



UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO

UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

*“Esquema Estocástico Adaptivo de Vecindad
y Decremento de Temperatura
para Segmentación Markoviana”*

T E S I S

QUE PARA OBTENER EL TÍTULO DE
INGENIERO EN COMPUTACIÓN

PRESENTA:

Roberto Bautista Guzmán

DIRECTOR DE TESIS:

Ing. Ludwin Ventura López López

CODIRECTOR DE TESIS:

Dr. Miguel Moctezuma Flores



MÉXICO, D.F.

NOVIEMBRE de 2007



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.

A mis padres.

Por los valores,
la paciencia,
el apoyo
y sobre todo;
el Amor que me han dado.

AGRADECIMIENTOS

Quiero agradecer a mi director de tesis y amigo Ing. Ludwin Ventura López López por el empeño que puso en ayudarme a realizar este trabajo, por sus consejos y por su paciencia.

A los profesores de la Facultad de Ingeniería con los que tuve el gusto de tomar clases, pues considero haber aprendido mucho de ellos, y no sólo me refiero a los conocimientos propios de cada disciplina, si no, a algo que considero aún más importante; Las anécdotas y experiencias de vida tanto profesionales como no profesionales que compartieron en la clase, pues son cosas que no se encuentran en los libros ni en el mismísimo Google.

A mis amigos, a los que ya tenía antes de entrar a la Facultad de Ingeniería y a los que he venido haciendo desde que entre ahí. En especial a aquellos que me impulsaron y motivaron a terminar esta empresa y a seguir creciendo como ser humano.

ÍNDICE

1. INTRODUCCIÓN	5
1.1 FUNDAMENTACIÓN	5
1.2 DEFINICIÓN DEL PROBLEMA	5
1.3 OBJETIVO	7
1.4 APORTACIÓN DE LA TESIS	7
1.5 ORGANIZACIÓN DE LA TESIS	7
2. SEGMENTACIÓN DE IMÁGENES	9
2.1 GENERALIDADES	9
2.2 SEGMENTACIÓN BASADA EN REGIONES	9
2.2.1 Regiones crecientes	10
2.2.2 Separación y unión de regiones	11
2.3 SEGMENTACIÓN POR WATERSHEDS	12
2.3.1 Construcción de los diques	12
2.3.2 Algoritmo de segmentación por Watershed	15
3. CAMPOS ALEATORIOS DE MARKOV	17
3.1 GENERALIDADES	17
3.1.1 Sites y etiquetas	17
3.1.2 El problema del etiquetado	18
3.2 TOPOLOGÍAS	21
3.2.1 Sistemas de Vecindad	21
3.2.2 Cliques	22

3.3 PROPIEDADES DE MARKOV	23
3.3.1 Campos Aleatorios de Gibbs	24
3.3.2 Equivalencia Markov-Gibbs	26
3.4. MODELOS DE CAMPOS ALEATORIOS DE MARKOV	27
3.4.1 MODELO DE ISING	27
3.4.2 MODELO DE POTTS	28
4. RECOCIDO SIMULADO	30
4.1 GENERALIDADES	30
4.2 ESQUEMAS DE DECREMENTO DE TEMPERATURA	30
5. ESQUEMA DE PROCESAMIENTO	35
5.1 INTRODUCCIÓN	35
5.2 ANÁLISIS PREVIO DE LA IMAGEN	36
5.3 ESQUEMA GENERAL	39
5.4 SEGMENTACIÓN BAJO DIFERENTES ESQUEMAS DE VECINDADES	40
5.4.1 Segmentación en 4-vecindad	40
5.4.2 Segmentación en 8-vecindad	41
5.5 SEGMENTACIÓN BAJO ESQUEMAS DE DECREMENTO DE TEMPERATURA	41
5.5.1 Decremento Geométrico	41
5.5.2 Decremento Logarítmico	42
5.6 PROPUESTA DE MODELO VECINDAD-DECREMENTO DE TEMPERATURA	43

5.6.1 Segmentación 4-vecindad y 8-vecindad con decremento geométrico	43
5.6.2 Segmentación 4-vecindad y 8-vecindad con decremento logarítmico	46
5.6.3 Esquema Estocástico Adoptivo de Vecindad y Decremento de Temperatura.	49
5.7 EVALUACIÓN DE RESULTADOS	59
6. CONCLUSIONES	69
7. REFERENCIAS	71

1 INTRODUCCIÓN

1.1 FUNDAMENTACIÓN

Dentro de los diferentes modelos de segmentación se encuentran los contextuales y los no contextuales. Los tipos de segmentación contextuales tienen mejor desempeño sobre las imágenes que contienen ruido porque incorporan la información contextual de los píxeles analizados, haciendo que la segmentación resultante sea más homogénea. Dentro de este tipo se encuentra la segmentación utilizando la teoría de los Campos Aleatorios de Markov.

Dentro de la segmentación basada en los Campos Aleatorios de Markov (CAM) existen diferentes parámetros o factores que regulan el proceso, y depende de ellos que la segmentación resulte exitosa de acuerdo a nuestros criterios. La optimización de los CAM está ligada al Recocido Simulado. Para realizar una segmentación dentro del CAM se define un *sistema de vecindades* de los *sites* y para desarrollar el recocido simulado se define el *decremento de temperatura*.

Estos factores, de no ser elegidos adecuadamente, pueden producir que el resultado final sea una imagen sub-segmentada o, por el contrario, sobre-segmentada. Por tal motivo mi propuesta es que el esquema de vecindad y decremento de temperatura se modifique a lo largo de la segmentación de acuerdo a los resultados obtenidos a lo largo del proceso.

El esquema típico de la segmentación basada en la teoría de los Campos Aleatorios de Markov, es mediante asignación estática del sistema de vecindades y del decremento de temperatura. Se empieza el proceso con un cierto sistema de vecindad y se termina el proceso con ese mismo sistema de vecindad. De manera similar, se empieza el proceso bajo un cierto esquema de decremento de temperatura y se termina el proceso bajo el mismo. En el esquema típico el proceso no puede detectar los cambios que se van dando en el resultado de la segmentación de iteración en iteración.

1.2 DEFINICIÓN DEL PROBLEMA

La segmentación se considera como un problema de etiquetado en el marco de estudio de los Campos Aleatorios de Markov. A cada uno de los *sites* $s \in S$ en una *lattice* regular, le corresponderá una etiqueta en el dominio L de etiquetas.

Cada *site* tiene un conjunto de posibles etiquetas L , cada una de estas etiquetas corresponde a las clases en las que se desea asignar a los *sites*. Los *sites* están

relacionados entre sí por el sistema de vecindades. Las vecindades más comunes en segmentación de imágenes son las de *4-vecindad* y *8-vecindad*.

Un Campo Aleatorio discreto de Markov sobre una trama finita se define como una colección de variables aleatorias, que corresponden a los *sites* de una *lattice*, cuya distribución de probabilidad es tal que la probabilidad condicional de que una variable determinada tenga un valor particular, dado los valores del resto de las variables, es idéntica a la probabilidad condicional dados los valores del campo en un conjunto pequeño de *sites*, es decir, su *vecindad*.

Dependiendo de la cantidad de *sites* vecinos que se utilicen para modelar el CAM, las regiones homogéneas de segmentación se harán presentes como *sobre-segmentaciones* y *sub-segmentaciones*, de tal manera que la elección del tamaño de vecindad no resulta trivial.

Los CAM están ligados a un método específico de optimación, el *recocido simulado*. El algoritmo de recocido simulado está basado en una analogía entre el recocido de sólidos y la problemática de resolver problemas de optimación combinatoria de gran escala.

Recocido denota un proceso de calentamiento de un sólido a una temperatura en la que sus partículas (*sites*) deformadas cristalizan para producir nuevos granos (*segmentaciones*). La temperatura de recocido o de recristalización, depende del tipo de material, del grado de deformación del mismo.

En el proceso de enfriamiento la temperatura se baja poco a poco. Cada imagen tiene determinada una temperatura crítica que pertenece a una función conexa de energía. Existen principalmente dos esquemas de decremento de temperatura; Decremento *geométrico* y decremento *logarítmico*. La elección del esquema de decremento permite alcanzar esta temperatura y tener un mínimo global.

De acuerdo al esquema de decremento de temperatura se obtienen *sobre-segmentaciones* para decrementos pequeños y *sub-segmentaciones* para decrementos grandes. Para lograr segmentaciones correctas es necesario modificar el esquema de temperatura a lo largo de la segmentación.

En el presente esquema, el problema de etiquetado entre clases de los píxeles en las imágenes, se ataca como una segmentación basada en Campos Aleatorios de Markov. Esto se formula en el marco del teorema de Bayes. En este caso, la solución óptima es definida por el etiquetado Máximo A Posteriori (MAP) y se calcula minimizando la *energía a posteriori* por medio de un algoritmo de minimización global como lo es el *recocido simulado*.

En general, para resolver el problema de segmentación bajo el esquema MAP-CAM, después de haber definido las funciones de potencial para los diferentes cliques, es necesario minimizar la función de energía *a posteriori*. La solución se

encuentra en el espacio de configuraciones, siendo un problema combinatorio, se utiliza el algoritmo de optimización de recocido simulado para encontrar dicha solución global.

El modelo propuesto es adaptivo, pues conforme a los cambios que se vayan presentando en las soluciones parciales de la segmentación, se cambiara de un esquema de 4-vecindad a uno de 8-vecindad y de un esquema de decremento logarítmico a uno geométrico. Así mismo el modelo determina cuando la imagen se ha segmentado adecuadamente y termina el proceso.

1.3 OBJETIVO

La presente tesis tiene como objetivo proponer un esquema de segmentación de imágenes de percepción remota basado en la teoría de los Campos Aleatorios de Markov (CAM). Esta propuesta incorpora una estrategia que combina la selección adaptiva de *vecindad* y *decremento de temperatura* usando el algoritmo de optimización de *recocido simulado*.

1.4 APORTACIÓN DE LA TESIS

La presente tesis hace aportaciones en los siguientes temas:

- Técnicas de procesamiento de imágenes.
- Segmentación basada en Campos Aleatorios de Markov.
- Proponer un esquema de segmentación de imágenes basado en la selección adaptiva de vecindad y decremento de temperatura.

1.5 ORGANIZACIÓN DE LA TESIS

La presente tesis esta organizada en seis capítulos, a continuación se describirá el contenido de los capítulos dos al seis:

- El capítulo dos es relativo a la segmentación de imágenes, se explicará el concepto de segmentación y se revisarán algunos métodos de segmentación como son: segmentación por regiones crecientes y segmentación por separación y unión de regiones.
- En el capítulo tres se dará una introducción a al teoría de los campos aleatorias de Markov. Se definirán los conceptos de *sites* y *cliques*, se estudiaran las diferentes topologías de los sistemas de vecindad y algunos modelos de los Campos Aleatorios de Markov.

- En el capítulo cuatro se verá el modelo de optimación de recocido simulado y su aplicación a dos esquemas de decremento de temperatura; logarítmico y geométrico
- En el capítulo cinco se analizarán varios esquemas de segmentación, haciendo combinaciones entre distintos tipos de vecindad y distintos tipos de decremento de temperatura, para finalmente proponer un modelo adaptivo en cuanto a estos dos factores mencionados.
- Finalmente en el capítulo seis concluimos el presente trabajo de tesis.

2 SEGMENTACIÓN DE IMÁGENES

2.1 GENERALIDADES

Uno de los temas más importantes en el procesamiento de imágenes es la segmentación, mediante la cual una imagen es reducida a piezas de información representadas por regiones homogéneas, las cuales tendrán en común una cierta afinidad establecida por el tipo de análisis requerido.

Los algoritmos de segmentación de imágenes están basados principalmente en una de dos propiedades de los valores de intensidad: similitud y discontinuidad. En la primera categoría el método es dividir la imagen de acuerdo a regiones similares basándose en un conjunto de criterios definidos. En la segunda categoría, el método es dividir la imagen basándose en cambios abruptos de los niveles de intensidad.

A continuación se explicaran dos métodos de segmentación de gran aceptación: segmentación basada en regiones y segmentación por watershed. Ambos métodos son del tipo contextual, del mismo tipo que es la segmentación Markovia, por eso su importancia de estudio.

2.2 SEGMENTACIÓN BASADA EN REGIONES

El objetivo de este método es el de separar la imagen en regiones. Considerando que R representa la imagen completa, el proceso segmenta la imagen R en subregiones, R_1, R_2, \dots, R_n , que cumplen con las siguientes reglas:

- (a) $\bigcup_{i=1}^n R_i = R$
- (b) R_i es una región relacionada, $i=1,2,\dots,n$
- (c) $R_i \cap R_j = \emptyset$ para toda i y j , $i \neq j$

La condición (a) indica que la segmentación debe de ser completa, o sea, cada píxel de la imagen debe de pertenecer a una de las n regiones. La condición (b) indica que los puntos en una región deben de tener una conexión en una cierta forma predefinida. La condición (c) indica que las regiones deben de ser disjuntas, complementando con lo mencionado en la condición (a) quedaría el siguiente enunciado; cada píxel de la imagen debe de pertenecer a una y sólo una de las n regiones.

2.2.1 Regiones crecientes

Este método se basa en agrupar píxeles o subregiones con ciertas características para posteriormente hacer que estas subregiones crezcan. El procedimiento básico es empezar con grupos de puntos, a los que se les denomina *semillas*, a los cuales se les irán agregando píxeles vecinos que cumplan con propiedades similares a las de cada semilla, como puede ser el nivel de gris.

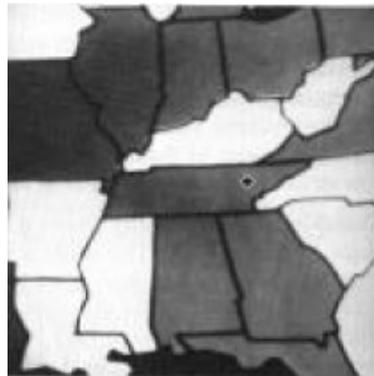
La selección de los puntos de inicio depende por lo general de la naturaleza del problema, en caso de que no se tenga información a priori cada píxel es procesado con las mismas propiedades, esto hará que después de una primera etapa se puedan tener píxeles agrupados en regiones. En caso de que se logren estas agrupaciones podrán servir como grupos semillas para su crecimiento.

Un punto importante en este método es el formular adecuadamente la regla para parar el crecimiento. Básicamente se deberá parar cuando no se encuentre ningún píxel vecino que cumpla con el criterio para ser incluido en la región. Otros criterios que se pueden tomar para mejorar el algoritmo de segmentación son el tamaño, la forma de la región y el parecido entre el píxel candidato y la zona en crecimiento. Estos criterios adicionales sólo aplican si se tienen conocimientos de los resultados deseados.

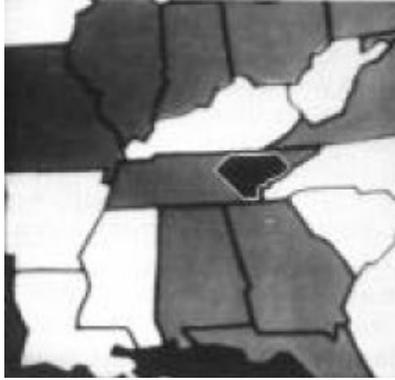
En las siguientes imágenes se muestra una segmentación de este tipo, iniciando en la imagen 2.1 con una semilla y terminando el proceso en la imagen 2.4 con una región encontrada.



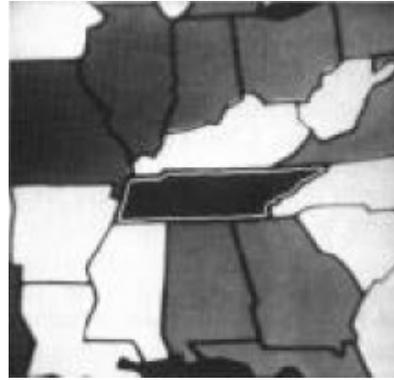
Se inicia con una semilla
imagen 2.1 obtenida de [02]



Crecimiento de la región
Imagen 2.2 obtenida de [02]



Crecimiento de la región
Imagen 2.3 obtenida de [02]



Región encontrada
Imagen 2.4 obtenida de [02]

2.2.2 Separación y unión de regiones

Este procedimiento se basa en separar a una imagen en cuadrantes para posteriormente unir regiones adyacentes que cumplan con condiciones establecidas.

Se toma a la imagen completa imagen R sobre la cual operara la condición P . El método es dividir la imagen R en cuadrantes cada vez más pequeños hasta que se cumpla $P(R_i)=\text{Verdadero}$. Se empieza con la imagen completa, si $P(R)=\text{Falso}$, se divide la imagen en cuadrantes, y si P es Falso para cualquiera de estos cuadrantes se dividirá este a su vez en cuadrantes, y así se seguirá hasta haberse analizado cada cuadrante.

La siguiente imagen muestra como sería la separación en cuadrantes.

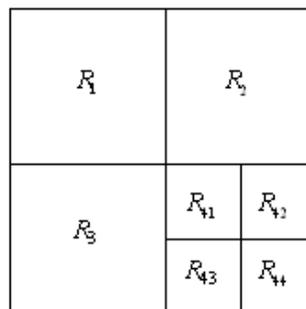


Imagen 2.5

Una vez terminada la etapa de separación seguiría la de unión, esto sería sólo entre regiones adyacentes cuyos píxeles combinados satisficieran la condición P . Esto sería, dos regiones adyacentes R_j y R_k se unirían sólo si $P(R_j \cap R_k)=\text{Verdadero}$.

2.3 SEGMENTACIÓN POR WATERSHEDS

La segmentación por watersheds incorpora algunos conceptos utilizados en la segmentación basada en regiones y en ocasiones produce segmentaciones más estables.

Este concepto se basa en visualizar a la imagen en tres dimensiones, dos espaciales y la escala de valores de grises. Se reconocen tres tipos de puntos bajo esta visualización tridimensional de la imagen: (a) los puntos pertenecientes a un mínimo regional; (b) puntos en los que una gota de agua, si fuese puesta en uno de estos puntos, caería a un mínimo regional; (c) puntos en los que si se vertiera agua, ésta caería igualmente a más de un mínimo regional. El grupo de puntos que satisfagan la condición (b) para un cierto mínimo regional es llamado *contenedor* o *watershed* de ese mínimo. Los puntos dados por la condición (c) son las partes más “altas” en nuestro modelo tridimensional y son llamados *watershed lines*.

El objetivo del algoritmo de segmentación basado por Watersheds es precisamente encontrar las watershed lines. La idea para encontrar estas líneas es la siguiente: Se supone que de cada mínimo regional emergiera agua a una taza uniforme. Cuando el agua en cada watershed ha aumentado su nivel y esta apunto de emerger, se construye un “dique” para impedirlo, el agua sigue aumentado de nivel hasta que se llega a un punto en que sólo se puede observar los diques. Y son precisamente estos diques los que hacen la segmentación y a los que se les llama watershed lines.

2.3.1 Construcción de los diques

El método utilizado para la construcción de las líneas de división es mediante la expansión del agua dentro de los contenedores. El siguiente ejemplo muestra el procedimiento.

