correcta, esto se debe a varias razones: 1) que la imagen original está tomada con mala luz y con el lente sucio por lo que tiene mucha basura, 2) no tiene suficiente contraste por lo que la lesión tiene tonos de gris muy parecidos a una parte de la región del fondo, ya que el fondo no es homogéneo, esto tiene como resultado que en la segmentación se detecta parte del fondo como parte de la lesión.

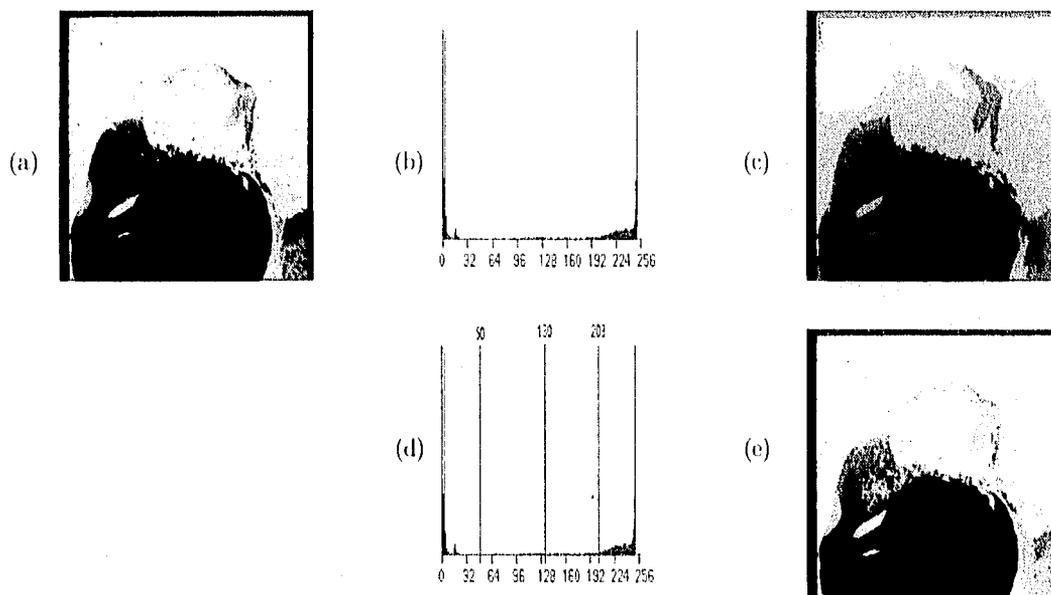


Figura 3.6: (a) Corte histológico de cerebro de rata con lesión en la corteza cerebral; (b) Histograma de (a); (c) Imagen Clasificada de (a) por el método propuesto; (d) Histograma de (a) con umbrales por OTSU; (e) Imagen Clasificada por OTSU.

En la Figura 3.6d se muestra el mismo histograma con los umbrales calculados por OTSU, en este caso tres umbrales porque son cuatro regiones, y en la Figura 3.6e el resultado de la segmentación por OTSU, donde vemos que la clase fondo y la clase lesión

están detectadas como una misma región, y en la que debería ser la clase corteza están mezcladas las tres clases restantes.

Finalmente probamos, como otro ejemplo de aplicación, una imagen de 256×256 píxeles y 2^8 niveles de gris, esta vez se trata de una imagen de resonancia magnética nuclear de cerebro humano que se muestra en la Figura 3.7a. En la Figura 3.7b está el resultado de la segmentación para tres clases, y en la Figura 3.7c tenemos la imagen clasificada resultado de segmentar nuevamente una máscara de la clase blanca de la Figura 3.7b. Debe notarse que el algoritmo es capaz de recobrar el detalle del tejido blando del cerebro.

Con este último ejemplo lo que queremos mostrar es que si mantenemos fijo el número de clases a buscar, en este caso tres clases, el método puede aplicarse iterativamente sobre una clase hasta conseguir el detalle deseado.

