

03063



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**“DISEÑO DE UN ALGORITMO RECOCIDO
SIMULADO PARA LA SEGMENTACIÓN DE
IMÁGENES MULTIESPECTRALES”**

T E S I S

QUE PARA OBTENER EL GRADO DE:

**MAESTRO EN CIENCIAS
(COMPUTACIÓN)**

P R E S E N T A:

EDGAR RICARDO LÓPEZ GALVÁN

DIRECTOR DE TESIS: DR. BORIS ESCALANTE RAMÍREZ

México, D.F. junio de 2005.

m. 345975



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos

A LA UNAM Y AL POSGRADO

Por que gran parte de lo que soy es gracias a ellos.

A MI ESPOSA Y FUTURO HIJO

Por todo su apoyo, comprensión y cariño.

AL DR. BORIS

Por su apoyo y dirección para la realización de de este trabajo

A MIS PADRES Y HERMANOS

Por su tiempo y paciencia, y por que siempre están para soportarme.

A TODOS MIS AMIGOS Y COMPAÑEROS DE POSGRADO

Por su apoyo, solidaridad y por sobre todo su amistad.

A MIS COMPAÑEROS DE TRABAJO

Por su apoyo

Autoriza a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE:

Edgar R. Lopez

FECHA:

17 / Junio / 2003

FIRMA:

[Firma manuscrita]

Resumen

El presente trabajo se desarrolló con la finalidad de proponer un algoritmo que permita realizar la segmentación de una imagen multiespectral en un número de k regiones homogéneas, tomando en cuenta la información espacial y espectral que proporcionan los píxeles y su contexto.

Para ello se empleó la Teoría de los Campos Aleatorios de Markov, misma que se ha venido empleando en diferentes áreas de visión y del procesamiento de imágenes desde los años 80's y actualmente sigue siendo tema de estudio e investigación.

Se utiliza el algoritmo "Recocido Simulado", para atacar el problema de optimización combinatoria que se presenta en el Modelo de Segmentación propuesto, y se realiza un análisis sobre las consideraciones que diferentes autores hacen sobre el algoritmo, enfocándonos en los parámetros geométricos (Temperatura inicial, β , ley de evolución, criterio de equilibrio y criterio de congelación).

La aportación de este trabajo se centra en las consideraciones que se hacen en el algoritmo "Recocido Simulado" para su aplicación en el proceso de segmentación de imágenes multiespectrales.

Para evaluar y comparar los resultados se empleó el algoritmo de k -means y diferentes imágenes sintéticas, a través de la matriz de confusión. Una vez evaluado el algoritmo se validó utilizando una imagen Landsat 7.

Índice general

1. Introducción	1
1.1. Motivación	2
1.2. Planteamiento del problema	4
1.3. Organización de capítulos	4
2. Estado del Arte	6
2.1. Conocimiento base	6
2.1.1. Imagen digital	7
2.1.2. Imágenes multiespectrales	8
2.2. Segmentación	9
2.2.1. Técnicas orientadas a pixel.	10
2.2.2. Técnicas orientadas al contexto.	11
2.2.3. Técnicas orientadas al contorno.	12
3. Campos Aleatorios de Markov	13
3.1. Problema de etiquetado	13
3.1.1. Problemas de etiquetado en visión	14
3.1.2. Clasificación del problema del etiquetado en visión. . .	15
3.1.3. Etiquetado con restricciones de contexto	16
3.2. Campos Aleatorios de Markov	17
3.2.1. Sistema de vecinos y clique	17
3.2.2. Campos Aleatorios de Markov	19

Capítulo 1

Introducción

Desde la antigüedad el hombre ha buscado clasificar a los seres y objetos de su alrededor, y de esta forma ha logrado grandes avances en el entendimiento y estudio de la naturaleza y del mundo en general. Por ejemplo, los seres vivos se dividieron en reinos, géneros, familias, especies, etc., de acuerdo a ciertas características que los hacen similares y otras que los distinguen.

Con los avances de la informática y la tecnología, se ha despertado la inquietud por lograr automatizar el proceso de clasificación. La rama de la informática conocida como procesamiento de imágenes digitales ha tenido mucho que ver al respecto. La medicina, astronomía, geografía, etc., son áreas que en la actualidad requieren de manera importante del procesamiento de imágenes digitales. Imágenes de tomografía computarizada, resonancia magnética, ultrasónicas, satelitales y de radar son empleadas para apoyo en dichas áreas de estudio.