Como se muestra en [GO03], la imagen 2.6 muestra a dos contenedores en la etapa n-1 de inundación, la siguiente imagen, imagen 2.7, muestra el siguiente paso, n, donde el agua se ha derramado de un contenedor a otro. Entre estos dos contenedores un dique debe de ser construido para evitar que el agua de ambos contenedores se mezcle.



Figura 2.6

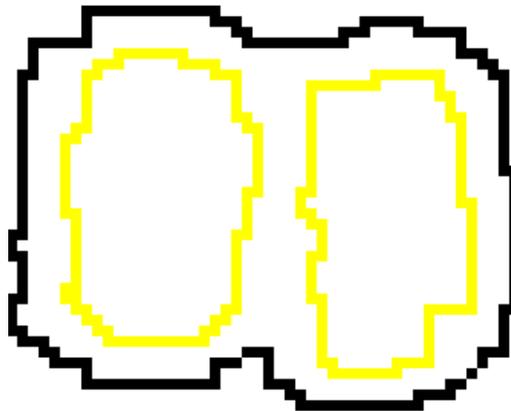


Figura 2.7

Se llamara q a lo mostrado en la figura 2.7, el agua que se ha mezclado de los dos contenedores adjuntos por su sobrellenado. Ahora, el método para la construcción del dique consiste en expandir cada contenedor mostrado en la figura 2.7, la expansión se realiza sobre los píxeles de los bordes como se muestra en la figura 2.8. Hay dos condiciones que debe de cumplir este procedimiento: (1) La expansión debe de estar restringida al espacio ocupado por q , y (2) la expansión no se puede llevara a cabo en puntos donde el agua de ambos contenedores se junte.

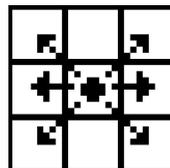


Figura 2.8

La figura 2.9 muestra la expansión de ambos contenedores. En gris claro se tiene la primera expansión que cumple satisfactoriamente ambas condiciones. La segunda expansión, en gris más fuerte, cumple con la primera condición pero en los puntos donde no se cumple la segunda condición es el lugar perfecto para construir el dique, que es la línea central. El dique debe de tener el valor en la escala de grises más alto para que cuando la inundación termine solamente los diques sobresalgan.

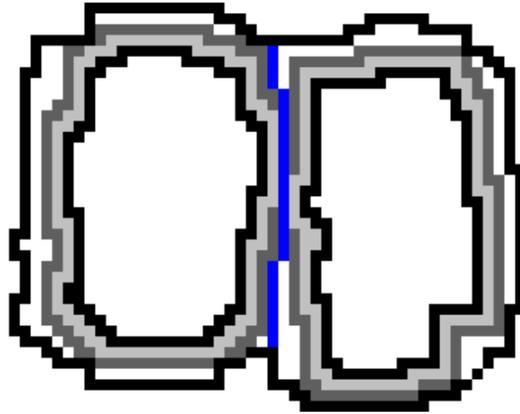


Figura 2.9

En las siguientes figuras se muestra una imagen en escala de grises, su respectiva representación topográfica y la segmentación resultante por watersheds.

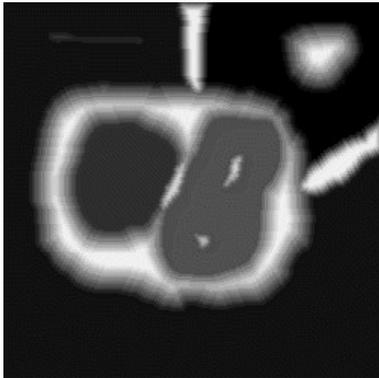
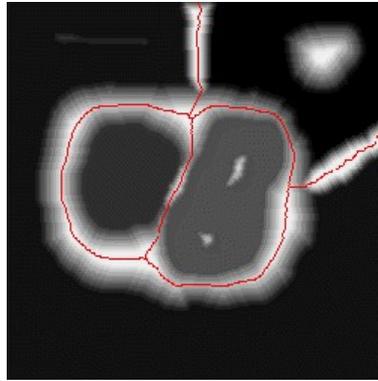


Imagen en escala de grises
Figura 2.10 obtenida de [03]



Representación topográfica de la imagen
Figura2.11 obtenida de [03]



Resultado de la segmentación
Figura 2.12 obtenida de [03]

2.3.2 Algoritmo de segmentación por Watershed

Sean M_1, M_2, \dots, M_R grupos de coordenadas que representan los mínimos regionales de la imagen, la cual se denotara por $g(x,y)$. Sea $C(M_i)$ el grupo de coordenadas que representa los puntos en el contenedor asociado con el mínimo M_i . Y sea $T[n]$ el grupo de coordenadas (s,t) para los cuales $g(s,t) < n$, siendo n el nivel de inundación o nivel del agua. Geométricamente $T[n]$ es el grupo de coordenadas de puntos en la imagen $g(x,y)$ que se encuentra por abajo del plano $g(x,y)=n$.

Lo anterior es para tener conocimiento de que puntos quedan debajo del plano $g(x,y)=n$ conforme se incrementa el nivel de agua (n aumenta).

Sea $C_n(M_i)$ el grupo de coordenadas de puntos del contenedor asociado con el mínimo M_i en el nivel n . Con lo que se tiene la siguiente igualdad

$$C_n(M_i) = C(M_i) \cap T[n] \quad 2-1$$

Sea $C[n]$ la unión de las porciones de los contenedores inundados para el nivel n :

$$C[n] = \bigcup_{i=1}^R C_n(M_i) \quad 2-2$$

Y la unión de todos los contenedores es:

$$C[\max+ 1] = \bigcup_{i=1}^R C(M_i) \quad 2-3$$

Donde $max + 1$ es el nivel máximo de n .

Al ir incrementando el nivel del agua tanto los puntos en $T[n]$ como en $C_n(M_i)$ no son reemplazados, y sólo pueden incrementar o permanecer igual. Por lo cual se tiene que $C[n-1]$ es un subconjunto de $C[n]$ y el algoritmo procede de forma recursiva asumiendo que en el nivel de inundación n , $C[n-1]$ ya se ha construido.

Si Q representa los grupos de agua en los contenedores en $T[n]$, entonces para cada $q \in Q[n]$ hay tres posibilidades:

- (1) $q \cap C[n-1]$ es vacío.
- (2) $q \cap C[n-1]$ contiene una porción de agua de uno de los contenedores.
- (3) $q \cap C[n-1]$ contiene agua de más de un contenedor.

Con esto $C[n]$ es formado recursivamente a partir de $C[n-1]$. La condición (1) ocurre cuando es encontrada una región mínima, en este caso la porción de agua q es incorporada a $C[n-1]$ para formar $C[n]$. La condición (2) ocurre cuando la porción de agua q se encuentra en uno de los contenedores, nuevamente q es incorporada a $C[n-1]$ para formar $C[n]$. Finalmente la condición (3) representa que dos o más porciones de agua de diferentes contenedores están por mezclarse, un dique se debe de construir (con el método mostrado en la sección anterior) para evitar el desbordamiento.

El algoritmo de segmentación por watersheds suele dar buenos resultados porque las segmentaciones siempre son bien definidas, nunca deja líneas de segmentación incompletas.

3 CAMPOS ALEATORIOS DE MARKOV

3.1 GENERALIDADES

En el capítulo anterior se mostraron dos técnicas ampliamente usadas en la segmentación de imágenes, basada en regiones y por watersheds, con el fin de ambientarse en el tema y para servir de referencia a la segmentación Markoviana, que es la utilizada en esta tesis. A partir de este capítulo y hasta el quinto se explicara toda la teoría básica relacionada con los campos aleatorios de Markov y su aplicación a la segmentación de imágenes.

Los Campos Aleatorios de Markov (CAM) son de gran importancia en sistemas de visión artificial por el hecho de que la información que manejan es contextual. Los resultados proporcionan herramientas para codificar restricciones contextuales en la probabilidad a priori. Y el uso de restricciones contextuales es imprescindible para un posible sistema de visión.

Los Campos Aleatorios de Markov proporcionan una manera conveniente y constante de modelar entidades dependientes del contexto tales como píxeles de la imagen y otras características espaciales correlacionadas. Su uso práctico es en gran parte a la equivalencia entre las distribuciones de probabilidad de los Campos Aleatorios de Markov y de Gibbs establecida por el teorema de Hammersley-Clifford.

Los modelos basados en Campos Aleatorios de Markov no son un método de segmentación en si mismos, pero son un modelo estadístico que puede ser usado dentro de los métodos de segmentación. En el tratamiento de imágenes, se utilizan frecuentemente para tomar en cuenta el hecho que la mayoría de los píxeles pertenecen a la misma clase a la que pertenecen sus píxeles vecinos.

Los estudios basados en CAM han tenido éxito al modelar problemas de visión de bajo nivel, como son restauración de imágenes, segmentación, reconstrucción de superficies y análisis de texturas. El uso en visión de alto nivel como el reconocimiento de objetos de los CAM, se ha incrementado recientemente.

3.1.1 Sites y Etiquetas

Los problemas de visión se pueden considerar como problemas de etiquetado en los cuales la solución a un cierto problema es un grupo de etiquetas asignadas a los píxeles de la imagen. A continuación se muestra la teoría de los Campos Aleatorios de Markov extraída de [LI01].

Un problema de etiquetado puede especificarse en términos de un conjunto de sites y de un conjunto de etiquetas. Sea S un conjunto discreto de m sites:

$$S = \{1, 2, \dots, m\} \quad 3-1$$

Un site puede representar un punto o una región bajo el espacio euclidiano como un píxel de una imagen. En una imagen 2D de tamaño $n \times n$ el conjunto de localizaciones de los píxeles se puede denotar por:

$$S = \{(i, j) \mid 1 \leq i, j \leq n\} \quad 3-2$$

Una etiqueta es un evento que puede ocurrir en un site. Sea L el conjunto de etiquetas, las cuales pueden ser continuas o discretas. En el caso continuo, el conjunto de etiquetas puede corresponder al conjunto de números reales o a un intervalo de éste:

$$L_c = [X_l, X_h] \subset R \quad 3-3$$

En el caso discreto, una etiqueta asume un valor discreto del conjunto de M valores:

$$L_d = \{1, \dots, M\} \quad 3-4$$

Para un conjunto ordenado de etiquetas se puede definir una medida numérica (cuantitativa) de semejanza entre dos de sus etiquetas. En el caso de un conjunto de etiquetas sin relación de orden, la medida de semejanza es simbólica (cualitativa), normalmente se elige el valor de igual y no igual. El orden de las etiquetas y su similitud no sólo categoriza problemas de etiquetado, sino que afecta también a la elección de los algoritmos de etiquetado que, a su vez, se refleja en la complejidad computacional.

3.1.2 El problema del etiquetado

El problema de etiquetado es el asignar una etiqueta del conjunto L a cada site del conjunto S . Por ejemplo, en un problema de detección de bordes en una imagen, para cada $i \in S$, donde S contiene a todos los píxeles de la imagen, será asignada una etiqueta del conjunto $L = \{\text{borde}, \text{no borde}\}$.

Se llamara etiquetado a:

$$f = \{f_1, \dots, f_m\} \tag{3-5}$$

donde a cada estado S se le ha asignado una de las etiquetas de L. Cuando a cada site le es asignado una única etiqueta, f_i se ve como una función con dominio S e imagen L. Se puede ver como un mapeo de S a L:

$$f : S \rightarrow L \tag{3-6}$$

Un mapeo con etiquetado discreto se puede ver en la figura 3.1.

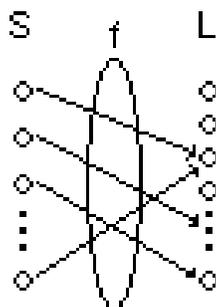


Figura 3.1

En la terminología de Campos Aleatorios de Markov un etiquetado es llamado una configuración. Entonces una imagen se puede ver como un caso específico de configuración, pues esta representada por un conjunto determinado de etiquetas.

Cuando todos los sites tienen el mismo conjunto de etiquetas de L, el conjunto de todas las configuraciones posibles es el producto Cartesiano.

$$F = L \times L \times \dots \times L = L^m \tag{3-7}$$

siendo m el tamaño de S. En un problema con m sites y M etiquetas existe un total de M^m posibles configuraciones en F.

Un problema de etiquetado en visión se puede clasificar, en términos de homogeneidad y continuidad, dentro de las cuatro categorías siguientes:

- PE1: Sites homogéneos con conjunto de etiquetas continuo.
- PE2: Sites homogéneos con conjunto de etiquetas discreto.
- PE3: Sites no homogéneos con conjunto de etiquetas discreto.
- PE4: Sites no homogéneos con conjunto de etiquetas continuo.

Las dos primeras categorías caracterizan el proceso de bajo nivel realizado sobre las imágenes observadas y las dos últimas el proceso de alto nivel realizado.

La restauración de imágenes es un ejemplo de PE1. El conjunto S de sites corresponde a los píxeles de la imagen, y el conjunto L de etiquetas se encuentra en el intervalo real. El problema de restauración es el de estimar la imagen real de una imagen con ruido.

La restauración de imágenes binarias es un buen ejemplo de PE2. La única diferencia de este tipo de problemas con el anterior, es que en éste L es un conjunto discreto de etiquetas.

Otro ejemplo que entra en esta categoría es la segmentación de imágenes. Una imagen es dividida en regiones mutuamente excluyentes, donde cada una tiene propiedades uniformes y homogéneas con valores significativamente diferentes a las regiones vecinas. Una de estas propiedades es el nivel de gris.

La detección de bordes es también un PE2. A Cada site, localizado entre dos píxeles vecinos, se le es asignado una de dos etiquetas {borde, no borde} si existe una diferencia significativa entre los dos píxeles.

El reconocimiento de objetos basándose en sus características es un PE3. Cada site indica una característica de la imagen, como un punto o un segmento de línea. Las etiquetas son discretas y cada una de ellas indica una característica del modelo. La configuración resultante es un mapeo de las características de la imagen al objeto modelo.

La estimación de la orientación a partir de un conjunto de correspondencia de puntos puede ser formulado como un PE4. Un site puede asumir el valor de una matriz real, y una etiqueta sería una transformación admisible (ortogonal, en perspectiva).

Para un problema de etiquetado discreto con m sites y M etiquetas, existen M^m configuraciones posibles. Para un problema de etiquetado continuo las configuraciones son infinitas. Sin embargo sólo algunas de ellas son buenas soluciones y muy pocas más son soluciones óptimas en términos de nuestros criterios. Como definir las soluciones óptimas y como encontrarlas son temas de lo que se llama aproximación a la optimización de etiquetado en visión.

3.2 TOPOLOGÍAS

Las topologías de los Campos Aleatorios de Markov describen las relaciones contextuales entre sites, sirviendo como bases para establecer distribuciones de probabilidad.

3.2.1 Sistemas de Vecindad

Los sites en el conjunto S están interrelacionados en lo que se denomina un sistema de vecindad. Un sistema de vecindad esta definido como:

$$N = \{N_i | \forall i \in S\} \tag{3-8}$$

donde N_i es el conjunto de sites vecinos a i, para los que se cumple las siguientes propiedades:

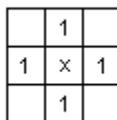
- (1) $i \notin N_i$; un site no es vecino de sí mismo
- (2) $i \in N_{i'} \Leftrightarrow i' \in N_i$; la relación vecinal es mutua entre i' e i.

Para una lattice, rejilla, regular S, el conjunto de vecinos del site i está definido como conjunto de sites dentro de un radio r desde i.

$$N_i = \{i' \in S | [dist(pixel_{i'}, pixel_i)]^2 \leq r, i' \neq i\} \tag{3-9}$$

donde $dist(A,B)$ es la distancia Euclidiana entre A y B, y r es un entero.

En el sistema de vecinos de primer orden, también llamado sistema de 4-vecindad, cada site interior tiene 4 vecinos, como se muestra en la figura 3.2, denotándose por x el site considerado.



4-vecindad
Figura 3.2

En el sistema de vecinos de segundo orden, también llamado sistema de 8-vecindad, cada site interior cuanta con 8 vecinos, como se muestra en la figura 3.3.

2	1	2
1	x	1
2	1	2

8-vecindad
Figura 3.3

Siguiendo en este sentido se puede ver en la figura 3.4 los sistemas de vecinos hasta el orden 5, $n = 1, \dots, 5$.

5	4	3	4	5
4	2	1	2	4
3	1	x	1	3
4	2	1	2	4
5	4	3	4	5

Distintas vecindades
Figura 3.4

Cuando están ordenados los elementos en S el conjunto de vecinos se puede determinar más explícitamente. Por ejemplo, si $S = \{1, \dots, m\}$ es un conjunto ordenado de sites y sus elementos indican los píxeles de una imagen en 1D un site interior $i \in \{2, \dots, m-1\}$ cuenta con dos vecinos, $N_i = \{i-1, i+1\}$, y los sites de las fronteras (1 y m) tienen uno cada uno, $N_1 = \{2\}$ y $N_m = \{m-1\}$. Si se habla de una rejilla regular correspondiente a los píxeles de una imagen en dos dimensiones de tamaño $n \times n$, cada site interno (i,j) tienen cuatro vecinos, $N_{i,j} = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}$, los estados de la frontera tiene tres y los de las esquinas dos.

En el caso de que S sea irregular, el conjunto de vecinos de i , N_i , se define mediante la expresión (3-9), que son los sites contenidos dentro del radio r.

3.2.2 Cliques

Sabiendo que S es un conjunto de sites y N un sistema de vecindad, entonces, el par (S, N) constituye un grafo, donde N establece al relación de los nodos de S de acuerdo al sistema de vecinos definido.

Un clique c para (S, N) se define como un subconjunto de S, que puede consistir de un site, de dos, de tres y así sucesivamente, que se denotan por C_1, C_2, \dots , donde:

$$C_1 = \{i | i \in S\} \tag{3-10}$$

$$C_2 = \{\{i, i'\} | i' \in N_i, i \in S\} \tag{3-11}$$

$$C_3 = \{\{i, i', i''\} | i, i', i'' \in S\} \tag{3-12}$$

El conjunto de todos los cliques para (S,N) es: $= C_1 \cup C_2 \cup C_3 \dots$

La figura 3.5 muestra distintos tipos de cliques para sistemas de vecinos de primer orden



Figura 3.5

La figura 3.6 muestra distintos tipos de cliques, que junto con los mostrados en la figura 3.5, corresponden al sistema de vecinos de segundo orden.



Figura 3.6

3.3 PROPIEDADES DE MARKOV

Sea $F = \{F_1, \dots, F_m\}$ una familia de variables aleatorias definida en el conjunto de S, en donde cada variable aleatoria F_i toma un valor f_i en L. Se llamara Campo Aleatorio a la familia de variables aleatorias F. Se utilizara la notación $F_i = f_i$ para denotar que F_i toma el valor de f_i , y al evento en conjunto $F=f$, donde f es una configuración de F. Para un conjunto discreto de etiquetas L, la probabilidad de que una variable aleatoria F_i tome el valor f_i es denotada por $P(F_i = f_i)$ o simplemente $P(f_i)$.