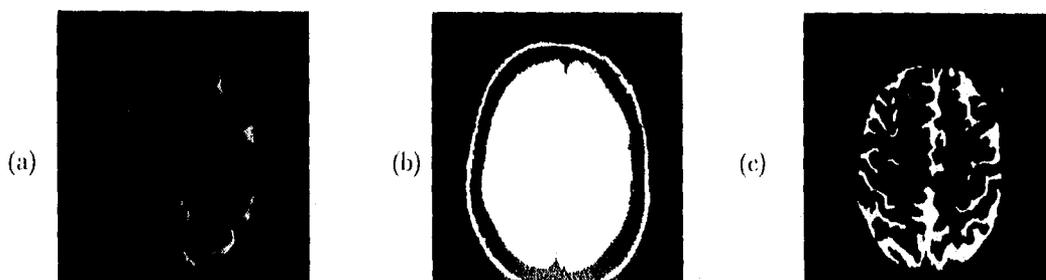


Figura 3.7: (a) Resonancia Magnética Nuclear de cerebro humano; (b) Imagen Clasificada de (a) con tres clases; (c) Imagen Clasificada de una máscara de la clase blanca de (b) con tres clases.

Capítulo 4

Conclusiones

Hemos desarrollado un algoritmo que estima, basado en información global y local, los parámetros de las regiones presentes en una imagen, y adapta automáticamente los umbrales de las distribuciones utilizando información local.

El algoritmo no detecta de manera automática el número significativo de clases contenidas en la imagen, es necesario dar de antemano el número de clases a segmentar, por lo que para esto se requiere de un conocimiento *a priori* sobre la imagen en cuestión.

Este algoritmo trabaja con mucha exactitud cuando tenemos histogramas que contienen modas significativas. Como pudimos ver en los ejemplos de validación y de aplicación. Estamos de acuerdo que si esto no ocurre, debe utilizarse otro método de detección de modas local.

Cabe mencionar también que el hecho de que el método sea iterativo, como se muestra en el ejemplo de la Figura 3.7, sugiere que es posible generalizarlo de manera que podamos encontrar desde el principio del análisis espectral, y refinándolo con el análisis espacial, un número variable de clases. La manera de hacerlo totalmente automático sería:

- Paso 1. Partir en principio de un número fijo de clases, por ejemplo dos.
- Paso 2. Probar, utilizando la prueba de Kolmogorov-Smirnov, si las clases encontradas son diferentes.
- Paso 3. Para cada clase diferente, volver al Paso 1. Parar cuando ya no existan clases diferentes.

No pudimos encontrar la segmentación correcta para las imágenes lesionadas de los cortes cerebrales de rata debido a las malas condiciones de iluminación de las mismas. Sin embargo, pudimos demostrar en base a los resultados obtenidos, que el método espacial y el método de crecimiento de regiones son herramientas poderosas cuando contamos con las modas correctas.

Capítulo 4

Conclusiones

Hemos desarrollado un algoritmo que estima, basado en información global y local, los parámetros de las regiones presentes en una imagen, y adapta automáticamente los umbrales de las distribuciones utilizando información local.

El algoritmo no detecta de manera automática el número significativo de clases contenidas en la imagen, es necesario dar de antemano el número de clases a segmentar, por lo que para esto se requiere de un conocimiento *a priori* sobre la imagen en cuestión.

Este algoritmo trabaja con mucha exactitud cuando tenemos histogramas que contienen modas significativas. Como pudimos ver en los ejemplos de validación y de aplicación. Estamos de acuerdo que si esto no ocurre, debe utilizarse otro método de detección de modas local.

Cabe mencionar también que el hecho de que el método sea iterativo, como se muestra en el ejemplo de la Figura 3.7, sugiere que es posible generalizarlo de manera que podamos encontrar desde el principio del análisis espectral, y refinándolo con el análisis espacial, un número variable de clases. La manera de hacerlo totalmente automático sería:

- Paso 1. Partir en principio de un número fijo de clases, por ejemplo dos.
- Paso 2. Probar, utilizando la prueba de Kolmogorov-Smirnov, si las clases encontradas son diferentes.
- Paso 3. Para cada clase diferente, volver al Paso 1. Parar cuando ya no existan clases diferentes.