La segmentación de imágenes tiene su origen en estudios psicológicos que indican la preferencia de los humanos por agrupar regiones visuales en regiones homogéneas, en términos de proximidad, similitud y continuidad. La

homogeneidad puede estar basada en propiedades tales como textura, color y distribución de las densidades de los objetos de la imagen.

La segmentación es un paso imprescindible en diversos procesos del tratamiento de imágenes, algunas aplicaciones son:

- Determinación de superficies y volúmenes
- Reconstrucción tridimensional
- Diagnóstico automático (detección de defectos en materiales, tejidos anómalos, células, etc.)
- Reducción de información en transmisión, almacenamiento.

1.1. Motivación

En el campo del procesamiento digital de imágenes, la segmentación es uno de los procesos más importantes. Actualmente para apoyo en el diagnóstico en medicina [13], imágenes de resonancia magnética, tomografía computarizada y mamografía digital, juegan un papel crucial, automatizando o facilitando la delimitación de estructuras anatómicas y otras regiones de interés por mencionar algunas aplicaciones.

En la percepción remota, además de sus evidentes aplicaciones en cuestiones militares y estratégicas, la segmentación de imágenes posibilita la detección de fallas geológicas, el análisis de composición del suelo, la evaluación de recursos naturales, seguimiento de la evolución de las cosechas, permite conocer la distribución de cultivos, o identificar daños como plagas o sequías, prevención y seguimiento de incendios forestales, etc. [15].

El mayor esfuerzo dedicado al análisis multiespectral ha sido en esta última disciplina “la percepción remota” definida como “la adquisición de in-

formación sobre un objeto sin tener contacto físico con él". Esta tiene sus orígenes desde mediados del siglo XIX, cuando en 1859 se tomaron las primeras fotografías aéreas utilizando como plataforma los recién inventados globos aerostáticos [15]. Pero no fue sino hasta la década de los sesenta cuando de la mano de la carrera espacial tiene un impulso definitivo. El uso de los satélites para obtener datos (imágenes multiespectrales) de la superficie y la atmósfera terrestre sustituyó a las cámaras fotográficas por sensores electrónicos que permiten captar diferentes longitudes de onda del espectro electromagnético.

Las imágenes multiespectrales pueden verse como datos obtenidos por sensores que permiten captar la radiación que emiten, absorben o reflejan los objetos. Con el avance de la carrera espacial se han desarrollado sensores que cada vez captan más información, algunos ejemplos [44] son de resolución intermedia: Landsat7 con 8 bandas, Spot5 con 5, IRS con 6, de alta resolución: Ikonos2 y QuickBird2 con 3 bandas, Orb-View4 con 4, por mencionar algunos, además, ahora también los llamados sensores hiperespectrales como: HSI, ASTER, AVIRIS, CASI y HYDICE que trabajan entre 100 y 200 bandas.

Es evidente que los algoritmos de segmentación de imágenes deben ir de la mano del avance tecnológico. La teoría de los Campos Aleatorios de Markov es una rama de la probabilidad que permite analizar dependencias espaciales o contextuales de fenómenos físicos [32]. El presente trabajo propone una forma de realizar la segmentación utilizando la información de la respuesta espectral y la información contextual para imágenes multiespectrales.

1.2. Planteamiento del problema

Una considerable porción del procesamiento de imágenes digitales está enfocada a la segmentación. En este proceso la imagen es particionada en *regiones*. Este proceso es igual de importante para las imágenes multiepectrales, en cuyo caso el rango el análisis no esta limitado a una banda o al espectro visible. Partiendo de la importancia y necesidad de mejores algoritmos para la segmentación de imágenes multiespectrales,

trataremos de diseñar un algoritmo que permita particionar una imagen multiespectral en regiones homogéneas y deseamos usar la información espectral y contextual que proporcionan los píxeles en las diferentes bandas que componen la imagen.

1.3. Organización de capítulos

El primer capítulo contiene la introducción y las razones de éste trabajo. En el Capítulo 2 se dan las bases necesarias para la comprensión del marco de la segmentación en el procesamiento digital de imágenes, las diferentes técnicas que se emplean, así como las aplicaciones e importancia de este proceso.