Una familia de variables aleatorias F se dice que es un Campo aleatorio de Markov sobre la vecindad N de S, si y sólo si las dos siguientes condiciones se satisfacen:

$P(f) > 0$	Positividad	3-13
$P(f_i f_{s-\{i\}}) = P(f_i f_{N_i})$	Markovianidad	3-14

donde $S-\{i\}$ es el conjunto de etiquetas en todos los sites exceptuando a $\{i\}$ y

$$f_{N_i} = \{f_i | i' \in N_i\} \quad 3-15$$

es el conjunto de etiquetas que pertenecen a los sites de la vecindad i .

La condición (3-13) es para exigir que F sea un Campo Aleatorio. La condición (3-14) es la llamada característica local, sólo las etiquetas vecinas tienen interacción entre ellas. En otras palabras, un Campo Aleatorio de Markov se define como una colección de variables aleatorias, cuya distribución de probabilidad es tal que la probabilidad condicional de que una variable determinada tenga un valor particular, dado los valores del resto de las variables, es idéntica a la probabilidad condicional dados los valores de los vecinos.

Un CAM puede tener la característica de homogeneidad. Se dice que es homogéneo si $P(f_i | f_{N_i})$ es independiente de la posición relativa del site i en S .

3.3.1 Campos Aleatorios de Gibbs

Un conjunto de variables aleatorias F es un Campo Aleatorio de Gibbs en S con respecto a N si y sólo si sus configuraciones siguen una distribución de Gibbs. Una distribución de Gibbs tiene la siguiente forma:

$$P(f) = \frac{1}{Z} \exp\left(-\frac{1}{T} U(f)\right) \quad 3-16$$

donde

$$Z = \sum_{f \in F} e^{-\frac{1}{T} U(f)} \quad 3-17$$

es una constante de normalización. T es una constante llamada temperatura. $U(f)$ es la función de energía definida por:

$$U(f) = \sum_{c \in C} V_c(f) \quad 3-18$$

que es la suma de los cliques potenciales $V_c(f)$, dependiendo su valor de la configuración local del clique c .

Dos de las propiedades de los Campos Aleatorios de Gibbs (CAG) son su homogeneidad y su isotropía; Se dice que es homogéneo si $V_c(f)$ es independiente de la posición relativa del clique c . Y se dice que es isotrópico si V_c es independiente de de la orientación de c . Por simplicidad las distribuciones de CAG se consideran homogéneas e isotrópicas pues son más fáciles de especificar.

En la expresión (3-16) $P(f)$ mide la probabilidad de ocurrencia de una configuración en particular. Las configuraciones más probables son aquellas con energías más bajas. La temperatura T controla la forma de la distribución, cuando ésta es alta, todas las configuraciones tienden a estar igualmente distribuida. Por el lado contrario, cuando es cercana a cero, la distribución se concentra alrededor del mínimo global de la energía.

Para los problemas de etiquetado discreto, una función clique potencial $V_c(f)$ puede ser especificada por un número de parámetros. Por ejemplo, sea $f_c = (f_i, f_{i'}, f_{i''})$ la configuración local en el clique triple $c = \{i, i', i''\}$, f_c toma un número finito de estados y, en consecuencia, $V_c(f)$ toma un número finito de valores.

La energía de una distribución de Gibbs puede ser expresada, por conveniencia, como una suma de varios términos, cada uno correspondiente a cliques de de distinto tamaño:

$$U(f) = \sum_{\{i\} \in C_1} V_1(f_i) + \sum_{\{i, i'\} \in C_2} V_2(f_i, f_{i'}) + \sum_{\{i, i', i''\} \in C_3} V_3(f_i, f_{i'}, f_{i''}) + \dots \quad 3-19$$

Como V_1 , V_2 y V_3 no están en función de i , i' o i'' , se considera que la expresión (3-19) es una distribución de Gibbs homogénea.

Un caso importante es cuando se consideran los cliques de tamaño dos como máximo. En este caso la energía se puede expresar de la siguiente forma:

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'}) \quad 3-20$$

Considerando la ecuación anterior la probabilidad condicional puede escribirse como:

$$P(f_i | f_{N_i}) = \frac{e^{-[V_i(f_i) + \sum_{i' \in N_i} V_2(f_i, f_{i'})]}}{\sum_{f_i \in L} e^{-[V_i(f_i) + \sum_{i' \in N_i} V_2(f_i, f_{i'})]}} \quad 3-21$$

3.3.2 Equivalencia Markov-Gibbs

Los Campos Aleatorios de Markov se caracterizan por sus propiedades locales (Markovianas) mientras que los Campos Aleatorios de Gibbs se caracterizan por sus propiedades globales (la distribución de Gibbs). Es el teorema de Hammersley-Clifford el que establece la equivalencia entre estos dos tipos de características.

El teorema menciona que F es un Campo Aleatorio de Markov en S con respecto a N si y sólo si F es un campo Aleatorio de Gibbs en S con respecto a N.

El teorema anterior se demostrara probando que un Campo Aleatorio de Gibbs es un Campo Aleatorio de Markov. Sea P(f) una distribución de Gibbs en S con respecto al sistema de vecinos N. Entonces teniendo en cuenta la siguiente probabilidad condicional:

$$P(f_i | f_{S-\{i\}}) = \frac{P(f_i, f_{S-\{i\}})}{P(f_{S-\{i\}})} = \frac{P(f)}{\sum_{f'_i \in L} P(f')} \quad 3-22$$

siendo f' cualquier configuración que contenga todos los sites de f excepto posiblemente i.

Escribiendo $P(f) = Z^{-1} x e^{-\sum_{\alpha \in C} V_{\alpha}(f)}$ tenemos:

$$P(f_i | f_{S-\{i\}}) = \frac{e^{-\sum_{\alpha \in C} V_{\alpha}(f)}}{\sum_{f'_i} e^{-\sum_{\alpha \in C} V_{\alpha}(f')}} \quad 3-23$$

Si se divide C en dos conjuntos A y B, A siendo el conjunto de cliques que contienen i y B el conjunto de cliques que no contienen i. La expresión (3-23) se puede reescribir como:

$$P(f_i | f_{S-\{i\}}) = \frac{[e^{-\sum_{\alpha \in A} V_{\alpha}(f)}][e^{-\sum_{\alpha \in B} V_{\alpha}(f)}]}{\sum_{f'_i} \{[e^{-\sum_{\alpha \in A} V_{\alpha}(f')}][e^{-\sum_{\alpha \in B} V_{\alpha}(f')}] \}} \quad 3-24$$

como $V_c(f) = V_c(f')$ para cualquier claqué c que no contenga i , $e^{-\sum_{c \in B} V_c(f)}$ se cancela del numerador y denominador. Por lo tanto, esta probabilidad depende sólo de los potenciales de los cliques que contienen i :

$$P(f_i | f_{S-\{i\}}) = \frac{e^{-\sum_{c \in A} V_c(f)}}{\sum_{f'_i} e^{-\sum_{c \in A} V_c(f')}} \quad 3-25$$

La expresión de arriba muestra que sólo depende de las etiquetas vecinas a i . Con Esto se prueba que un Campo Aleatorio de Gibbs es un Campo Aleatorio de Markov.

3.4. MODELOS DE CAMPOS ALEATORIOS DE MARKOV

3.4.1 Modelo de Ising

En este modelo son únicamente dos etiquetas las que dan la información de contexto. Debido a su simplicidad y a su bajo coste computacional son ampliamente utilizados. Se codifican en la función de energía de Gibbs como funciones clique potenciales para cada par de sites. Para cliques de máximo dos sites la función de energía queda:

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'}) \quad 3-26$$

que es un caso especial de la expresión (3-19), a la que llamaremos energía de segundo orden. Con una selección apropiada de V_1 y V_2 se puede especificar un Campo Aleatorio de Markov.

Cuando $V_1(f_i) = f_i G_i(f_i)$ y $V_2(f_i, f_{i'}) = \beta_{i,i'} f_i f_{i'}$, donde $G(x)$ es una función arbitraria y $\beta_{i,i'}$ son constantes que reflejan la interacción entre los sites i e i' , la función de energía es:

$$U(f) = \sum_{\{i\} \in C_1} f_i G_i(f_i) + \sum_{\{i,i'\} \in C_2} \beta_{i,i'} f_i f_{i'} \quad 3-27$$

El modelo anterior recibe el nombre de auto-modelo. Si las f_i del auto-modelo toman valores en el conjunto de etiquetas discreto $L=\{0,1\}$, se dice que se tiene un modelo auto-logístico. La correspondiente función de energía es de la forma:

$$U(f) = \sum_{\{i\} \in C_1} \alpha_i f_i + \sum_{\{i,i'\} \in C_2} \beta_{i,i'} f_i f_{i'} \quad 3-28$$

Cuando el sistema de vecindad es el de 4-vecinos en una lattice 2D el modelo auto-logístico recibe el nombre de modelo Ising. La probabilidad condicional para el modelo Ising es:

$$P(f_i | f_{N_i}) = \frac{e^{\alpha_i f_i + \sum_{i' \in N_i} \beta_{i,i'} f_i f_{i'}}}{\sum_{f_i \in \{0,1\}} e^{\alpha_i f_i + \sum_{i' \in N_i} \beta_{i,i'} f_i f_{i'}}} = \frac{e^{\alpha_i f_i + \sum_{i' \in N_i} \beta_{i,i'} f_i f_{i'}}}{1 + e^{\alpha_i + \sum_{i' \in N_i} \beta_{i,i'} f_{i'}}} \quad 3-29$$

Cuando la distribución es homogénea $\alpha_i = \alpha$ y $\beta_{i,i'} = \beta$ sin importar i ni i' .

3.4.2 Modelo de Potts

El modelo de Potts o modelo logístico multinivel, es el modelo de Ising generalizado para más de dos etiquetas. Hay $M>2$ etiquetas discretas en el conjunto $L=\{1,\dots,M\}$. En este tipo de modelos un clique potencial depende del tipo c (tamaño y orientación) del clique y de la configuración local $f_c = \{f_i | i \in c\}$. Para cliques que contienen más de un site las funciones de clique potenciales se definen como:

$$30 \quad V_c(f) = \begin{cases} \zeta_c & \text{si todos los sites en } c \text{ tiene la misma etiqueta} \\ -\zeta_c & \text{en otro caso} \end{cases} \quad 3-$$

donde ζ_c es el potencial para los cliques de tipo c . Para cliques de un sólo site, depende de la etiqueta asignada al site:

$$V_c(f) = V_c(f_i) = \alpha_I \quad \text{si } f_i = I \in L_d \quad 3-31$$

donde α_I es potencial para la etiqueta I.

Se asumirá que un modelo de Potts es de segundo orden, como en la expresión (3-26), si los parámetros α (para cliques de un estado) y β (para cliques de dos estados) son distintos de cero, estos parámetros y los cliques a los que están asociados se muestran en la figura 4.1.

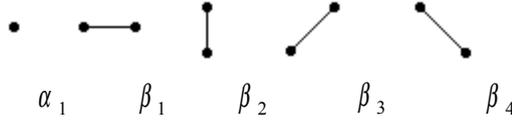


Figura 4.1

La función clique potencial para cliques de dos sites es:

$$V_2(f_i, f_{i'}) = \begin{cases} \beta_c & \text{si los sites en el clique } \{i, i'\} = c \in C_2 \\ & \text{tienen la misma etiqueta} \\ -\beta_c & \text{en caso contrario} \end{cases} \quad 3-32$$

donde C_2 es el conjunto de cliques de dos estados. Para un sistema de 4-vecinos hay cuatro tipos de cliques de dos estados y, por lo tanto, cuatro parámetros β . Si el modelo se considera isotópico los cuatro parámetros toman el mismo valor y la expresión (3-32) se simplifica. Debido a su simplicidad, el modelo de Potts es ampliamente utilizado para modelar regiones y texturas. La probabilidad condicional de este tipo de modelos es expresada como:

$$P(f_i = I | f_{N_i}) = \frac{e^{-\alpha_I - \beta n_i(I)}}{\sum_{I=1}^M e^{-\alpha_I - \beta n_i(I)}} \quad 3-33$$

donde $n_i(I)$ es el número sites en N_i que están etiquetados por I.

4 RECOCIDO SIMULADO

4.1 GENERALIDADES

Los modelos de optimación se agrupan en modelos determinísticos y modelos aleatorios o estocásticos. Los modelos determinísticos reciben a un conjunto de variables, cuyo comportamiento o valor en el futuro se supone cierto, es decir, no sujetos a variación. Un ejemplo de este tipo de método es; El Método de Newton. En cuanto a los modelos estocásticos, se caracterizan por que una o más variables pueden tener diferentes valores futuros, y cada una de ellas tiene una cierta probabilidad de ocurrencia. Algunos ejemplos de estos métodos son; Recocido Simulado y Algoritmos Genéticos.

4.2 ESQUEMAS DE DECREMENTO DE TEMPERATURA

El modelo de decremento de temperatura que utilizaremos en el presente trabajo será el de Recocido Simulado que como se menciona es del tipo estocástico.

El término *recocido* es tomado de la metalurgia y se refiere al proceso de llevar a un sólido a un estado de equilibrio térmico. El principio es el siguiente: Dada una sustancia, no todas las moléculas tienen la misma energía, sino que se encuentra en diferentes niveles, el menor de los cuales se denomina estado fundamental. Si la sustancia está a 0°K todas sus moléculas se encuentran en su estado fundamental, si está a temperaturas mayores las moléculas ocupan estados superiores de energía.

El proceso de recocido en metalurgia se basa en reblandecer un sólido a una temperatura elevada y luego enfriarlo lentamente hasta que las partículas, por sí solas, se van colocando en el *estado fundamental del sólido*. El proceso pasa por diferentes fases cada vez a menores temperaturas.

En optimización el recocido simulado se realiza tomando la solución actual en una cierta iteración, esta solución es perturbada para producir una nueva solución que puede reemplazar a la anterior o no. El reemplazo será aceptado o no en función de una regla de aceptación. El proceso se repite iterativamente hasta que se considera que se llegó al óptimo, esto es cuando la solución converge.

El criterio de metrópolis es la regla de aceptación que se utiliza en el recocido simulado y es la siguiente:

- Condición 1

Si la nueva configuración x' tiene una energía menor que la anterior x , $U(x') < U(x)$, la reemplaza. $x = x'$

- Condición 2

Si no, la probabilidad de que x' se acepte es

$$P(x = x') = \exp\left(\frac{\Delta U}{T}\right) \quad 4-1$$

donde

$\Delta U = (U(x') - U(x))$ y T es la temperatura actual.

En la práctica la Condición 2 se realiza mediante la generación de un número aleatorio, n , en el intervalo $(0,1)$ el cual es comparado con la expresión 4-1:

Si $n < P(x=x')$, x' reemplaza a x como solución actual. $x = x'$

Si $n \geq P(x=x')$, x se usa nuevamente como solución para la próxima iteración.

El diagrama de flujo del algoritmo de metrópolis es el mostrado en la figura 4-1.

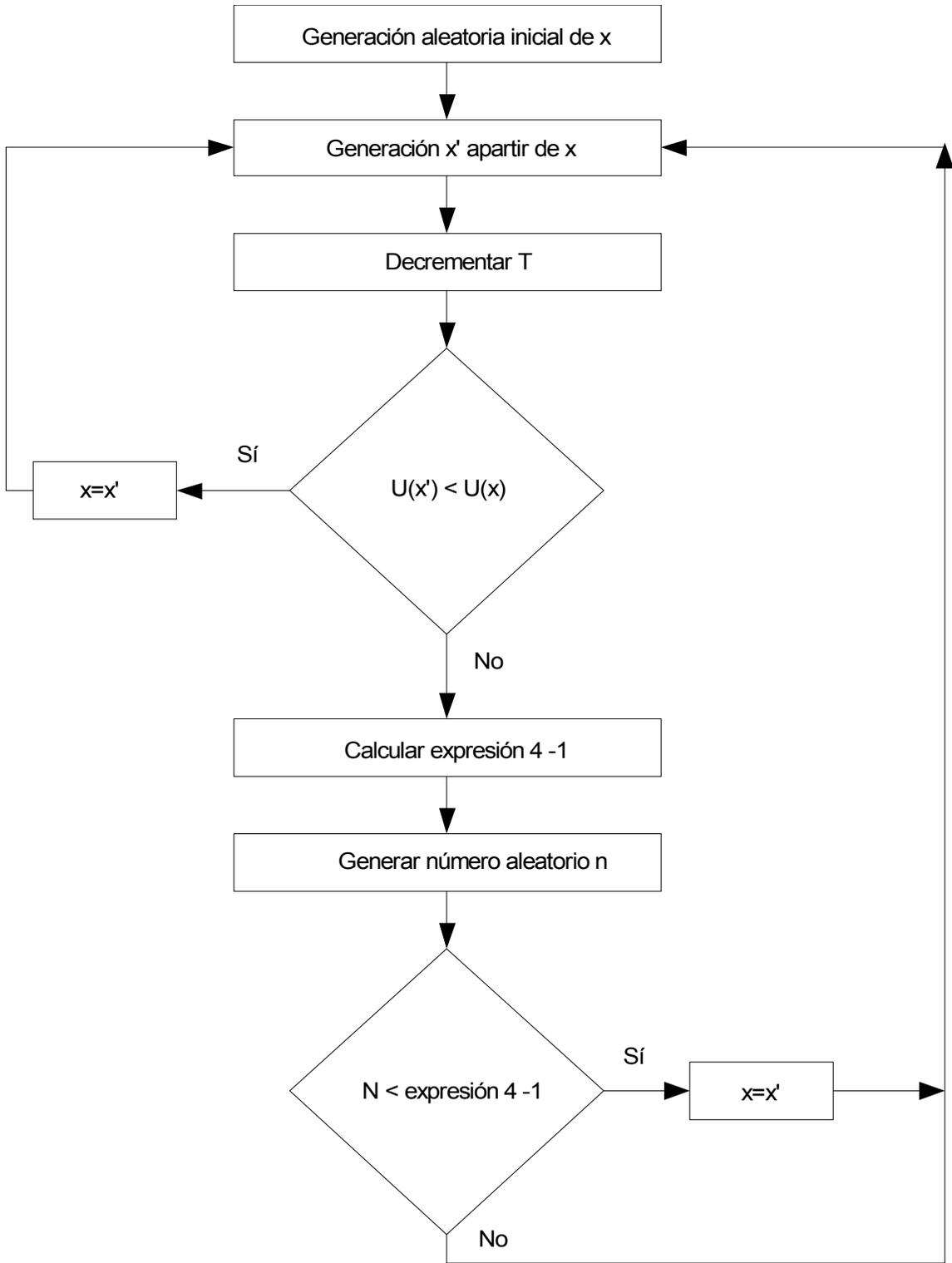


Diagrama de flujo del algoritmo de Recocido Simulado
 Figura 4.1

Como se ha visto en la parte inicial de este capítulo el factor de temperatura T se va disminuyendo conforme se avanza en las iteraciones. El algoritmo inicialmente para valores grandes de T puede aceptar soluciones con un valor objetivo mucho mayor a la solución actual, más sin embargo, cuando T decrece únicamente pequeñas desviaciones serán aceptadas y finalmente, cuando el valor de T se aproxima a cero, no se aceptarán desviaciones. Este hecho significa que el algoritmo de recocido simulado puede escapar de mínimos locales dando una solución global.