No pudimos encontrar la segmentación correcta para las imágenes lesionadas de los cortes cerebrales de rata debido a las malas condiciones de iluminación de las mismas. Sin embargo, pudimos demostrar en base a los resultados obtenidos, que el método espacial y el método de crecimiento de regiones son herramientas poderosas cuando contamos con las modas correctas.

Capítulo 4

Conclusiones

Hemos desarrollado un algoritmo que estima, basado en información global y local, los parámetros de las regiones presentes en una imagen, y adapta automáticamente los umbrales de las distribuciones utilizando información local.

El algoritmo no detecta de manera automática el número significativo de clases contenidas en la imagen, es necesario dar de antemano el número de clases a segmentar, por lo que para esto se requiere de un conocimiento *a priori* sobre la imagen en cuestión.

Este algoritmo trabaja con mucha exactitud cuando tenemos histogramas que contienen modas significativas. Como pudimos ver en los ejemplos de validación y de aplicación. Estamos de acuerdo que si esto no ocurre, debe utilizarse otro método de detección de modas local.

Cabe mencionar también que el hecho de que el método sea iterativo, como se muestra en el ejemplo de la Figura 3.7, sugiere que es posible generalizarlo de manera que podamos encontrar desde el principio del análisis espectral, y refinándolo con el análisis espacial, un número variable de clases. La manera de hacerlo totalmente automático sería:

- Paso 1. Partir en principio de un número fijo de clases, por ejemplo dos.
- Paso 2. Probar, utilizando la prueba de Kolmogorov-Smirnov, si las clases encontradas son diferentes.
- Paso 3. Para cada clase diferente, volver al Paso 1. Parar cuando ya no existan clases diferentes.

No pudimos encontrar la segmentación correcta para las imágenes lesionadas de los cortes cerebrales de rata debido a las malas condiciones de iluminación de las mismas. Sin embargo, pudimos demostrar en base a los resultados obtenidos, que el método espacial y el método de crecimiento de regiones son herramientas poderosas cuando contamos con las modas correctas.

Hemos enviado el artículo de este trabajo para ser evaluado en la Conferencia Internacional de Procesamiento de Imágenes (ICIP'96) que organiza la IEEE, y que se llevará a cabo en septiembre de 1996. Anexamos como apéndice a este trabajo una copia de dicho artículo.

4.1 Tendencias al futuro

Las tendencias al futuro serán, por lo tanto, refinar el método de detección de modas cuando éstas no son significativas. Una manera más eficiente tal vez sería no basarnos en la información espectral del histograma sino hacer desde el principio un análisis espacial local para su detección, de manera que podamos además encontrar automáticamente el número de clases en cada imagen [15]; es decir, que el método detecte números variables de clases según la imagen en cuestión, sin necesidad de tener información *a priori* de la imagen.

Una vez resuelto este problema, el siguiente paso será dedicarnos al área de *análisis* y realizar la medición de las densidades ópticas de los cortes control y los cortes experimentales lesionados de rata. La medición de la densidad óptica en los cortes de rata se hace sobre la tinción del tejido, el problema del investigador es poder hacer mediciones cuantitativas sobre esta tinción. La tinción histológica de estos cortes en particular, revela la actividad enzimática de Acetil- colinesterasa (AChE). Esta actividad constituye un marcador indirecto de la liberación de acetilcolina en las diversas áreas cerebrales. Este neurotransmisor se encuentra estrechamente vinculado con las funciones de aprendizaje y memoria mediadas por la corteza cerebral. La tinción es, pues, función de una actividad fisiológicamente involucrada en los mecanismos asociativos del cerebro. Las manipulaciones experimentales que dan lugar a correspondientes alteraciones colinérgicas permiten extraer información valiosa sobre la participación de esta circuitería en las funciones del cerebro. La cuantificación de las alteraciones colinérgicas experimentales, medidas mediante el nivel histológico de tinción de AChE como la densidad óptica, el área de la lesión y la ubicación, constituirían para el investigador una herramienta de gran utilidad.