El Capítulo 3 presenta toda la teoría necesaria para comprender los Campos Aleatorios de Markov, comienza describiendo el problema de etiquetado, del cual la segmentación de imágenes forma parte, continúa con la definición formal de los Campos Aleatorios de Markov y de Gibbs, la equivalencia entre ambos y termina mostrando los modelos más empleados.

En el Capítulo 4 se formula el modelo para las segmentación de imágenes multiespectrales empleando los Campos Aleatorios de Markov, y se propone

el algoritmo Recocido Simulado para minimizar la distribución de Gibbs y aproximar al estimador máximo a posteriori, obteniendo de esta manera una segmentación que desde en regiones homogéneas se aproxima o converge a la segmentación ideal.

Los resultados y el análisis del algoritmo propuesto se presentan en el Capítulo 5, empezando por los resultados en imágenes de una banda y posteriormente para imágenes multispectrales, mostrando para ambos casos la segmentación hecha con k-means un algoritmo que no considera la información contextual. Así también se muestra la segmentación realizada con el algoritmo de Recocido Simulado considerando los parámetros más sencillos del esquema de enfriamiento y los resultados haciendo ciertas consideraciones sobre dicho esquema. En el capítulo 6 presento mis conclusiones.

Capítulo 2

Estado del Arte

2.1. Conocimiento base

Comencemos por definir el concepto de imagen, la Real Academia Española define imagen como “Figura, representación, semejanza y apariencia de algo”, la forma más común de una imagen es la representación que se forma en nuestro cerebro de las cosas que vemos a través de nuestros ojos.

La retina de un ojo humano está cubierta de células fotorreceptoras, la luz que reflejan los objetos se enfoca en la retina pasando a través de la cornea y el cristalino y estas células fotorreceptoras la absorben. Dichas células generan impulsos nerviosos que viajan al cerebro a través del nervio óptico. La frecuencia de estos impulsos codifica la brillantez de la luz incidente.

Las células fotorreceptoras se clasifican en dos tipos, conos y bastones llamadas así por su apariencia física. Los bastones son más sensibles y nos proveen de una visión nocturna. Los conos son de tres tipos, la diferencia estriba en la fotoquímica que emplean para convertir la luz en impulsos nerviosos. Los conos dividen la porción visible del espectro electromagnético en tres bandas: rojo, verde y azul. Por esta razón, estos colores son conocidos

como colores primarios de la visión humana [27].

2.1.1. Imagen digital

Una imagen digital es la imagen que percibió un sistema de visión como el del ser humano, almacenada en algún medio o dispositivo para que esta pueda ser procesada en una computadora.

En Lira [34] se muestran los elementos que permiten generar una imagen digital. De manera general, supóngase que se tiene una fuente de radiación que incide en la escena de estudio, los objetos que componen la escena reflejan una parte de la radiación, ésta es captada por unos sensores que poseen un subsistema de enfoque y conducción de radiación, la radiación es conducida hasta un grupo de sensores frente a los cuales se encuentra un conjunto de filtros, cuya función es seleccionar intervalos de energía $\{\Delta E_1, \Delta E_2, \dots \Delta E_\gamma\}$ o de longitudes de onda $\{\Delta \lambda_1, \Delta \lambda_2, \dots \Delta \lambda_\gamma\}$, para cada intervalo de energía los sensores miden la cantidad de radiación. La señal generada por los sensores se introduce a un convertidor analógico digital que la cuantifica empleando una escala que usualmente va de 0 a 255, para así formar la imagen digital. La Figura 1 muestra lo antes mencionado.

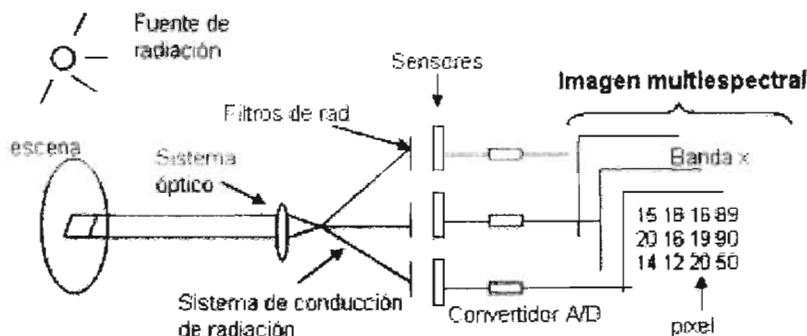


Figura 1. Generación de una imagen digital multispectral.