Heurísticamente se tienen los siguientes esquemas de decremento:

- *Decremento de tipo geométrico*

$$T^{(iter+1)} = \alpha T^{(iter)} \quad 4-2$$

con α en el rango $(0,1)$ y siendo $iter$ el número de iteración en curso. Las evidencias empíricas señalan que valores de α comprendidos entre 0.8 y 0.99 dan los mejores resultados.

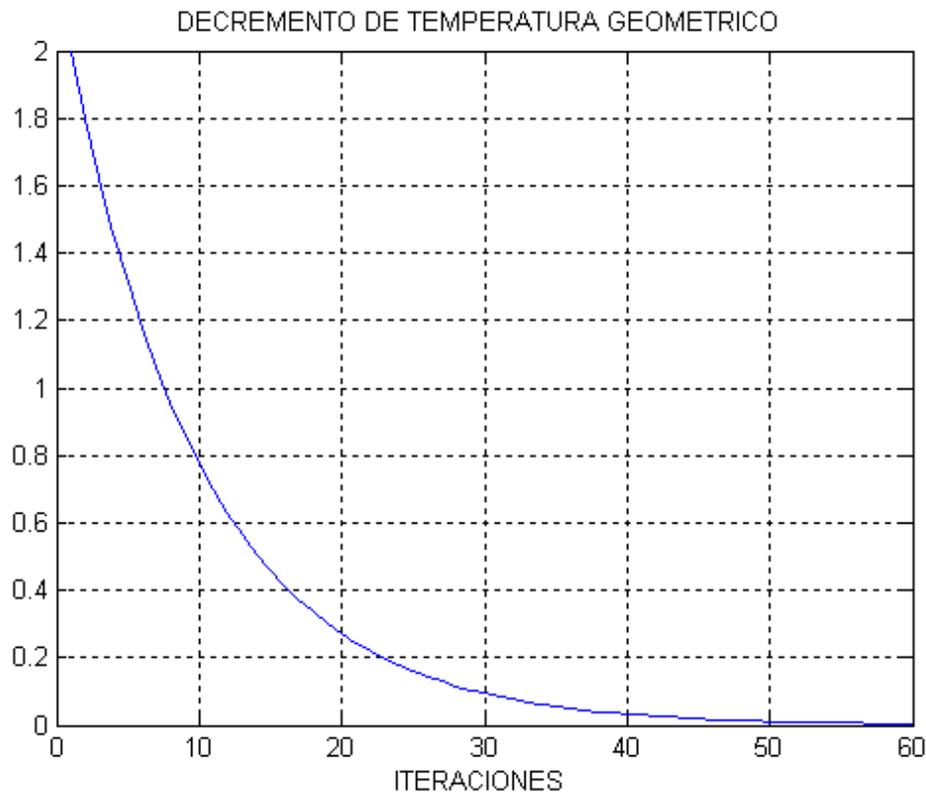


Figura 4.2

- *Decremento Logarítmico*

$$T^{(iter)} = \frac{T_0}{\ln(iter + 1)} \quad 4.3$$

donde T_0 es la temperatura inicial.

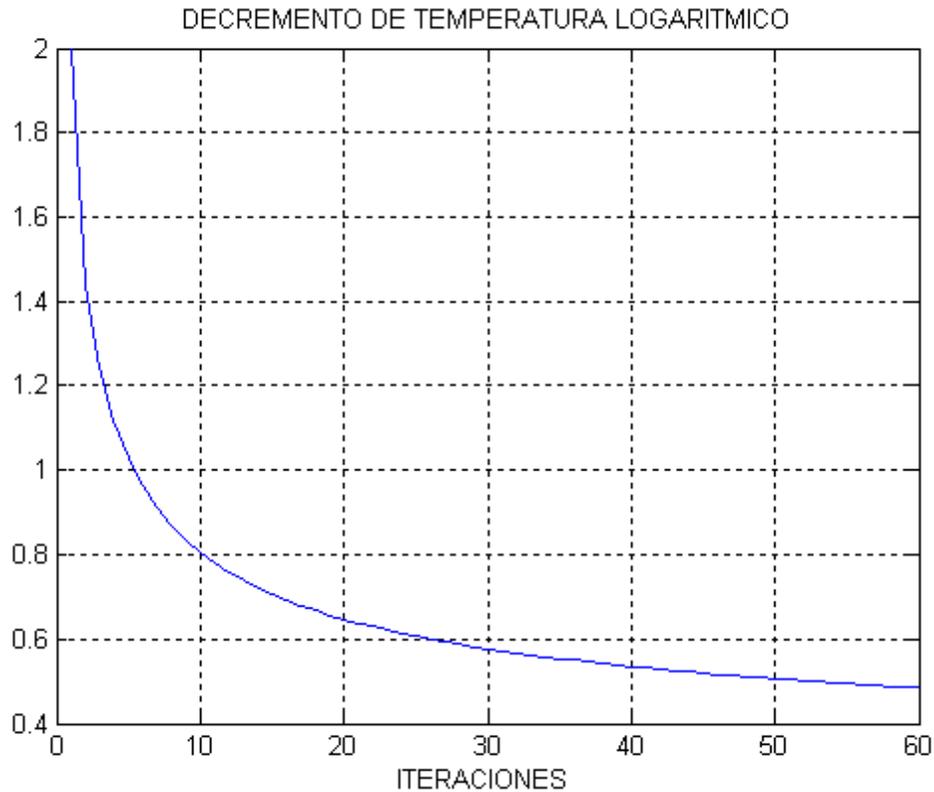


Figura 4.3

Es importante analizar cual de estos dos esquemas da mejores resultados y es precisamente en este punto donde la presente tesis toma importancia, ya que la elección adecuada de uno de estos esquemas, o una combinación de ellos, puede arrojar mejores resultados en el proceso de segmentación; tanto en la reducción de iteraciones (reducción de tiempo) como en la precisión de la segmentación.

5 ESQUEMA DE PROCESAMIENTO

5.1 INTRODUCCIÓN

En el presente capítulo se presentará el esquema propuesto para la segmentación de imágenes digitales. En particular la imagen con la que se trabajara a lo largo del capítulo es la fotografía satelital mostrada en la figura 5.1 [04]. Esta imagen satelital fue adquirida por el sensor satelital de Radar de Apertura Sintética. La imagen muestra un glaciar, el área gris representa porciones de hielo y el área oscura cuerpos de agua líquida. La imagen esta en un formato de mapa de bits de 500 por 960 pixeles en escala de grises de 256 tonos.

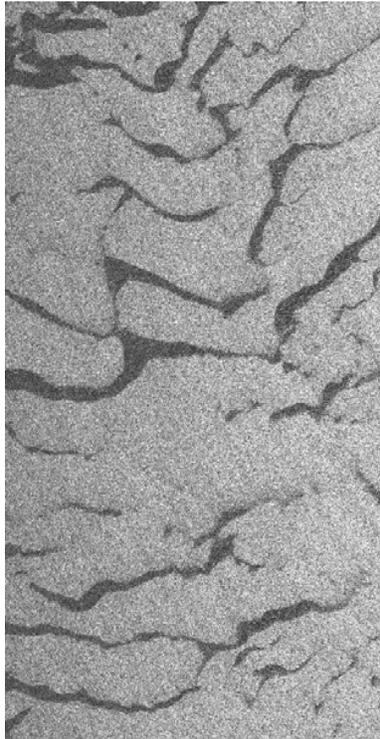


Imagen original de 500x960 pixeles
Figura 5.1

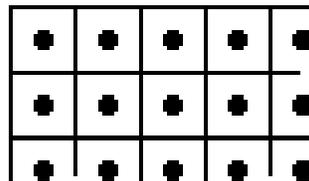
El esquema propuesto en este trabajo se aplicará en imágenes de Radar de Apertura Sintética. A simple vista, la discriminación entre clases no resulta tarea fácil, la necesidad de una segmentación se justifica a que en algunas partes de la imagen no se puede distinguir con precisión la parte de hielo y la de agua, por lo tanto es necesario un algoritmo de segmentación para poder clasificar que partes de la imagen son agua y cuales hielo utilizando un criterio contextual.

El esquema de segmentación utilizado en la presente tesis es Markoviano, el cual en este tipo de imágenes es apropiado, dada la naturaleza ruidosa de estas. Las imágenes adquiridas por el Radar de Apertura Sintética son interesantes en su análisis ya que presentan un comportamiento ruidoso dada la naturaleza coherente de las ondas de radar. El ruido que se presenta presenta la característica de ser de naturaleza multiplicativa y recibe el nombre de ruido speckle.

5.2 ANÁLISIS PREVIO DE LA IMAGEN

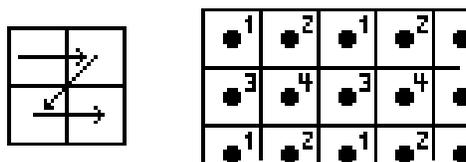
Se realizó un análisis previo de la imagen al dividirla en cuatro sub-bandas. Este análisis acelera el proceso de segmentación tomando en cuenta todos los pixels presentes en la imagen.

La forma en que es dividida la imagen en cuatro bandas se muestra en los siguientes esquemas; La figura 5.2 muestra esquemáticamente los pixeles de la esquina superior izquierda de una imagen, la cual queremos repartir en cuatro bandas, cada punto representa un pixel.



Vista esquemática de los pixeles de una imagen
Figura 5.2

Cada cuatro pixeles adyacentes (que forman un cuadrado) se clasifican de acuerdo con el siguiente ejemplo. Se numeran los pixeles como se muestra en la figura 5.3



Numeración de los pixeles
Figura 5.3

cada una de estas cuatro clasificaciones formarán una nueva imagen que tendrá como dimensión una cuarta parte que la original. La figura 5.4 muestra otra forma de ver esta clasificación y la figura 5.5 muestra como a partir de este criterio

podemos construir cuatro subimágenes, simplemente poniendo contiguamente los pixels según su clasificación en una nueva matriz.

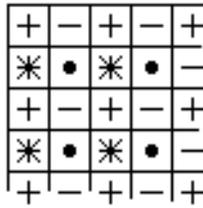
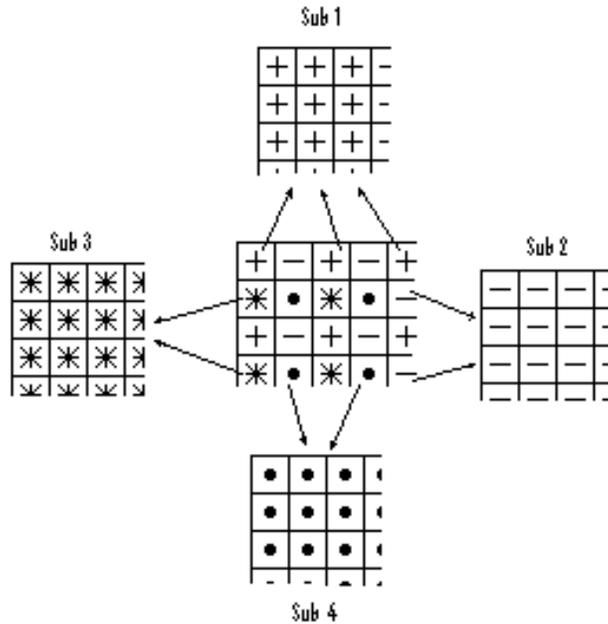
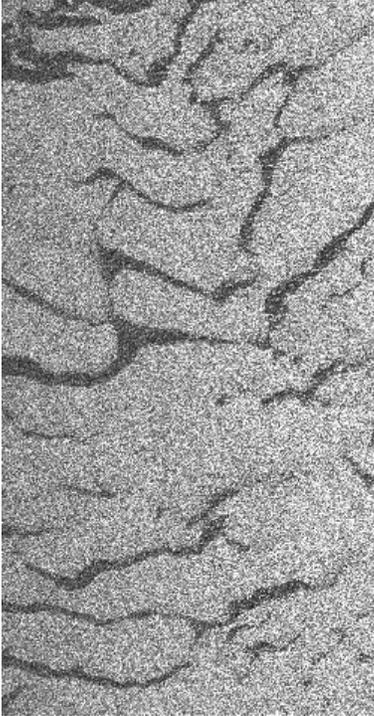


Imagen Original
Figura 5.4

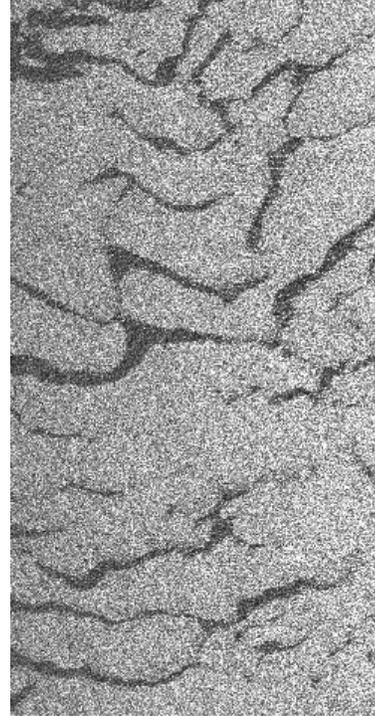


Bandas generadas a partir de la imagen original
Figura 5.5

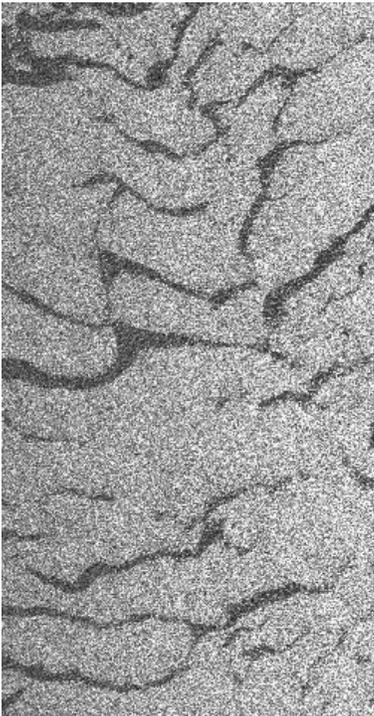
Una vez teniendo las cuatro imágenes generadas a partir de la imagen original, se ejecutará el algoritmo propuesto sobre ellas, tomando secuencialmente una de ellas en cada iteración y a partir del resultado de cada una se tomará como la entrada en la siguiente iteración, por lo tanto, cada cuatro iteraciones se volverá a empezar con la primera y así sucesivamente. Como la imagen original se ha dividido en cuatro, es recomendable que el número de iteraciones totales sea múltiplo de 4. Esta forma de proceso de la imagen no solo ahorra el tiempo en las iteraciones del cómputo, si no que mantiene la información original de la imagen (de forma distribuida) durante todo el proceso, cosa que no pasa cuando se obtiene un promedio de la imagen para reducirla. Las siguientes imágenes muestran la división de la imagen original (figura 5.1) en cuatro subimágenes. Aparentemente las cuatro imágenes son iguales pero vistas en detalle se aprecia la diferencia.



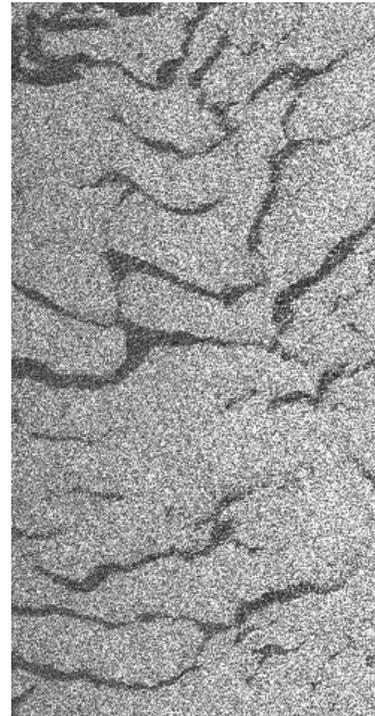
Sub-banda superior izquierda 125x240
Figura 5.6



Sub-banda superior derecha 125x240
Figura 5.7



Sub-banda inferior izquierda 125x240
Figura 5.8



Sub-banda inferior derecha 125x240
Figura 5.9

5.3 ESQUEMA GENERAL

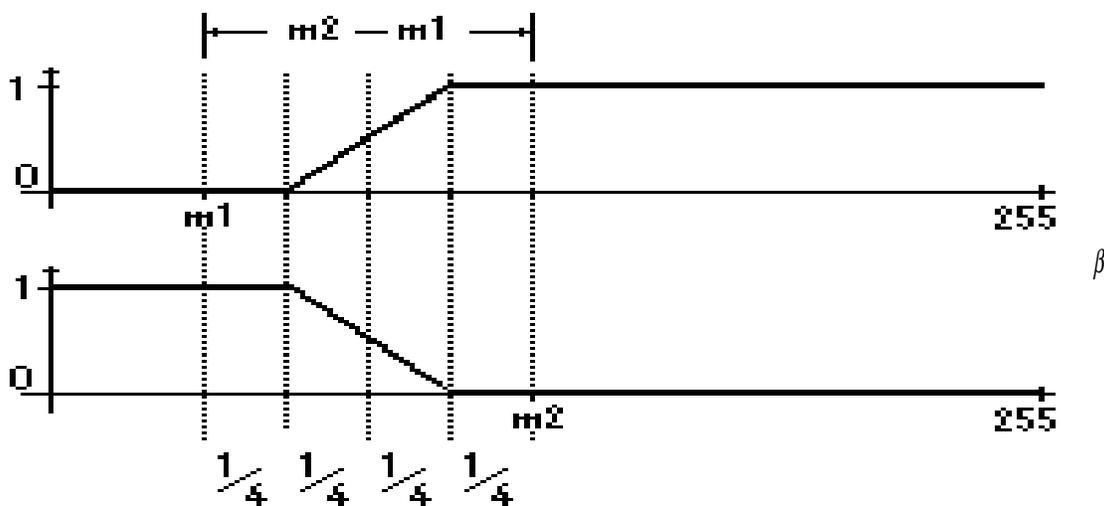
En el presente esquema, la solución óptima esta dada por la probabilidad Máxima A Posteriori (MAP). La aproximación MAP esta compuesta por la suma del modelo de probabilidad conjunta $P(X|Y)$ y la probabilidad a *priori* $P(Y)$.

El modelo de probabilidad conjunta $P(X|Y)$ esta asociado al modelo de energía conjunta $U(X|Y)$. Este modelo de energía es definido mediante ventanas de entrenamiento que calculan la media correspondiente a cada clase.

El modelo de probabilidad a *priori* $P(Y)$ esta asociado al modelo de energía a *priori* $U(Y)$. Las restricciones contextuales de este modelo se basan en la teoría de los Campos Aleatorios de Markov. En esta tesis se considera que el Campo Aleatorio de Markov es homogéneo e isotrópico, por lo que el parámetro β asociado a los cliques es independiente de la posición y orientación del clique.

Dicho lo anterior, el primer paso es crear los modelos de energía conjunta. Se toman muestras de las distintas clases y se calcula sus respectivos promedios, en nuestro estudio son dos clases (hielo y agua).

Una vez teniendo los promedios de las clases se crean sus funciones de energía normalizadas a de 0 a 1, como se muestra en la figura 5.10.