Bibliografía

- [1] Ahuja, N., Rosenfeld, A., "A note on the use of second-order gray-level statistics for threshold selection.", *Proc. IEEE Trans. Systems Man Cybernet.*, vol. SMC-8, pp. 895- 899, 1978.
- [2] Bhanu, B., Feucher, O., "Segmentation of images having unimodal distributions.", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, pp. 408-419, 1982.
- [3] Boukharouba, S., Rebordao, J.M., Wendel, P.L., "An amplitude segmentation method based on the distribution function of an image.", *Computer Vision, Graphics and Image Processing.*, vol. 29, pp. 47- 59, 1985.
- [4] Brice, C., Fennema, C., "Scene analysis using regions.", *Artificial Intelligence.*, vol. 1, pp. 205- 226, 1975.
- [5] Carlton, C.G., Mitchell, O.R., "Image Segmentation Using Texture and Gray Level", *Proc. IEEE Conf. Pattern Recognition and Image Processing*, Troy, New York, no. 6-8, pp. 387- 391, 1977.
- [6] Chen, Patrick C. , Pavlidis, Theodosios, "Segmentation by Texture Using a Co-Occurrence Matrix and a Split-and-Merge Algorithm", *Computer Graphics and Image Processing*, vol. 10, pp. 172- 182, 1979.
- [7] Chien, Y.P., Fu, K.S., "Preprocessing and feature extraction of picture patterns.", TR-EE 74-20, Purdue University, West Lafayette, Indiana, 1974.
- [8] Chow, C.K., Kaneko, T., "Boundary detection of radiographic images by a threshold method.", *Proceedings, IFIP Congress 71*, pp. 130- 134, 1971.
- [9] Davis, L.S., "A Survey of Edge Detection Techniques", *Computer Graphics and Image Processing*, vol. 4, pp. 248- 270, 1975.
- [10] Deravi, F., Pal, S.K., "Gray level thresholding using second-order statistics.", *Pattern Recognit. Lett.*, vol. 1, pp. 417- 422, 1983.

- [11] Doyle, W., "Operation useful for similarity-invariant pattern recognition", *J. Assoc. Comput. Mach.* vol. 9, pp. 259-267, 1962.
- [12] Duda, R.O., Hart, P.E., "Use of the Hough transformation to detect lines and curves in pictures.", *Commun. ACM.*, vol. 15, pp. 11-15, 1972.
- [13] Feldman, J.A., Yakimovsky, Y., "Decision theory and artificial intelligence. I. A semantics-based region analyzer.", *Artificial Intelligence.*, vol. 5, pp. 349-371, 1974.
- [14] Fu, K.S., Mui, J.K., "A Survey on Image Segmentation", *Pattern Recognition.*, vol. 13, pp. 3-16, 1981.
- [15] Garza, M., Meer, P., Medina, V., "Robust Retrieval of 3D Structures from Magnetic Resonance Images.", *Submitted: 13th International Conference on Pattern Recognition, 1996.*
- [16] González, Rafael, Woods, Richard E. *Digital Image Processing*, New York, Addison-Wesley, 1992.
- [17] Haralick, R.M., Shanmugam, K., Dinstein, I., "Textural Features for Image Classification", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, 1973.
- [18] Haralick, R.M., Shapiro, L.G. *Computer and Robot Vision Volume 1*, New York, Addison-Wesley, 1992.
- [19] Hinkle, Dennis E., Wiersma, W., Jurs, Stephen G. *Applied Statistics for the Behavioral Sciences*, Boston, Houghton Mifflin Co., 1988.
- [20] Horowitz, S.L., Pavlidis, T., "Picture Segmentation by Directed Split-and-Merge Procedure", *Proc. Second Int. Joint Conf. Pattern Recognition*, pp. 424-433, 1974.
- [21] Hueckel, M., "An operator which locates edges in digital pictures.", *J. Ass. Comput. Mach.*, vol. 20, pp. 113-125, 1971.
- [22] Hueckel, M., "A local operator which recognizes edges and lines.", *J. Ass. Comput. Mach.*, vol. 20, pp. 634-647, 1973.
- [23] Kelly, M., "Edge detection by computer using planning.", *In Machine Intelligence.*, vol. VI, pp. 397-409, Edinburgh University Press, Edinburgh 1971.
- [24] Kirby, R.L., Rosenfeld, A., "A note on the use of (gray level, local average gray level) space as an aid in thresholding selection.", *IEEE Trans. Systems Man Cybernet.*, vol. SMC-9, pp. 860-864, 1979.