Definición 2.1.1. Una imagen digital es una función discreta $f(x, y) = z$, donde x e y son las coordenadas espaciales y z es el valor que toma la imagen en esas coordenadas.

También se puede definir una imagen digital como una matriz de dimensión genérica $N \times M$. Cada par de coordenadas denota un elemento de la imagen conocido como pixel (picture element). Cada pixel incorpora el nivel de radianza, es decir, z generalmente indica la intensidad de la fuente de radiación capturada por el sensor.

2.1.2. Imágenes multiespectrales

Una imagen multiespectral es un arreglo o serie de imágenes digitales, cada una de ellas contiene valores de intensidades de energía captados por sensores, la visión a color es la imagen multiespectral más familiar. El subsistema de formación de imágenes del ser humano es muy limitado en cuanto al tipo y energía de radiación perceptible [33], su sensibilidad se limita a una pequeña porción del espectro electromagnético.

Una imagen digital a color consta de tres planos, rojo (RED), verde (GREEN), y azul (BLUE). Es decir, que en este caso se habla de una función $f(x, y)$ que devuelve un vector de tres componentes, donde cada componente es el valor de intensidad correspondiente al plano en cuestión. Por ejemplo: $f(10, 50) = (140, 200, 145)$. Este vector indica que para las coordenadas (10, 50), el plano rojo tiene un valor de intensidad de 140, el plano verde de 200 y el azul de 145.

El mayor esfuerzo en el análisis multiespectral ha sido en el campo de la percepción remota [27]. La energía solar se transmite a la tierra en forma de ondas de luz, y es reflejada, absorbida o transmitida por los materiales de la superficie terrestre. La energía reflejada viaja e interactúa con la atmósfera,

y luego es registrada por un sensor montado ya sea en un avión o un satélite. Este registro es cuantificado como un número digital que se corrige y se almacena como un elemento de la imagen (pixel). Todos los pixeles forman una imagen del área de estudio.

2.2. Segmentación

Las imágenes digitales son una representación discreta de una escena que realmente es una forma continua, éstas están formadas por pixeles cuyo valor representa la respuesta espectral de los objetos.

El procesamiento de imágenes digitales es un área que abarca distintas etapas, desde la adquisición de datos de entrada, pasando por el mejoramiento, segmentación, hasta el análisis, clasificación e interpretación de la imagen. Como ya se mencionó, la segmentación es un paso imprescindible en diversos procesos del tratamiento de imágenes.

Definición 2.2.1. La segmentación es el proceso de particionar una imagen digital en regiones homogéneas disjuntas.

Una región es un conjunto de pixeles conectados, es decir, un conjunto donde todos los pixeles son adyacentes. Cuando un observador ve una escena, el proceso que realiza el sistema visual es esencialmente la segmentación de la escena para él o para ella.

Segmentar involucra categorizar los pixeles de una imagen digital en clases o temas. Cuando se tienen imágenes multiespectrales o hiperespectrales, generalmente se toma el patrón espectral de cada pixel para la segmentación. Este patrón se refiere al conjunto de medida de radiación obtenidas en varias bandas de longitud de onda para cada pixel. El reconocimiento de patrones espacial involucra la categorización de pixeles de la imagen sobre la

base de la relación espacial que guarda con los píxeles que lo rodean. Pueden considerarse aspectos como textura, proximidad, tamaño de rasgo, forma, direccionalidad, repetición, contexto, etc., e intenta imitar el análisis humano [31].

Las técnicas de segmentación [16, 39, 40, 41, 47] se pueden agrupar en aquellas donde la segmentación es realizada tomando sólo en cuenta la respuesta espectral, éstas son conocidas como técnicas orientadas a píxeles. Otras técnicas son aquellas que toman en cuenta la información contextual, es decir, la información entorno al píxel, éstas permiten modelar propiedades como textura, forma, etc.. Por último las suponen la existencia de continuidad en la frontera de los objetos, estas técnicas son conocidas como orientadas al contorno.