Funciones de energía
Figura 5.10

Las funciones tienen forma trapezoidal por el hecho de ser simplificaciones de funciones gaussianas. El motivo de basarse en funciones gaussianas es para aproximar a esta forma la distribución de probabilidad de las muestras.

Como se ve en la figura 5.10 un valor que se encuentre muy cerca de una media tendrá energía igual a cero. Y en cualquier punto la suma de las funciones de energía es igual a uno.

Una vez teniendo las funciones de energía se inicializa una solución $Y(0)$ con las mismas dimensiones de la imagen original y con valores aleatorios en el rango del número de clases en que se segmentará la imagen, que en nuestro caso es dos.

A partir de la solución aleatoria $Y(0)$ se ira generando, mediante el proceso Markoviano y el algoritmo de recocido simulado, la segmentación de la imagen como se muestra en la siguiente figura.

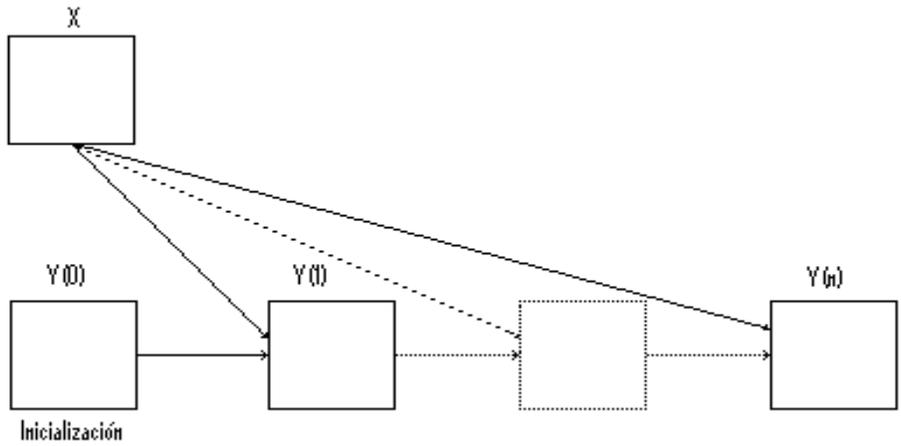


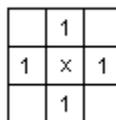
Figura 5.11

5.4 SEGMENTACIÓN BAJO DIFERENTES ESQUEMAS DE VECINDADES

Como se vio en el capítulo 3 existen diferentes sistemas de vecindades, pero son dos los principalmente utilizados; los de primer orden (4-vecindad) y los de segundo orden (8-vecindad).

5.4.1 Segmentación en 4-vecindad

En el sistema de vecinos de primer orden, también llamado sistema de 4-vecindad, cada site interior tiene 4 vecinos, como se muestra en la figura siguiente.



4-vecindad
Figura 5.12

En este tipo de segmentación las iteraciones se realizan de manera muy rápida, debido a que sólo se consideran 4 vecinos del site y no 8. Aunque podría pensarse que por no considerar toda la vecindad este tipo de segmentación arrojaría malos resultados, como se vera más adelante, este esquema de segmentación da muy buenos resultados en las primeras iteraciones del proceso de segmentación.

5.4.2 Segmentación en 8-vecindad

En el sistema de vecinos de segundo orden, también llamado sistema de 8-vecindad, cada site interior cuanta con 8 vecinos, como se muestra en la figura 5.13.

2	1	2
1	x	1
2	1	2

8-vecindad
Figura 5.13

Como se vera más adelante en las pruebas realizadas, este esquema de segmentación da mejores resultados en las iteraciones finales del proceso de segmentación, pues tiende a eliminar zonas heterogéneas.

5.5 SEGMENTACIÓN BAJO ESQUEMAS DE DECREMENTO DE TEMPERATURA

Como se mostró en el capítulo 4, la optimización del proceso de segmentación basada en el algoritmo *de recocido simulado* tiene como parámetro a la temperatura, la cual se va decrementando en cada iteración.

Los dos formas de decremento con las que trabajaremos en la presente tesis son; decremento geométrico y decremento logarítmico.

En ambos decrementos la temperatura inicial se considerara igual a 2, $T_0=2$.

5.5.1 Decremento Geométrico

El decremento geométrico es de la forma $T^{(iter+1)} = \alpha T^{(iter)}$, donde α la tomaremos con el valor de 0.9, $\alpha =0.9$. La curva de este decremento se observa en la figura 5.14

5.5.2 Decremento Logarítmico

El decremento logarítmico es de la forma $T^{(iter)} = \frac{T_0}{\ln(iter + 1)}$. La curva de este decremento se observa en la figura 5.15

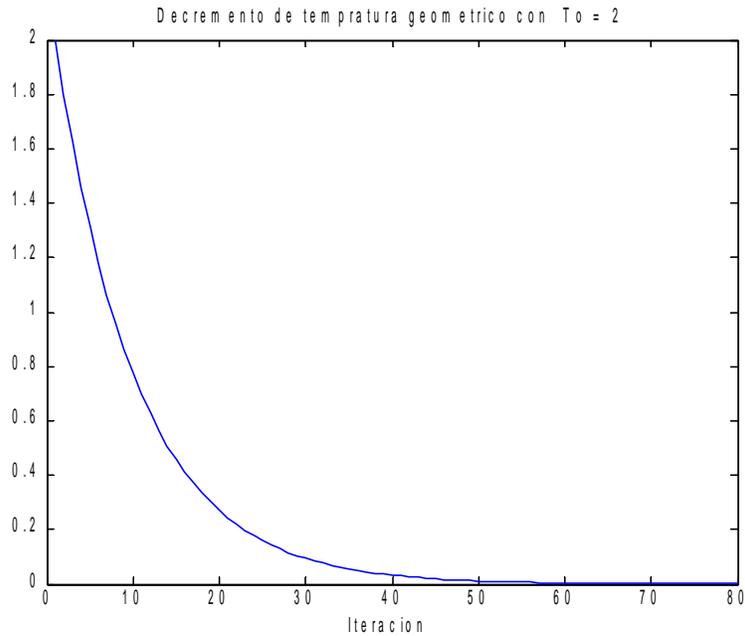


Figura 5.14

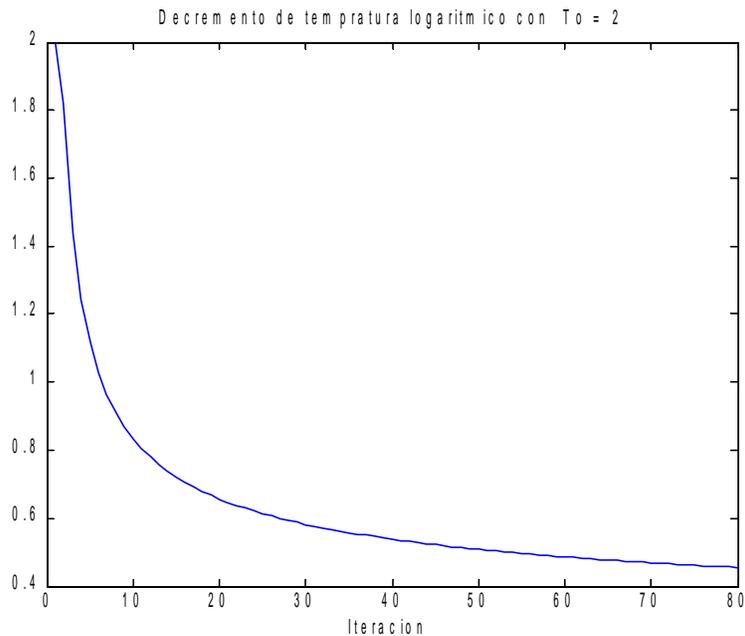


Figura 5.15

En las figuras de arriba se puede apreciar que el decremento logarítmico es más pronunciado o “acelerado” al principio de las iteraciones y conforme éstas avanzan el decremento se vuelve muy “lento”.

Por el contrario el decremento geométrico es menos vertical en las primeras iteraciones pero de la iteración 50 en adelante es prácticamente cero.

5.6 PROPUESTA DE MODELO VECINDAD-DECREMENTO DE TEMPERATURA

A continuación se muestran algunos ejemplos de segmentación aplicados a la figura 5.1, durante un proceso de 40 iteraciones, haciendo combinaciones entre distintos esquemas de vecindad y esquemas de decremento de temperatura.

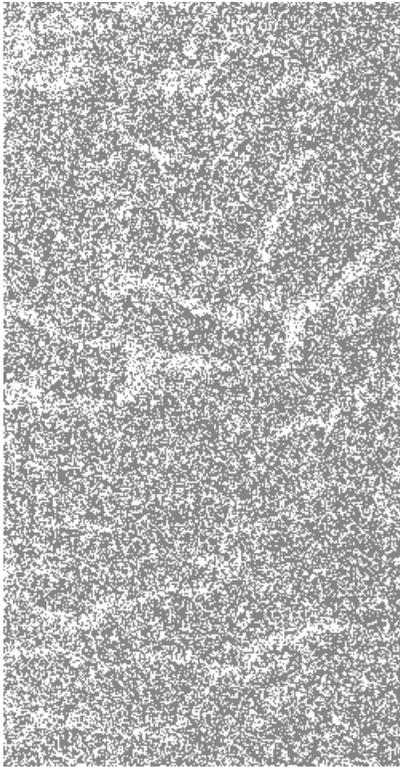
5.6.1 Segmentación 4-vecindad y 8-vecindad con decremento geométrico

a) Segmentación 4-vecindad decremento geométrico

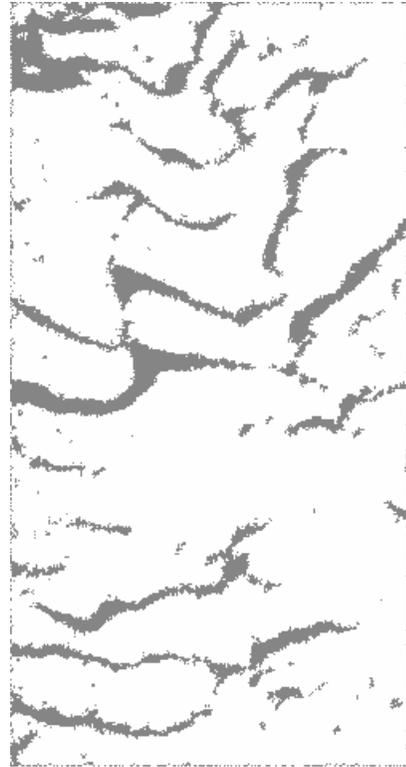
La imagen segmentada mostrada abajo se realizó mediante un esquema de segmentación de 4-vecindad y un decremento de temperatura geométrico (figura 5.14), el tiempo aproximado fue de 8.9 minutos (534 s).

La imagen de la izquierda muestra la solución generada hasta la cuarta iteración y la de la derecha, la solución generada después de 40 iteraciones. La gráfica 5.18 muestra el decremento de la energía de la imagen.

El decremento de energía representa el porcentaje de píxeles, sites, que cambian de valor respecto a la iteración pasada.



Resultado hasta la cuarta iteración
Figura 5.16



Después de 40 iteraciones
Figura 5.17

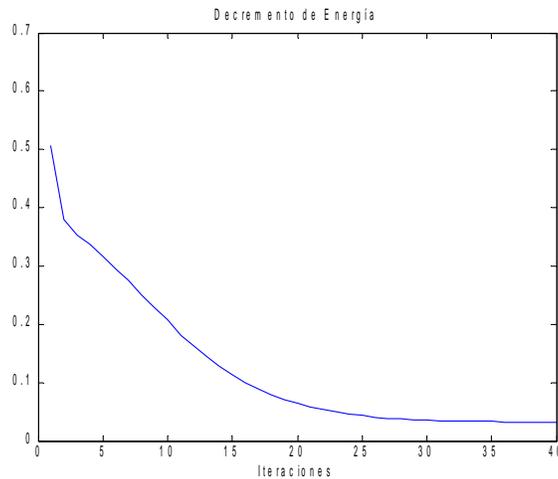
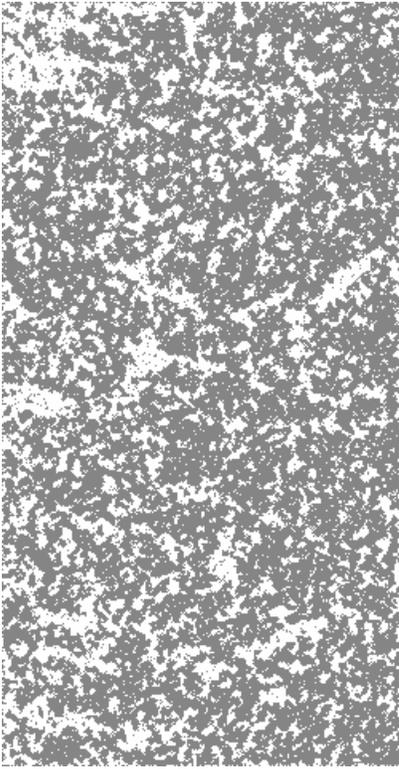


Figura 5.18

b) Segmentación 8-vecindad decremento geométrico

La imagen segmentada mostrada abajo se realizó mediante un esquema de segmentación de 8-vecindad y un decremento de temperatura geométrico, el tiempo aproximado fue de 10.9 minutos (654 s). La imagen de la izquierda muestra la solución general hasta la cuarta iteración y la de la derecha, la solución generada después de 40 iteraciones. La grafica 5.21 muestra el decremento de la energía de la imagen.



Resultado hasta la cuarta iteración
Figura 5.19



Después de 40 iteraciones
Figura 5.20

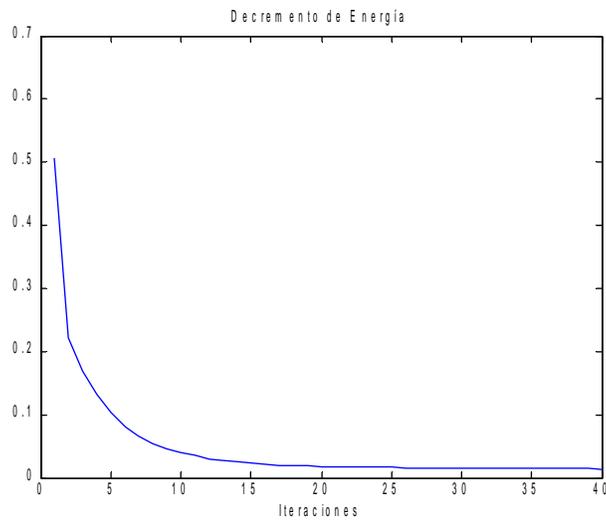


Figura 5.21

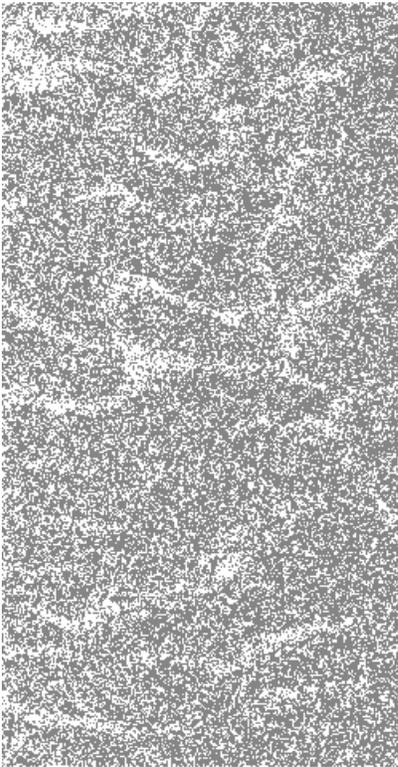
Se puede apreciar que la segmentación considerando el esquema de 8-vecindad es deficiente. Además se aprecia que la imagen está sobre segmentada, pues la gráfica del decremento de energía no disminuye de la iteración 25 en adelante.

5.6.2 Segmentación 4-vecindad y 8-vecindad con decremento logarítmico

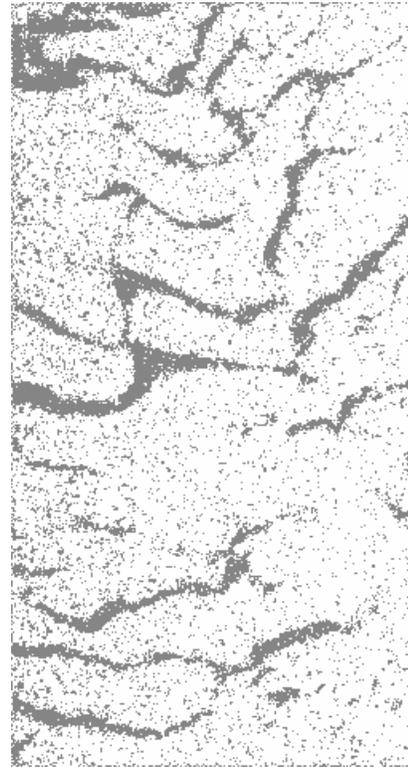
a) Segmentación 4-vecindad decremento logarítmico

La imagen segmentada mostrada abajo se realizó mediante un esquema de segmentación de 4-vecindad y un decremento de temperatura logarítmico (figura 5.15), el tiempo aproximado fue de 8.7 minutos (522 s).

La imagen de la izquierda muestra la solución generada hasta la cuarta iteración y la de la derecha, la solución generada después de 40 iteraciones. La grafica 5.24 muestra el decremento de la energía de la imagen.



Resultado hasta la cuarta iteración
Figura 5.22



Después de 40 iteraciones
Figura 5.23

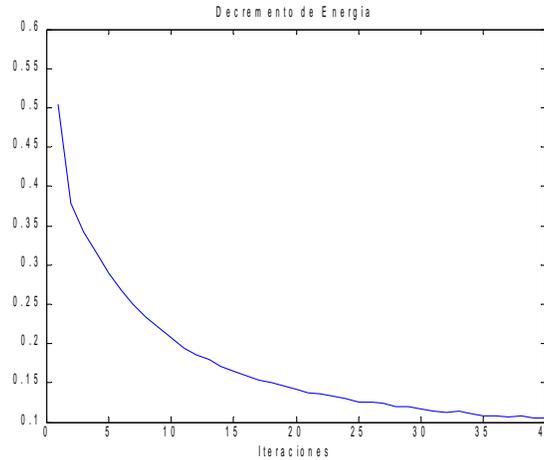
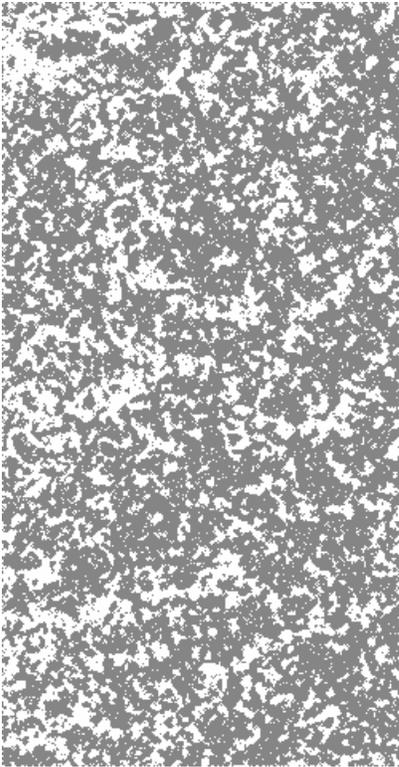


Figura 5.24

Después de 40 iteraciones la imagen aun contiene muchas regiones heterogéneas, gran cantidad de pequeñas secciones (puntos) oscuras, lo cual indica que la imagen aun no se ha segmentado, esto se puede ver también en la poca disminución que tiene la curva en la gráfica del decremento de energía. Observar que el decremento de energía mostrado en la figura 5.24 durante las 40 iteraciones nunca es menor de 0.1, mientras que esta situación se logra desde la iteración 20 de la figura 5.18, donde se tiene un decremento de temperatura geométrico.

b) Segmentación 8-vecindad decremento logarítmico

La imagen segmentada mostrada abajo se realizo mediante un esquema de segmentación de 8-vecindad y un decremento de temperatura logarítmico, el tiempo aproximado fue de 11 minutos. La imagen de la izquierda muestra la solución genera hasta al cuarta iteración y la de al derecha, la solución generada después de 40 iteraciones. La grafica 5.27 muestra el decremento de la energía de la imagen.