- [25] Kittler, J., Illingworth, J., "Minimum error thresholding.", *Pattern Recognition.*, vol. 19, pp. 41- 47, 1986.
- [26] Kohler, R., "A segmentation system based on thresholding.", *Computer Graphics and Image Processing.*, vol. 15, pp. 319- 338, 1981.
- [27] Lim, J.S., *Two-Dimensional Signal and Image Processing.*, Prentice-Hall, New Jersey, 1990.
- [28] Martelli, A., "Edge detection using heuristic search methods.", *Computer Graphics and Image Processing.*, vol. 1, pp. 169- 182, 1972.
- [29] Martelli, A., "An application of heuristic search methods to edge and contour detection.", *Communs ACM.*, vol. 14, pp. 73- 83, 1976.
- [30] Mason, D., Lauder, I.J., Rutoriz, D., Spowart, G., "Measurement of C-Bands in human chromosomes.", *Comput. Biol. Med.*, vol. 5, pp. 179- 201, 1975.
- [31] Meer, P., Jolion, J.M., Bataouche, S., "Robust Clustering with Applications in Computer Vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 13, no. 8, pp. 791- 802, 1991.
- [32] Muerle, J.L., Allen, D.C., "Experimental evaluation of techniques for automatic segmentation of objects in a complex scene.", *Pictorial Pattern Recognition (G.C. Cheng et al., Eds.)*, pp. 3-13, Thompson, Washington, 1968.
- [33] Otsu, Nobuyuki, "A Threshold Selection Method from Gray-Level Histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-9, no. 1, pp. 62- 66, 1979.
- [34] Panda, D.P., Rosenfeld, A., "Image Segmentation by Pixel Classification in (Gray Level, Edge Value) Space", *IEEE Transactions on Computers.* vol. C-27, pp. 875- 879, 1978.
- [35] Pavlidis, T., "Segmentation of pictures and maps through functional approximation.", *Computer Graphics and Image Processing.* vol. 1, pp. 360- 372, 1972.
- [36] Press, William H., Teukolsky, Saul A., Vetterling, William T., Flannery, Brian P. *Numerical Recipes in C. The Art of Scientific Computing*, New York, Cambridge University Press, 1988.
- [37] Prewitt, J.M., Mendelsohn, M.L., "The analysis of cell images.", *Ann. New York Acad. Sci.* vol. 128, pp. 1035- 1053, New York Acad. Sci., New York, 1966.
- [38] Pun, T., "A new method for gray-level picture thresholding using the entropy of the histogram.", *Signal Process.* vol. 2, pp. 223- 237, 1980.

- [39] Rosenfeld, A., Kak, A.C., *Digital Picture Processing*, Academic Press, New York, 1976.
- [40] Rosenfeld, A., "Interactive methods in image analysis.", *Proc. IEEE Conf. Pattern Recognition and Image Processing*, Troy, New York, no. 6-8, pp. 14- 20, 1977.
- [41] Rosenfeld, A., Smith, R.C., "Thresholding using relaxation.", *Proc. IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, pp. 598 - 606, 1981.
- [42] Rosenfeld, A., De la Torre, P., "Histogram concavity analysis as an aid in threshold selection.", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, pp. 598- 606, 1983.
- [43] Rousseeuw, Peter J., Leroy, Annick M. *Robust Regression & Outlier Detection*, New York, John Wiley & Sons, 1987.
- [44] Sahoo, P.K., Soltani, S., Wong, A.K., Chen, Y.C., "A Survey of Thresholding Techniques.", *Computer Vision, Graphics and Image Processing* vol. 41, pp. 233 - 260, 1988.
- [45] Southwell, R., "Relaxation Methods in Engineering Science, A Treatise on Approximate Computation.", Oxford University Press, London, 1940.
- [46] Tsai, W., "Moment-preserving thresholding: A new approach.", *Computer Vision, Graphics and Image Processing* vol. 29, pp. 377- 393, 1985.
- [47] Wang, S., Haralick, R.M., "Automatic multithreshold selection.", *Computer Vision, Graphics and Image Processing*, vol. 25, pp. 46- 67, 1984.
- [48] Wu, A.Y., Rosenfeld, A., "Threshold selection using quadtree.", *IEEE Trans. Pattern Anal. Mach. Intell.* vol. PAMI-4, pp. 90- 94, 1982.
- [49] Zucker, S.W., "SURVEY. Region Growing: Childhood and Adolescence.", *Computer Graphics and Image Processing*. vol. 5, pp. 382- 399, 1976.

Apéndice

Unsupervised Segmentation Based on Robust Estimation and Cooccurrence Data

M. E. Martínez-Pérez, M. Garza-Jimich

elena@greta.iimas.unam.mx

IIMAS, UNAM, MEXICO

Abstract

The accuracy of any image segmentation method depends on the correct estimation of the parameters of the different regions present in the image, as well as on the correct labelling of the pixels. Using robust estimators and relaxation labelling techniques, an unsupervised segmentation algorithm was developed. The mean gray value of each region is estimated from the histogram using robust clustering analysis. The gray level distribution of each individual region is approximated through the mean gray value cooccurrence data. The standard deviation of the gray levels of each region is estimated from this distribution using the Least Median of Squares (LMedS) robust estimator. The labelling of the pixels is done through an iterative relaxation region growing process, taking into account both spectral and spatial information. The method is tested in various images and validated with synthetic data, where it is shown that the known true parameters are recovered accurately.

1 Introduction. Assuming we have a set of sample data, in our case, pixels, which are the result of sampling and quantizing an image, drawn from a finite mixture of normal distributions, we wish to find descriptive measures that can predict each of the parameters μ and σ of the j different regions represented by such mixture. In classical statistics, the use of the sample mean \bar{x}_j as a measure of central tendency for the j th normal, and the sample standard deviation s_j as a measure of dispersion is very popular, due to their fast and explicit computation. Nevertheless, when the sample data contains even a small amount of outliers, these measures become very sensitive and may not estimate correctly the corresponding population parameters. This is precisely the case in image analysis, where conventional estimation techniques are unavoidably error prone. Thus, we propose the use of robust techniques combined with spatial analysis as an appropriate solution to overcome these problems.

2 Mean Gray Value Estimation. The mean gray value of the regions is estimated from the histogram using a variant of the minimum volume ellipsoid (MVE) robust estimator [3]. Let $h(i)$, $i = 1, \dots, n$ represent the histogram of an image with n gray levels, let $H = \sum h(i)$ represent the mass of the image, let $h_c(k)$, $k = k_{c1}, \dots, k_{c2}$ represent a cluster of h from gray value k_{c1} to gray value k_{c2} . Modes in h , assumed to correspond to the most frequent values of the different regions, are iteratively identified by locating the most compact cluster h_c [2] in h which contains qH data. The compactness of a cluster h_c is defined as $\sum h_c(k)/(k_{c2} - k_{c1})$, $k = k_{c1}, \dots, k_{c2}$. A cluster candidate to be removed is one which for two consecutive values of q , $q = 0.5, 0.45, \dots, 0.15$, the center of the interval containing the cluster is the same, or if q reached value 0.15. The cluster found is extracted, the center of the interval $[k_{c1}, k_{c2}]$ is assumed to be the estimator $\hat{\mu}_j$ of the mean gray value of region j , H is recomputed. The algorithm is repeated until the number of detected clusters exceeds an upper bound.

3 Standard Deviation Estimation. The standard deviation of the gray values of each region j cannot be estimated from the histogram, due to the fact that the histogram represents a mixture of PDF's. Therefore, the algorithm recovers an approximation of the PDF of each region j by counting the cooccurrence of the pixels with gray level $\hat{\mu}_j$ and their 8 neighbors [1]. The histogram thus obtained contains one distribution. Its standard deviation $\hat{\sigma}_j$ is estimated robustly [3] as:

$$\hat{\sigma}_j = 1.4826 \sqrt{\text{med } r_{ij}^2}, \quad r_{ij} = x_{ij} - \hat{\mu}_j.$$

4 Region Growing Algorithm The labelling algorithm was designed to solve the case when one gray level may belong to two different distributions, as shown in Fig. 1. Intervals I_1 and I_2 (see figure), which depend on the standard deviation of each distribution, are iteratively increased in size. For each iteration, a gray value p lying in such an interval is labelled as belonging to the region represented by such distribution, depending on the number of neighbors that have already been labelled as members of that region. Below we describe the algorithm in detail.