Resultado hasta la cuarta iteración
Figura 5.25



Después de 40 iteraciones
Figura 5.26

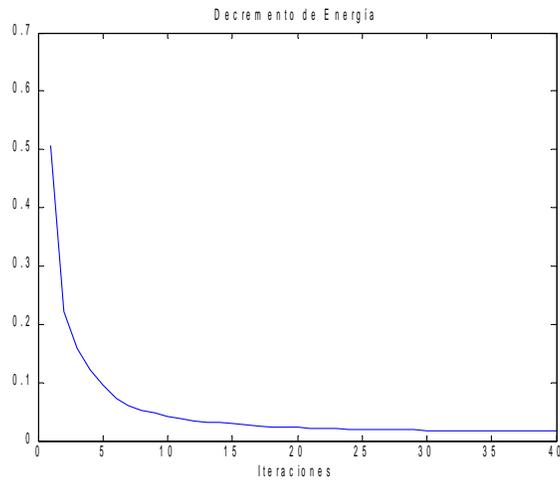


Figura 5.27

En la figura 5.26 se observa que nuevamente el esquema de segmentación en 8-vecindad da malos resultados, ahora con decremento de temperatura logarítmico.

5.6.3 Esquema Adaptivo

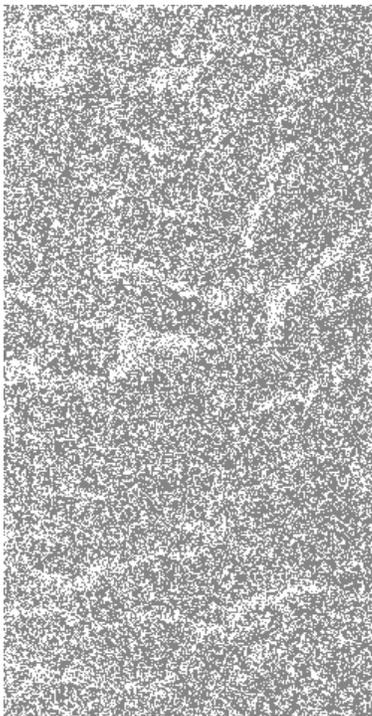
En las secciones anteriores se combinaron esquemas de vecindad y de decremento de temperatura, en la presente sección se intentara encontrar un modelo que utilice los esquemas anteriores y que sea capaz de cambiar entre estos esquemas al que se considere más apropiado, según sea el comportamiento del proceso de segmentación a lo largo de las iteraciones.

De los resultados de los algoritmos de la sección anterior (figuras 5.16 a 5.26) se puede deducir heurísticamente algunas características de los esquemas de segmentación en cuanto a la vecindad considerada para los cliques y, al decremento de temperatura considerado.

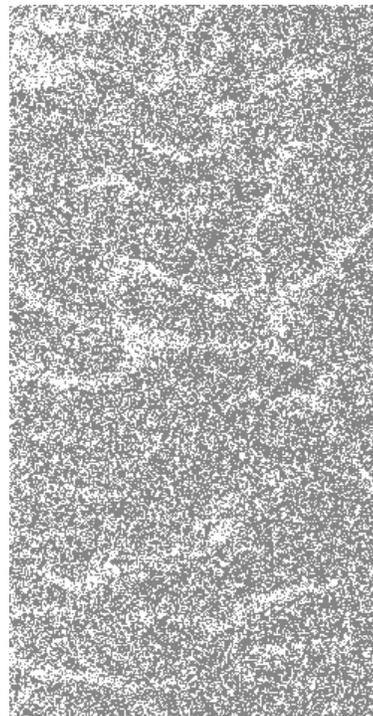
Respecto al decremento de temperatura

Primero se analizara la etapa inicial de la segmentación, resultado de las primeas 4 iteraciones. Después se vera los resultado finales, resultado obtenidos después de 40 iteraciones.

Observando las imágenes de las figuras 5.16 y 5.22 (decremento de temperatura geométrico y logarítmico respectivamente, que corresponden cada una al resultado obtenido después de 4 iteraciones en el esquema de 4-vecindad), a continuación vueltas a mostrar por comodidad, podemos ver que la correspondiente al decremento logarítmico (derecha) da mejores resultados.



Geométrico
Figura 5.16

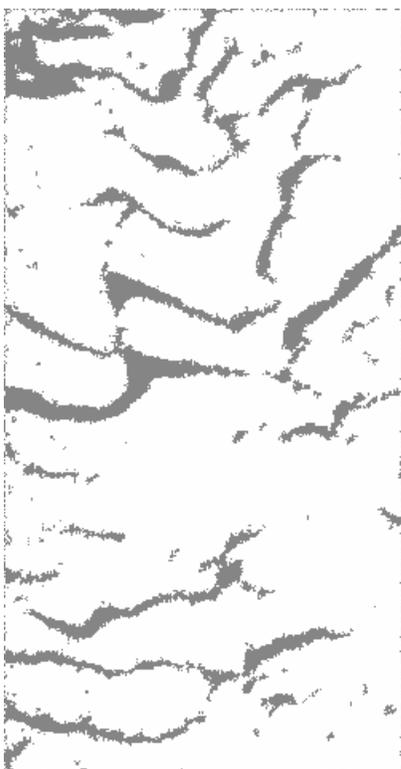


Logarítmico
Figura 5.22

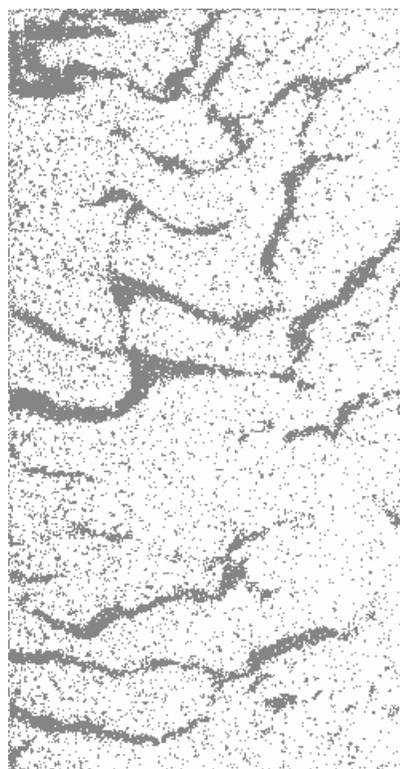
Ambas imágenes son muy parecidas pero observando con un poco de cuidado se puede apreciar que el resultado dado, en estas primeras iteraciones, por el modelo de decremento logarítmico está más definido, se pueden apreciar con un poco de mayor claridad los ríos que se muestran en la figura 5.1

Los resultados mostrados en las figuras 5.19 y 5.25 (de 8-vecindad) no dan muestra de una mejor segmentación de una sobre otra.

Viendo la etapa final de los esquemas de 4-vecindad, figuras 5.17 y 5.23 (decremento de temperatura geométrico y logarítmico respectivamente), a continuación vuelto a mostrar por comodidad, se puede apreciar que el la imagen resultado dada por el esquema de decremento geométrico (izquierda) es más homogénea, y por lo tanto mejor, que la dada por el de decremento logarítmico.



Geométrico
Figura 5.17



Logarítmico
Figura 5.23

Aunque el resultado final obtenido por el modelo de decremento geométrico es mejor que el dado por el de decremento logarítmico, aun ambas imágenes presentan zonas heterogéneas con una gran cantidad de puntos aislados.

Con las observaciones vistas respecto a los esquemas de decremento, de manera heurística se puede atrever a decir que en las primeras iteraciones del proceso de

segmentación, los resultados son mejores con decrementos del tipo logarítmico y que en las finales, por el contrario, los resultados son mejores con decrementos del tipo geométrico.

Superponiendo las graficas de decremento geométrico y logarítmico, figura 5.28, se puede ver que después de la iteración 8, la grafica de decremento geométrico (línea continua) decrece con mayor rapidez.

Se puede concluir, entonces, que entre más rápido decrezca la temperatura se obtienen mejores resultados.

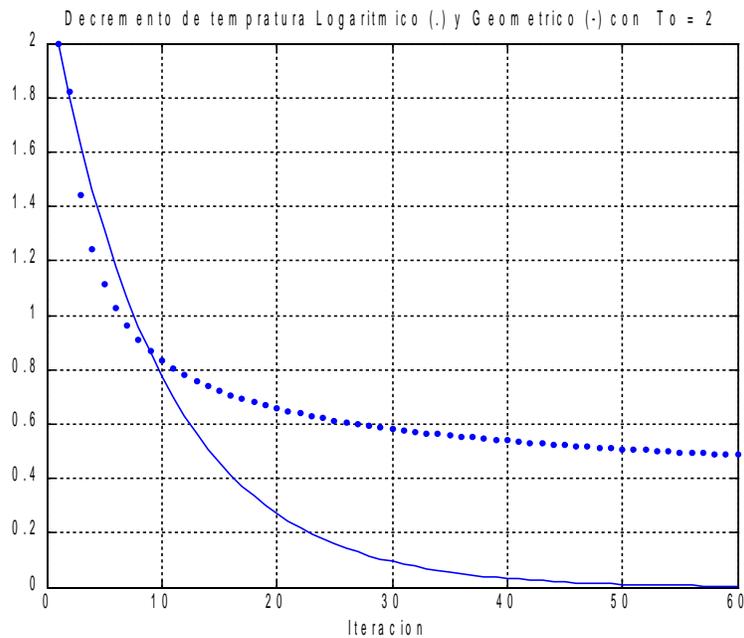


Figura 5.28

Entonces la propuesta en cuanto al decremento de energía es empezar con un decremento logarítmico y después cambiar a uno geométrico.

Este cambio de decremento se logra introduciendo una sentencia *if* dentro del loop de iteraciones:

```

if    Dec_Log < Dec_Geo
        Dec = Dec_Log
else
        Dec = Dec_Geo
end

```

La función de decremento utilizada para el modelo propuesto tiene la gráfica mostrada en la Figura 5.29.

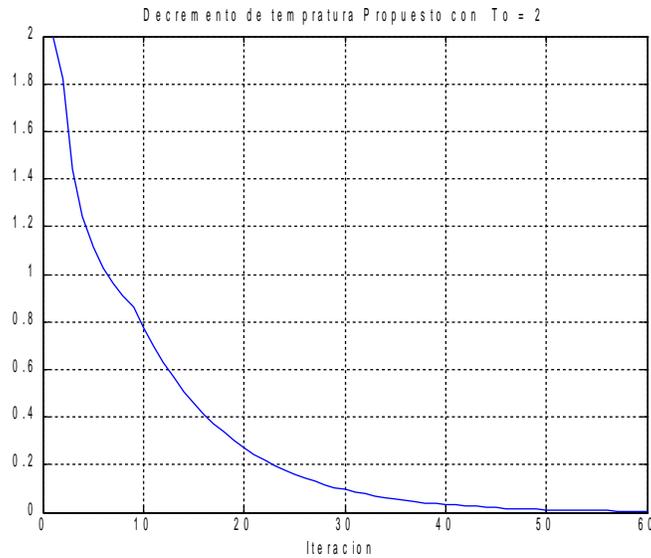
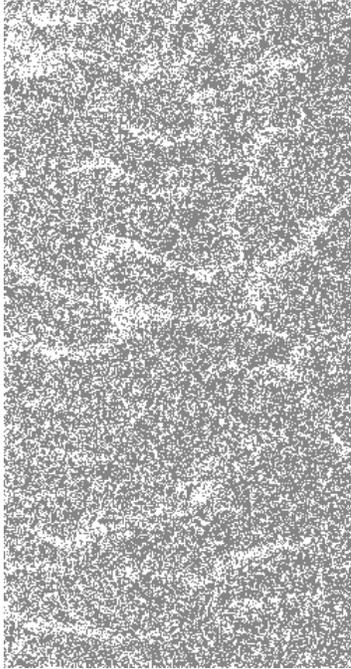


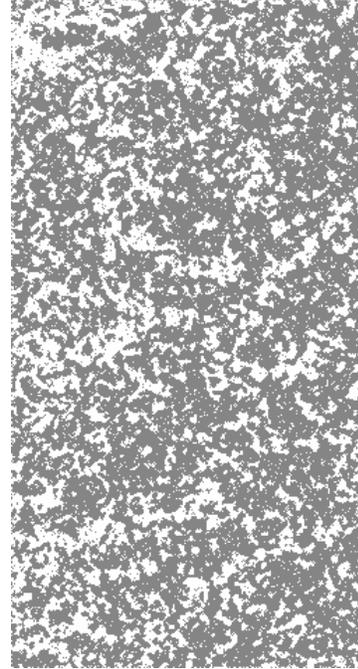
Figura 5.29

Respecto al sistema de vecindad

Mostrando nuevamente las figuras 5.22 y 5.25 (esquemas de 4-vecindad y 8-vecindad respectivamente) correspondientes a la cuarta iteración en esquema de decremento logarítmico, se puede ver que los mejores resultados se dan en el esquema de 4-vecindad (izquierda).



4- vecindad
Figura 5.22



8-vecindad
Figura 5.25

La figura 5.25 muestra que el iniciar el proceso de segmentación con un esquema de 8-vecindad da malos resultados. Y los problemas no son sólo al inicio, conforme se va avanzando en las iteraciones se va “acarreando” una mala segmentación que dista mucho de la imagen original. A continuación se ve la iteración final (40) bajo el esquema de 8-vecindad y decremento logarítmico, figura 5.26, y la imagen original, figura 5.1. Se puede apreciar que el resultado de la segmentación no concuerda con la original.



Resultado de esquema 8-vecindad
y decremento logarítmico
Figura 5.26

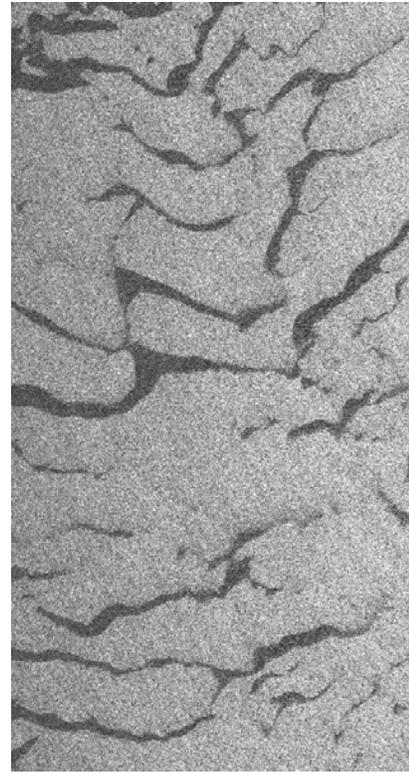


Imagen original

Figura 5.1

Sin embargo, se puede ver que la solución dada por el esquema de 8-vecindad está libre de puntos aislados, tiene regiones más homogéneas. Y es precisamente tener regiones homogéneas el objetivo de la segmentación.

Entonces de forma heurística se puede decir que en el proceso de segmentación es menester iniciar con un esquema de 4-vecindad y finalizar con uno de 8-vecindad. El esquema de 8-vecindad serviría únicamente para eliminar los puntos aislados y “suavizar” o “redondear” las regiones que ya se hallan definidas en las iteraciones de 4-vecindad.

El problema ahora reside en ¿Cuándo cambiar del esquema de 4-vecindad al de 8-vecindad?, ¿Después de cuántas iteraciones cambiar?

Mi propuesta está basada en el decremento de la energía de la imagen. Como se ve en la figura 5.18, la energía de la solución generada va disminuyendo de iteración en iteración.

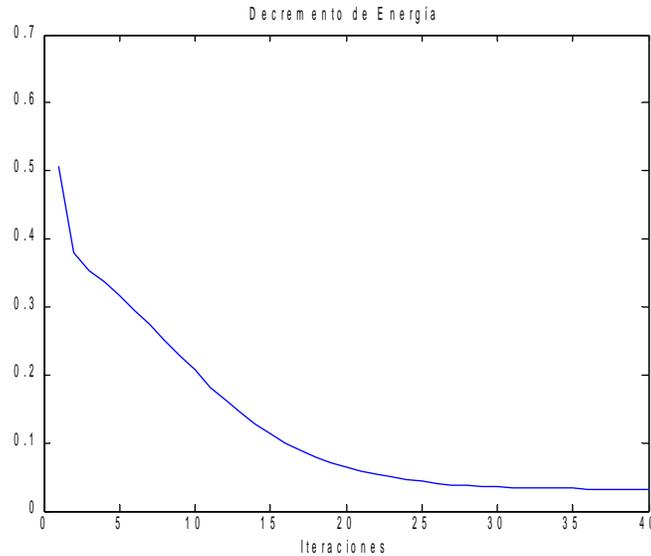


Figura 5.18

La energía aquí representa el cambio, en porcentaje, de la configuración actual de los valores de los cliques respecto a la configuración anterior. Conforme se va avanzando en el proceso de segmentación se van acabando las configuraciones posibles, la imagen se empieza a definir y el número de píxeles que cambian su valor de una iteración a otra disminuye.

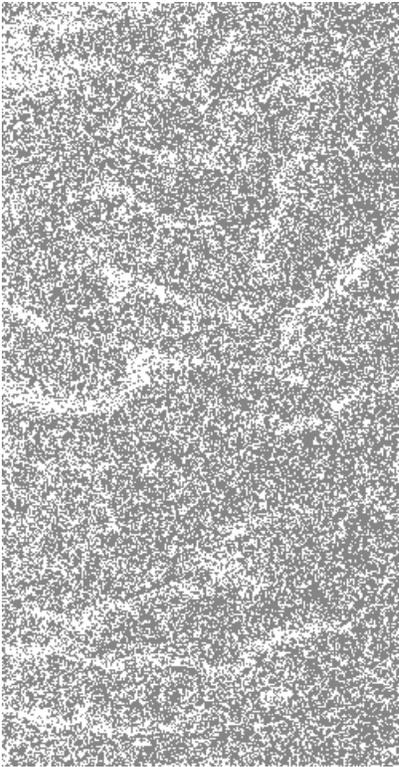
Si se observa la figura 5.18 se puede apreciar que después de la iteración 30 la curva deja de disminuir y se vuelve prácticamente constante. Se necesita establecer un criterio que determine a partir de que porcentaje de decremento de energía se empezará a utilizar el sistema de 8-vecindad.