Step 1. The algorithm starts by planting seeds for every region j . A pixel is a seed of region j if its gray value is in the interval $I_j = (\hat{\mu}_j - a\hat{\sigma}_j, \hat{\mu}_j + a\hat{\sigma}_j)$, $a = 1.0$ and if more than 50% of its neighbors lie in the same interval. The result of this procedure is that only those pixels in the inner areas of each region, with very high probability of belonging to the region, are labelled as seeds.

Once the seeds of all regions are planted:

Step 2. Regions are grown in the following manner: a pixel with gray value in interval I_j is labelled as belonging to class j if it has at least one neighbor already labelled as class j . A gray value which may belong to two different distributions is assigned to the one to which most of its neighbors belong, or in case of a tie, the pixel is assigned to the closest class (Mahalanobis distance). The growing procedure is repeated until no more pixels are aggregated.

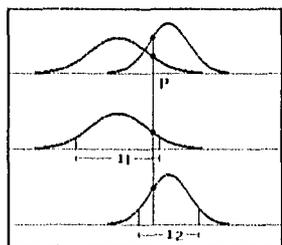


Figure 1: Labelling confusion.

The constraints are relaxed by incrementing a by 0.5 and step 2 is repeated for $a = 1.5, 2.0, \dots$ until all pixels have been labelled. The result is an adaptive thresholding, which depends both on spectral and spatial local information.

5 Experimental Results. Our method was validated with synthetic images and tested on images which are difficult to segment with classical methods. In these examples, the number of cluster to be found was 3. Fig. 2a represents a synthetic image containing three different regions produced with parameters shown in column 1 of Table 2d. Fig. 2b is the histogram from where the mean values were estimated. Fig. 2c is the resulting labelled image. Table 2d is the result of validation. Column 2 shows the accuracy of the estimated parameters of every region. Column 3 shows the statistics of the original gray values of those pixels labelled as belonging to each region.

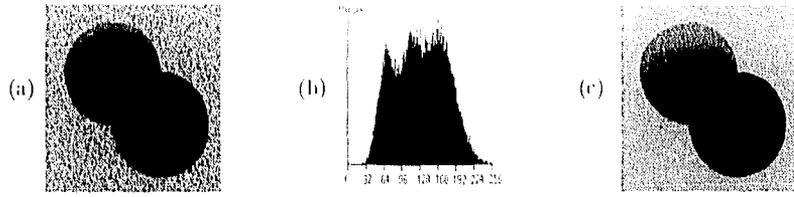
Fig. 3a represents another synthetic image containing two regions. Fig. 3b is its histogram. It can be seen it is the same histogram as Fig. 2b. The difference between this example and the previous one is the background contains two different populations spatially mixed, its weighted average mean μ_2 and pooled standard deviation σ_2 are shown in Table 3d. The initial analysis from the histogram produces 3 regions. During the spatial analysis they are reduced to 2 regions as it can be seen in the Table 3d and Fig. 3c, which is the resulting labelled image.

Finally Fig. 4a represents a rat brain image used in histology research. Due to the inadequate illumination conditions during acquisition, this image cannot be segmented with classical methods. Fig. 4b is its histogram. Fig. 4c the resulting labelled image. Such an accurate segmentation could not be obtained using conventional techniques. Fig. 4d is an axial magnetic resonance image of a human brain. Fig. 4e is the resulting labelled image after having segmented the image into 3 different regions. Fig. 4f is the resulting labelled image after segmenting again the white region in Fig. 4e. It should be noticed the algorithm is able to recover all the detail of the soft tissue in the brain.

6 Conclusion. We developed a powerful algorithm that estimates from global information the parameters of the regions present in an image, and adapts automatically the thresholds of the distributions based on pure local information. This algorithm performs very accurately when the histogram contains significant modes. We are aware that whenever this is not true, a local mode detection algorithm should be used. Work is being done to achieve this goal.

References

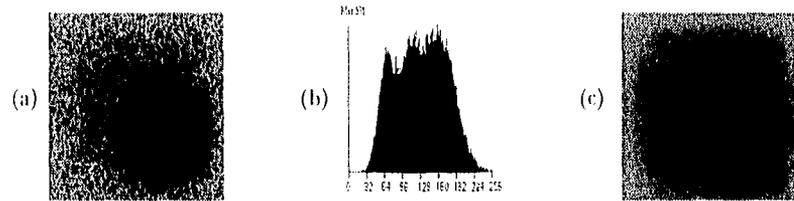
- [1] R. M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features for Image Classification", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, 1973.
- [2] P. Meer, J. M. Jolion, S. Bataonche, "Robust Clustering with Applications in Computer Vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 791-802, 1991.
- [3] Rousseeuw, Peter J., Leroy, Annick M. "Robust Regression & Outlier Detection", New York, John Wiley & Sons, 1987.



(d)

	Real Parameters	Estimated Parameters	Recovered Statistics
Dark Circle	$\mu_1 = 70.00$ $\sigma_1 = 15.00$	$\hat{\mu}_1 = 73.22$ $\hat{\sigma}_1 = 16.30$	$\bar{x}_1 = 70.15$ $s_1 = 15.19$
Light Circle	$\mu_2 = 110.00$ $\sigma_2 = 15.00$	$\hat{\mu}_2 = 113.65$ $\hat{\sigma}_2 = 25.94$	$\bar{x}_2 = 110.51$ $s_2 = 15.91$
Background	$\mu_3 = 160.00$ $\sigma_3 = 30.00$	$\hat{\mu}_3 = 163.85$ $\hat{\sigma}_3 = 29.65$	$\bar{x}_3 = 159.69$ $s_3 = 29.95$

Figure 2: (a) Original Image; (b) Histogram of (a); (c) Classified Image; (d) Statistics.



(d)

	Real Parameters	Estimated Parameters	Recovered Statistics
Circle	$\mu_1 = 70.00$ $\sigma_1 = 15.00$	$\hat{\mu}_1 = 72.02$ $\hat{\sigma}_1 = 21.49$	$\bar{x}_1 = 70.21$ $s_1 = 15.40$
Background 1	$\mu_2 = 110.00$ $\sigma_2 = 15.00$	$\hat{\mu}_{2'} = 147.45$ $\hat{\sigma}_{2'} = 27.03$	$\bar{x}_{2'} = 147.15$ $s_{2'} = 34.66$
Background 2	$\mu_3 = 160.00$ $\sigma_3 = 30.00$		

Figure 3: (a) Original Image; (b) Histogram of (a); (c) Classified Image; (d) Statistics.

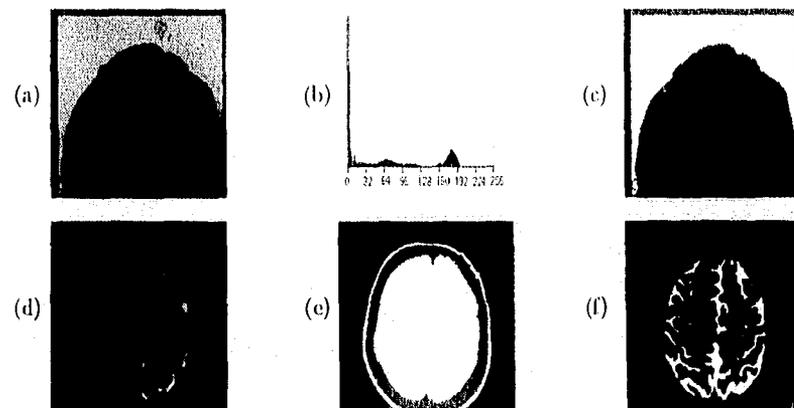


Figure 4: (a) Example 1; (b) Histogram of (a); (c) Classified Image of (a); (d) Example 2; (e) Classified Image of (d); (f) Classified Image of a mask of (e).