Una vez considerado el sistema de 8-vecindad en el proceso se puede establecer otro criterio que determine a partir de que porcentaje de decremento de energía, terminar con el proceso de segmentación. Así no es necesario fijar de inicio un número de iteraciones, si no que el modelo determina cuando se ha segmentado la imagen bajo los criterios establecidos.

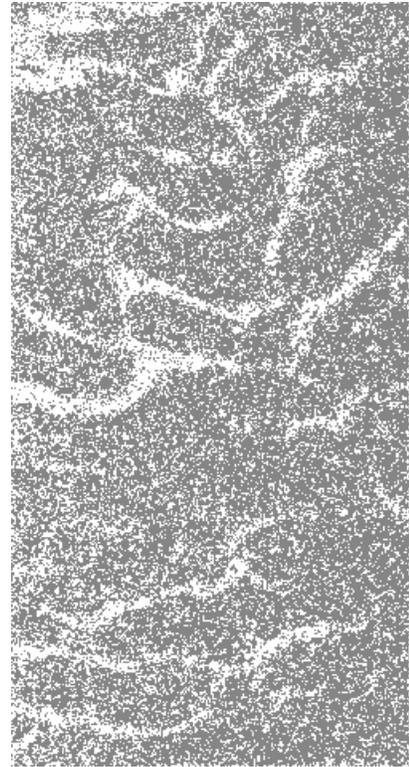
El modelo que se propone es tomar la pendiente (razón de cambio) entre el valor de la curva de energía actual y el valor de tres iteraciones atrás, y comparar esta pendiente con un valor umbral. Se escogió la tangente entre cada 4 puntos, porque, como se menciona al principio del capítulo, la imagen original fue dividida en cuatro subimágenes para acelerar el proceso de segmentación.

En la práctica se pudo ver que valores de umbrales 0.0005 dan buenos resultados.

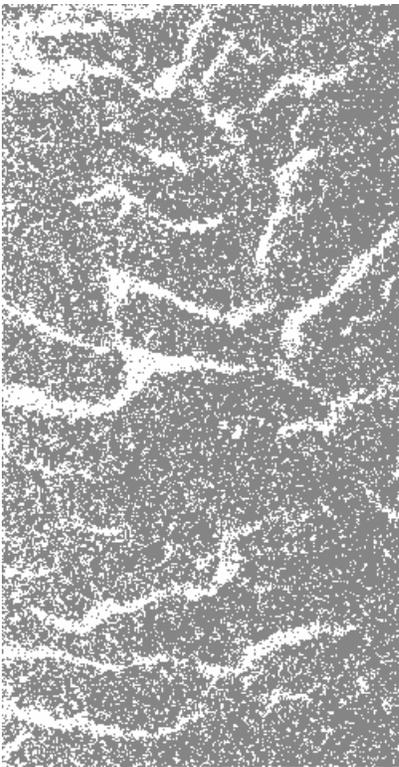
A continuación se muestra el proceso de segmentación (cada 4 iteraciones) propuesto con decremento de temperatura según la gráfica 5.29 y cambio de sistema de 4-vecindad a 8-vecindad según el criterio de la pendiente entre cada 4 puntos con umbral establecido de 0.0005.



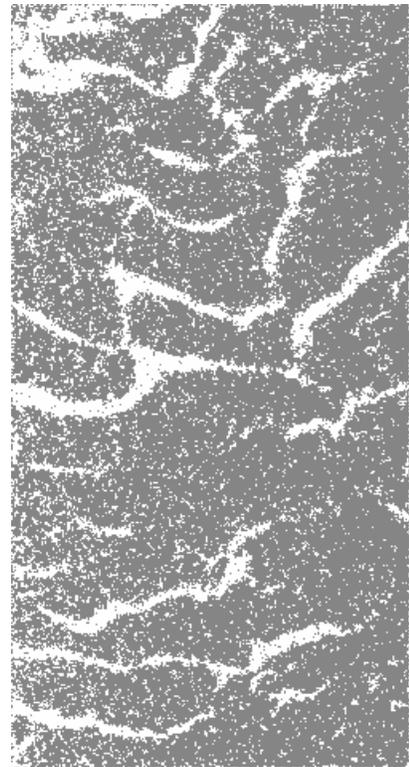
Iteración 4
Figura 5.30



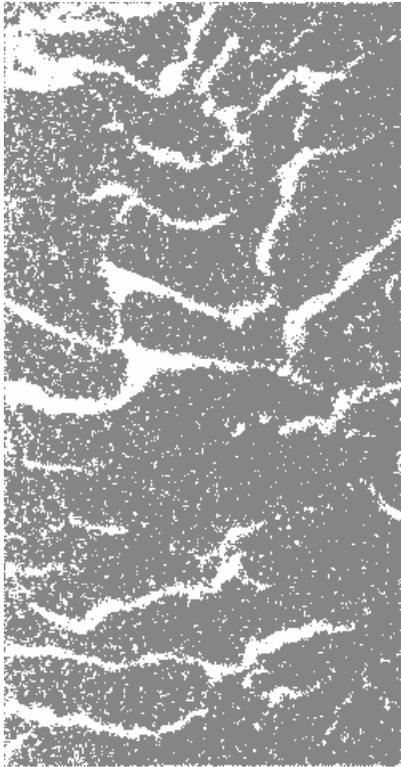
Iteración 8
Figura 5.31



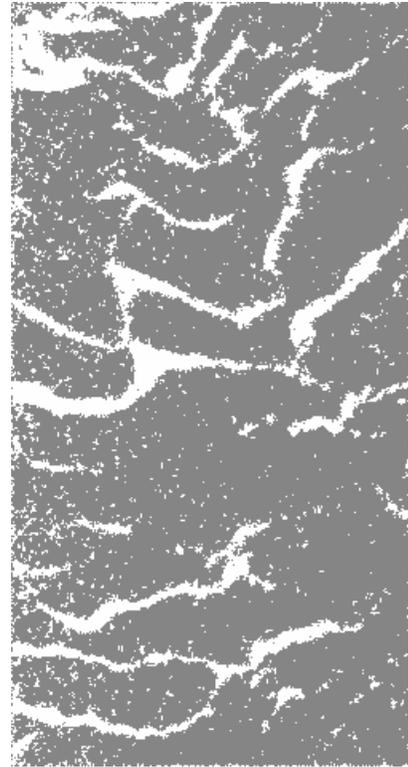
Iteración 12
Figura 5.32



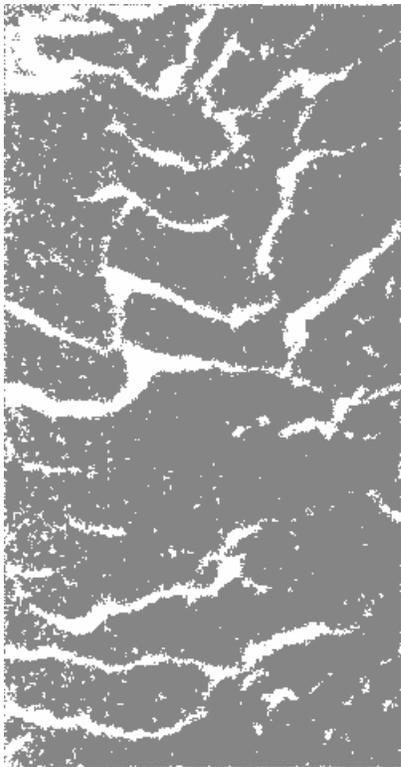
Iteración 16
Figura 5.33



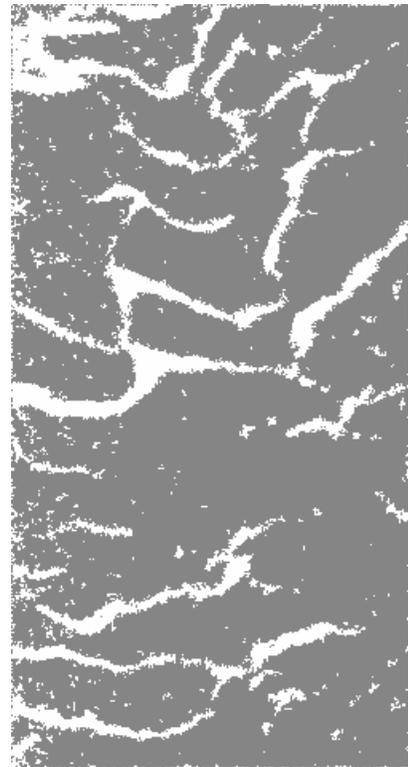
Iteración 20
Figura 5.34



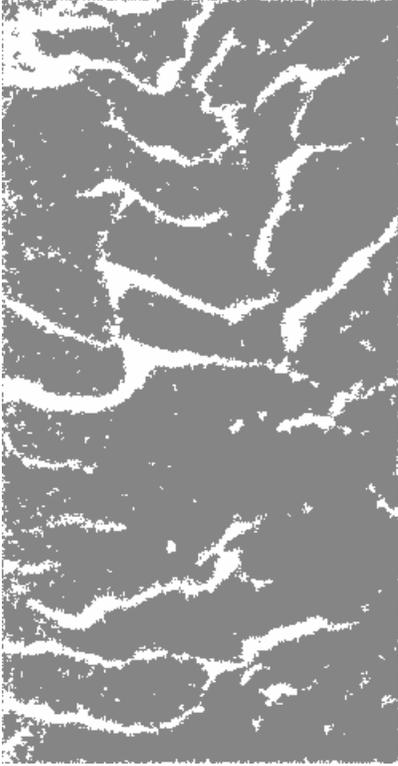
Iteración 24
Figura 5.35



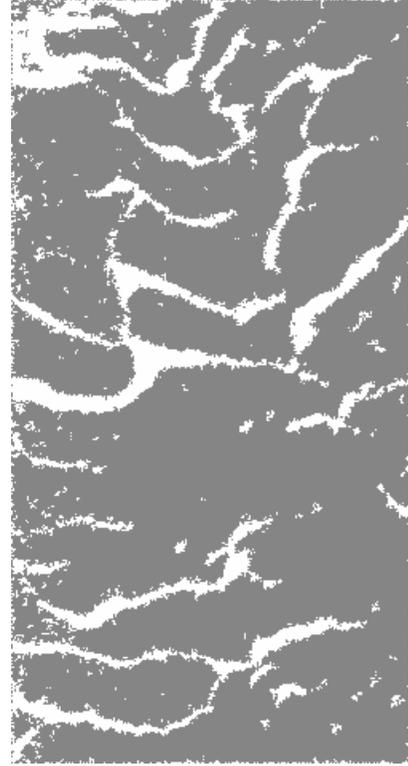
Iteración 28
Figura 5.36



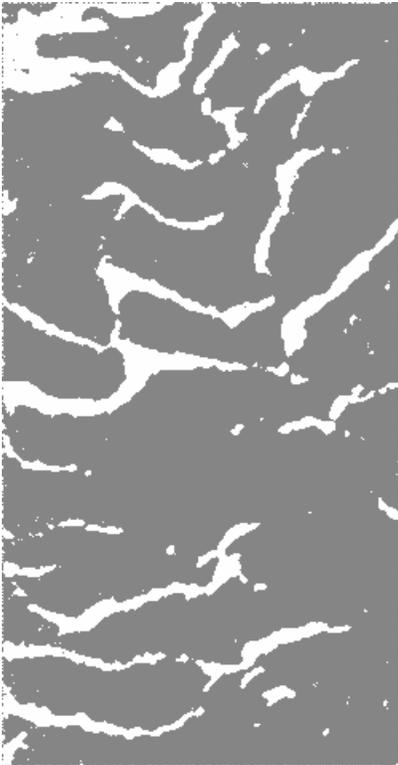
Iteración 32
Figura 5.37



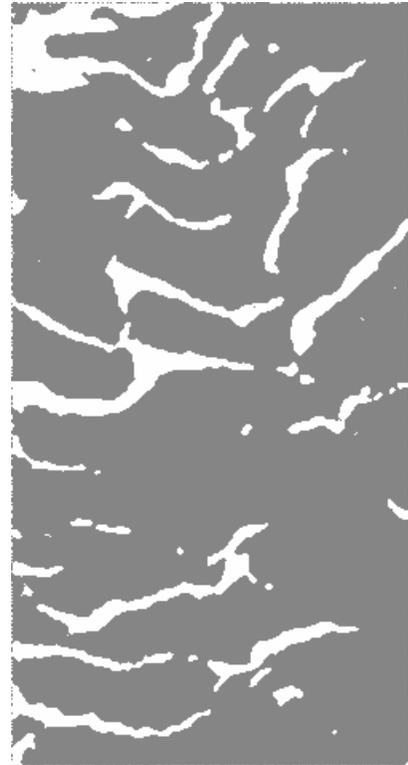
Iteración 36
Figura 5.38



Iteración 40
Figura 5.39



Iteración 44
Figura 5.40



Iteración 48
Figura 5.41



Iteración 50
Figura 5.42

5.7 EVALUACIÓN DE RESULTADOS

El proceso tuvo una duración aproximada de 12.1 minutos (726 s). De la figura 5.30 a 5.31 fueron las primeras 8 iteraciones con un decremento de temperatura logarítmico, en adelante el decremento fue geométrico. De la figura 5.30 a la 5.39 fueron las primeras 40 iteraciones con un sistema de 4-vecindad, en estas iteraciones es donde se lleva la mayor parte del tiempo.

De la figura 5.40 a 5.42 fueron las últimas 12 iteraciones (41-52) con un sistema de 8-vecindad, se ve claramente en estas imágenes como las regiones se “suavizan” y se eliminan puntos aislados. El modelo decidió 52 iteraciones para segmentar la imagen.

La figura 5.42 muestra el resultado final después de las 52 iteraciones que el modelo decidió para segmentar la imagen.

La grafica del decremento de temperatura (figura 5.43) muestra como hasta la iteración 40 la curva tiende a ser constante y partir de la iteración 41 hay un cambio en la curva, vuelve a disminuir un poco, debido al cambio de 4-vecindad a 8-vecindad.

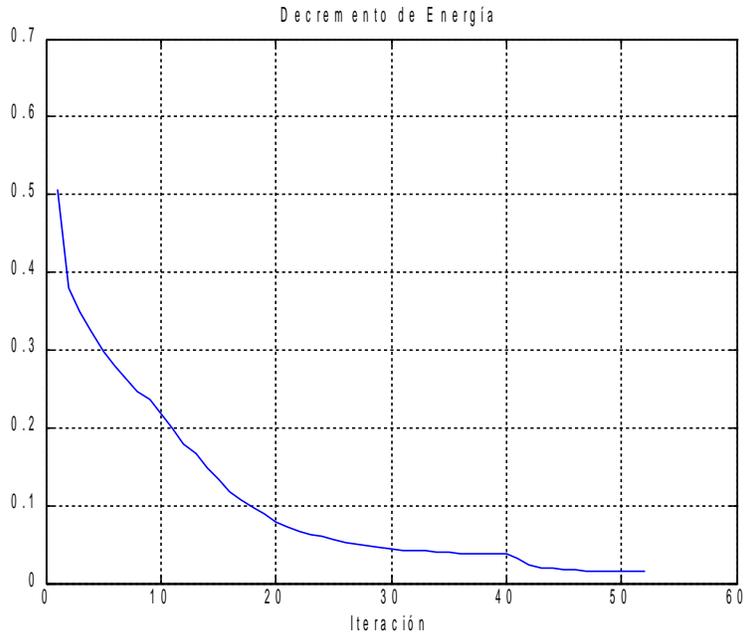


Figura 5.43

La deriva de la grafica anterior se muestra en la figura 5.44. Se puede apreciar que hasta la iteración 40 se alcanza el umbral de 0.0005, depuse se aprecia un cambio abrupto que representa la consideración del sistema de 8-vecindad. Nuevamente en la iteración 52 se vuelve a alcanzar el umbral y el proceso termina.

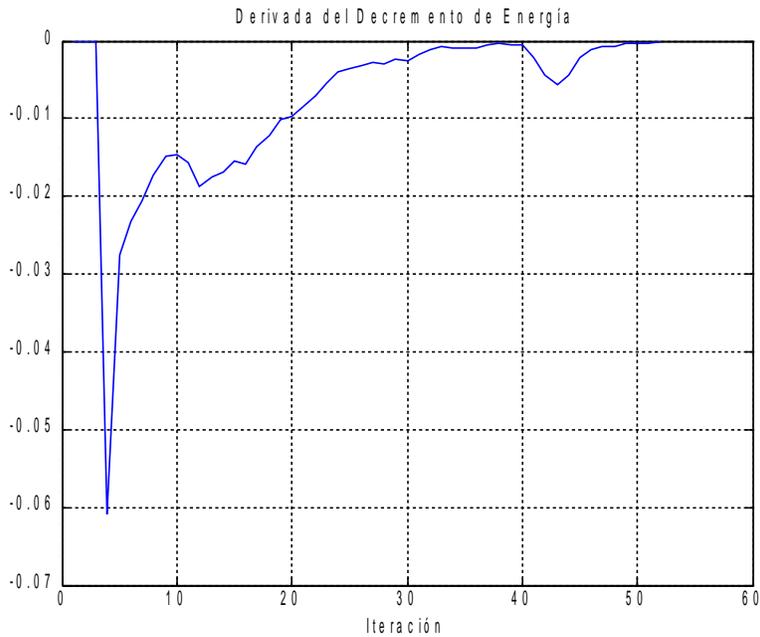
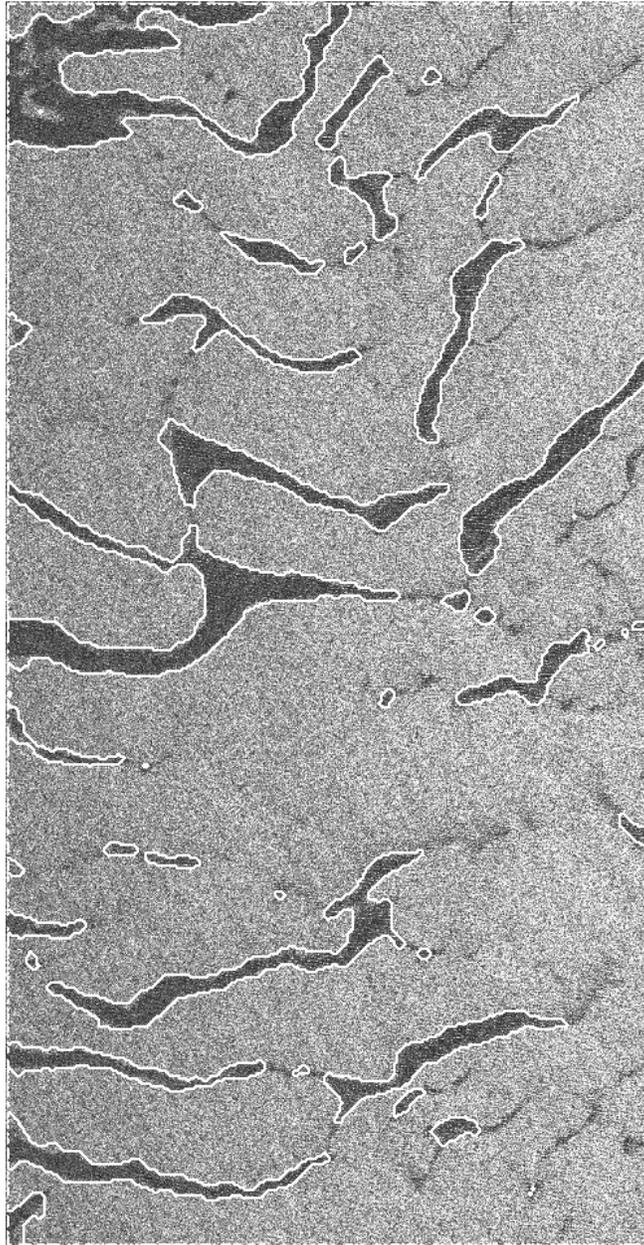


Figura 5.44

La siguiente figura muestra a la imagen original con los bordes, resultados del proceso de segmentación, superpuestos.



Resultado Final
Figura 5.45

En la imagen se puede ver que el resultado de la segmentación no es 100% perfecto, no obstante, los bordes señalan la mayor cantidad de regiones compuestas por agua en la imagen. Se debe tener en cuenta también que esta es una imagen “difícil” en el sentido de que existe mucha distorsión. Si se observa con cuidado la imagen original, se puede apreciar que la parte izquierda tiene tonos más oscuros que la derecha, lo que significa que las regiones de hielo no son del todo homogéneas.

Comparación de resultados entre el modelo propuesto y el modelo clásico

Imagen satelital de petróleo (parte oscura) en mar de 522x522 [04].

La siguiente imagen muestra la imagen original con la segmentación sobrepuesta obtenida con el modelo propuesto.

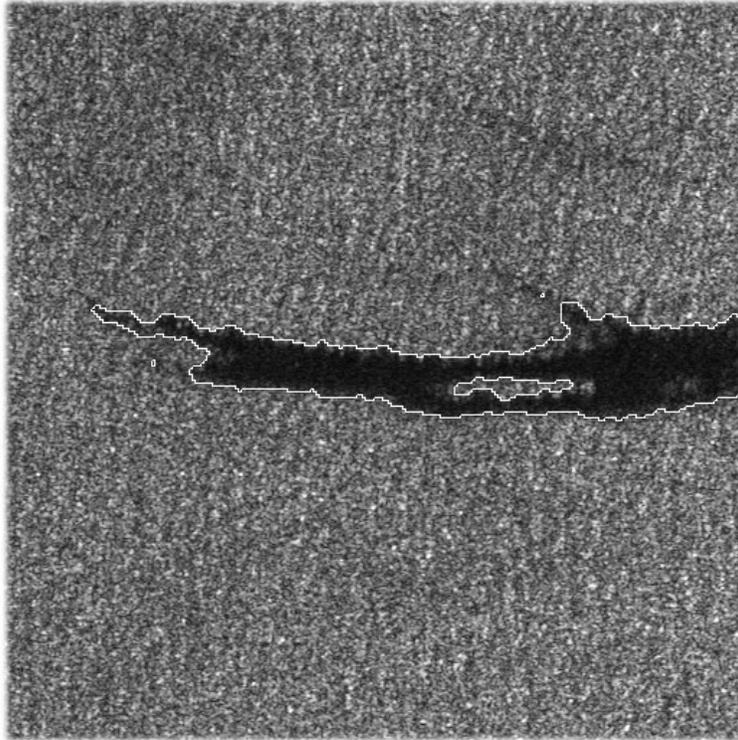


Figura 5.46

Segmentación sobrepuesta en la imagen original bajo el modelo propuesto.

El proceso determino 36 iteraciones y duro aproximadamente 290 segundos.

La siguiente imagen muestra el resultado obtenido al segmentar la misma imagen considerando únicamente una vecindad de 4-conexidad y un decremento de temperatura geométrico.

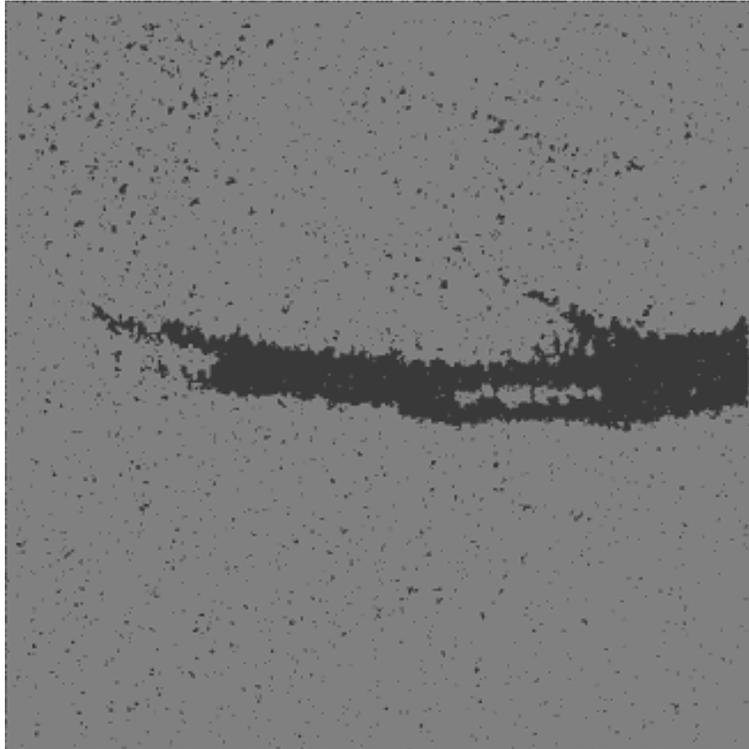


Figura 5.47
Resultado obtenido bajo el esquema clásico
(vecindades de 4-conexidad y decremento de temperatura geométrico)

Se estableció manualmente 36 iteraciones para tener un comparativo con el resultado anterior y el proceso duro aproximadamente 990 segundos.

Imagen satelital de petróleo (parte oscura) en mar de 522x522 [04].

La siguiente imagen muestra la imagen original con la segmentación sobrepuesta obtenida con el modelo propuesto.



Figura 5.48
Segmentación sobrepuesta en la imagen original bajo el modelo propuesto.

El proceso determino 40 iteraciones y duro aproximadamente 318 segundos.

La siguiente imagen muestra el resultado obtenido al segmentar la misma imagen considerando únicamente una vecindad de 4-conexidad y un decremento de temperatura geométrico.

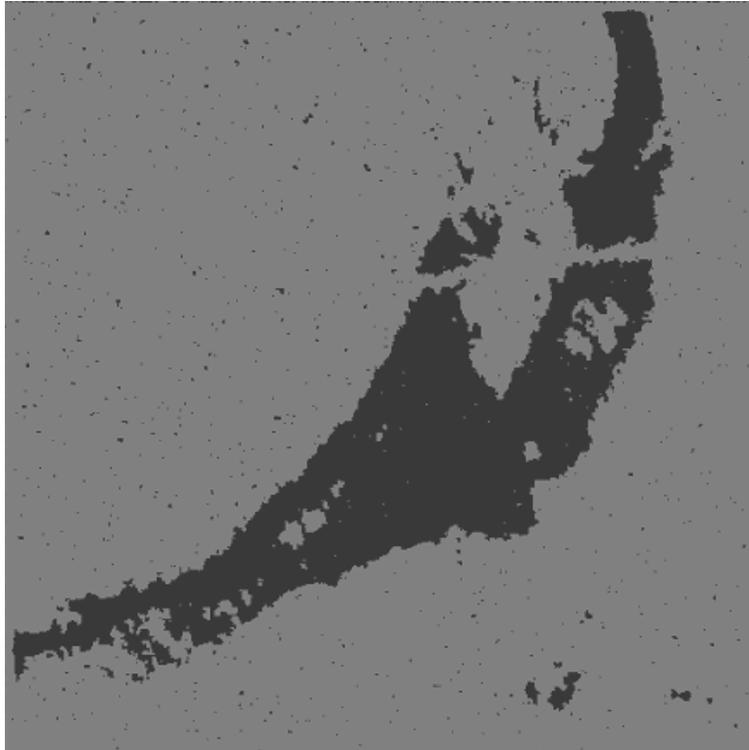


Figura 5.49
Resultado obtenido bajo el esquema clásico
(vecindades de 4-conexidad y decremento de temperatura geométrico)

Se estableció manualmente 40 iteraciones para tener un comparativo con el resultado anterior y el proceso duro aproximadamente 1129 segundos.

Imagen satelital de petróleo (parte oscura) en mar de 522x522 [04].

La siguiente imagen muestra la imagen original con la segmentación sobrepuesta obtenida con el modelo propuesto.

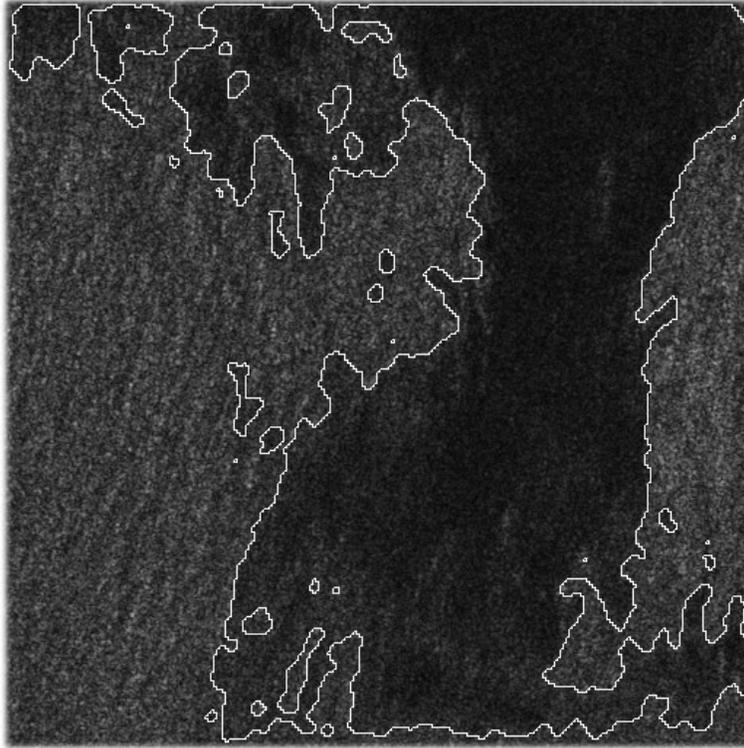


Figura 5.50
Segmentación sobrepuesta en la imagen original bajo el modelo propuesto.

El proceso determino 76 iteraciones y duro aproximadamente 594 segundos.

La siguiente imagen muestra el resultado obtenido al segmentar la misma imagen considerando únicamente una vecindad de 4-conexidad y un decremento de temperatura geométrico.

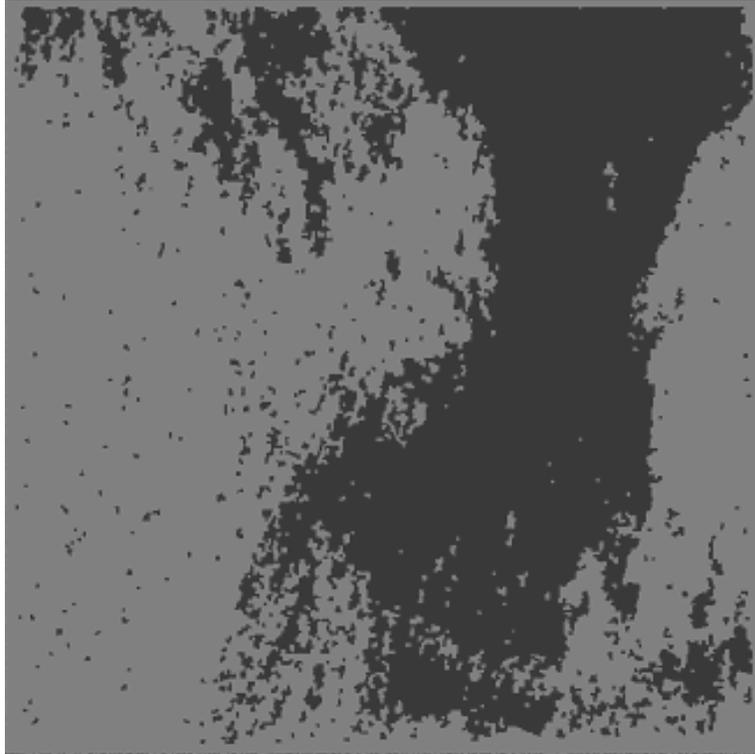


Figura 5.51
Resultado obtenido bajo el esquema clásico
(vecindades de 4-conexidad y decremento de temperatura geométrico)

Nuevamente, como en los ejemplos anteriores, se establecieron manualmente las iteraciones para tener un comparativo con el modelo propuesto, siendo 76 iteraciones y durando el proceso 2100 segundos.

Resumen de resultados

Imagen satelital	Dimensiones	Iteraciones	Tiempo (s) (modelo propuesto)	Tiempo (s) (modelo clásico)
Figura 5.45	500x960	52	726	-
Figura 5.46	522x522	36	290	990
Figura 5.48	522x522	40	318	1129
Figura 5.50	522x522	76	594	2100

Las imágenes anteriores nos muestran que el modelo adaptivo propuesto de vecindad da mejor resultado que el de una vecindad estática, en este caso de 4-conexidad. Y la tabla del resumen de resultados también nos muestra una reducción de los tiempos en el modelo propuesto.

6 CONCLUSIONES

Los resultados que se obtuvieron en la presente tesis muestran la utilidad de la segmentación basada en la teoría de Campos Aleatorios de Markov utilizando este esquema adoptivo de selección de vecindad y decremento de temperatura, pues con esta técnica las imágenes se pueden analizar, y así obtener información necesaria para la interpretación.

Para el estudio de la segmentación de la imagen primero se analizaron los esquemas de vecindad y de decremento de temperatura por separado. Observando la característica de los resultados dados por estos dos esquemas se propuso un modelo que a lo largo de la segmentación cambiara de un esquema a otro tanto en el *sistema de vecindad* como en el de *decremento de la temperatura*, dando como resultado mejores imágenes segmentadas.

En cuanto a los esquemas de decremento de temperatura fueron dos los analizados: el decremento tipo geométrico y el de tipo logarítmico. Se observó que los decrementos de temperatura geométricos dan buenos resultados después de las primeras iteraciones del proceso. Por el contrario, los decrementos de temperatura logarítmicos dan sus mejores resultados sólo al principio del proceso, pero si se mantiene este tipo de decremento durante el proceso se tendrá como resultado una imagen sub-segmentada (entiéndase por “imagen sub-segmentada” a una que después del proceso de segmentación presenta regiones heterogéneas).

Se puede decir que en las primeras iteraciones del proceso de segmentación, los resultados son mejores con decrementos del tipo logarítmico y que en las finales, por el contrario, los resultados son mejores con decrementos del tipo geométrico. Esto se apreció mejor al ver las graficas de estas dos funciones, la grafica logarítmica decrece con mayor rapidez durante las primeras iteraciones, después del punto de intersección entre estas dos graficas, la grafica geométrica es ahora la que decrece con mayor rapidez.

El criterio tomado para cambiar del decremento de temperatura logarítmico al decremento geométrico, en el modelo propuesto, fue el de tomar este último cuando su valor de temperatura fuera menor que la del otro, que ocurre después de la iteración 8. Sabiendo de antemano esto se puede crear un arreglo que contenga los valores de la temperatura que se utilizara.

En cuanto a los esquemas de vecindad asociados al píxel en estudio se restringió a dos casos; 8-conexidad y 4-conexidad. El primer esquema da como resultado regiones homogéneas, pero es inconveniente en las primeras iteraciones del

proceso debido a no conservar estructuras finas, lo cual conlleva a arrastrar los errores del inicio hasta el final. El segundo esquema proporciona detalles de la imagen pero las regiones que genera son heterogéneas.

Por tal motivo, el modelo de segmentación propuesto inicia bajo un esquema de sistema de vecindad de 4-conexidad, para incorporar detalles, y termina con un sistema de vecindad de 8-conexidad, para eliminar las regiones heterogéneas.

El criterio tomado para cambiar del esquema de 4-vecindad a 8-vecindad, en el modelo propuesto, fue el de tomar la pendiente (razón de cambio) entre el valor del cambio de la energía actual de la imagen generada hasta la iteración en curso y el valor del cambio de la energía de la imagen generada tres iteraciones atrás, y comparar esta pendiente con un valor umbral.

Este mismo criterio se utiliza para terminar el proceso. Una vez que se ha cambiado al esquema de 8-vecindad se vuelve a tomar la pendiente en cada iteración entre el valor del cambio de la energía actual de la imagen generada hasta la iteración en curso y el valor del cambio de la energía de la imagen generada tres iteraciones atrás. En caso de que el cambio de energía de la imagen actual respecto a sus anteriores tienda a ser constante, el algoritmo terminara el proceso para que no se genere una sobre-segmentación (entiéndase por “sobre-segmentación” a una imagen que después del proceso de segmentación pierde los detalles que poseía originalmente).

Otro punto importante para acelerar la convergencia del esquema propuesto, fue dividir la imagen a segmentar en cuatro subimágenes. Distribuyendo la información y analizando en cada iteración una subimagen.

7 REFERENCIAS

[LI01]Li Z. S., *Markov Random Field Modeling in Image Analysis*, Ed. Springer Verlag, 2001.

[DO01]Downsland K., Adenso-Díaz B., *Diseño de Heurísticas y Fundamentos de Recocido Simulado*, Revista Iberoamericana de Inteligencia Artificial, No. 20, pág. 34-52, 2001.

[GO03]Gonzalez R. C., Woods R. E., *Digital Image Processing*, Ed. Edisson-Wesley, pág. 610-621, Sep 2003.

[HA97]Hall D., Llinas J., *An Introduction to Multisensor Data Fusion*, Proceedings of the IEEE, Enero 1997, Vol. 85, No. 1.

[PU00]Pujol M., Rizo P., *Aplicación de los Modelos de Campos Aleatorios de Markov en Visión Artificial*, Revista Electrónica de Visión, Departamento de Ciencia de la Computación e Inteligencia Artificial, Número 4, Universidad de Alicante, Octubre 2000.

[MA94]Maravall Gómez-Allende, *Reconocimiento de Formas y Visión Artificial*, Ed. Addison-Wesley Iberoamericana, 1era. edición, 1994.

[01] Haralick R.M., Shanmugam K., Dinstein I., *Texture Features for Image Classification*, IEEE Trans. on Syst. Man Cybern., vol. 3.

[02] http://www.lfcia.org/~cipenedo/cursos/lp/Tema6/nodo6_3.html

[03] The Watershed Transformation - <http://cmm.ensmp.fr/~beucher/wtshed.html>

[04] Imágenes, cortesía de Flavio Parmiggiani, Institute of Atmospheric Sciences and Climate (ISAC) of the Italian National Research Council (CNR).