2.2.1. Técnicas orientadas a píxel.

Las técnicas orientadas a píxel son aquellas que asocian cada píxel de la imagen a una determinada clase según su nivel de gris. Para ello, es necesario determinar los rangos de intensidad que mejor representan a cada clase. Esta asignación de rangos puede haber sido fijada a-priori por el programador o puede ser realizada en línea empleando técnicas típicamente basadas en el análisis de histogramas. Una vez determinados los rangos, se realiza la asignación mediante la aplicación de un umbral.

Por mencionar las más clásicas encontramos a:

1. Umbrales. Es la técnica más simple, separa los elementos de la imagen basándose únicamente en la intensidad de éstos. El uso de umbrales ha sido usada en tomografía computarizada para segmentar huesos y tejidos blandos de otras estructuras [39]
2. Agrupamientos. En general son técnicas no supervisadas en donde uno tiene que generar las clases sin conocimiento a priori [7]. Algunos al-

goritmos clásicos con *k - means* [39, 42], y su variante k-means difuso [49]. ISODATA (Iterative Self-Organizing Data Analysis) [7]. Otra variante la presentan Shi et al.[45] planteando el problema de segmentación como un problema de partición de una gráfica G .

Si bien la implantación de éstos métodos suele llevar a soluciones relativamente rápidas y sencillas, adolecen el problema fundamental de ignorar las nociones de proximidad y conectividad, suponiendo que los objetos se distinguen únicamente por su valor de intensidad. Por ello, suelen limitar su área de aplicación a situaciones muy controladas donde se sabe con certeza que dichas condiciones se cumplen, como por ejemplo al pretender discernir la silueta de un objeto ante una iluminación trasera sobre un fondo claro semitransparente.

2.2.2. Técnicas orientadas al contexto.

Las técnicas orientadas al contexto explotan la noción de conectividad para agrupar zonas de la imagen con puntos de intensidad similar en entidades discernidas. Con objeto de eliminar los efectos del ruido, se han desarrollado métodos iterativos orientados a fundir entidades pequeñas en otras más grandes de valor medio de intensidad similar.

Entre dichos métodos cabe destacar el algoritmo “split and merge” [2, 24]. También cabe destacar los algoritmos basados en Programación Dinámica [2], la aplicación al problema de la teoría de Campos Aleatorios de Markov (Markov Random Fields) y el uso de la transformada “Watershed” dentro del ámbito de la Morfología Matemática [12]. El mayor inconveniente de estos métodos consiste en que son iterativos o recursivos, dificultando su implantación en sistemas que requieren un tiempo de respuesta rápido o donde se desea una configuración combinacional en “pipe-line”, reduciendo la posibilidad del paralelismo al dominio espacial. También existe una pequeña dificultad para establecer el criterio en el que una región pequeña es engullida

por otra grande, exigiendo un compromiso entre la eliminación de ruido y la pérdida de detalle. Otra dificultad es la de determinar con precisión la frontera entre regiones cuando las variaciones de intensidad son graduales, produciendo a veces un efecto de ambigüedad.

2.2.3. Técnicas orientadas al contorno.

Las técnicas orientadas al contorno se basan en la suposición de una correspondencia entre las discontinuidades en intensidad de una imagen y la frontera de los objetos contenidos en la misma. Esta suposición no es caprichosa sino que se basa en numerosos experimentos hechos con humanos y otros animales [30] que corroboran la importancia de las fronteras en los procesos de interpretación visual. Es bien sabido que los bordes pueden ser encontrados en imágenes de niveles de gris usando funciones gradiente o Laplacianas, entre las más conocidas se encuentran los operadores Roberts, Prewitt y Sobel, una revisión completa de éstos operadores la podemos encontrar en [19, 27, 46].

Por otra parte, resulta evidente pensar que dicha información no es suficiente en muchos casos, pudiendo llevar a situaciones de ambigüedad o incongruencia. Sin embargo, el mayor problema involucrado en esta suposición proviene de la naturaleza no biunívoca de la correspondencia aludida, debido a que si bien las fronteras de objetos suelen corresponder a gradientes de intensidad, existen muchos otros cambios de intensidad que no corresponden con la existencia de una frontera, sino con variaciones provocadas por la textura del objeto, por efectos de iluminación y sombras, o porque el color de la superficie no es uniforme. En cualquier caso, se puede decir que los contornos de una imagen proveen una buena base de información que en algunos casos será insuficiente y en otros desbordante, requiriendo por tanto algún tipo de tratamiento posterior.

Capítulo 3

Campos Aleatorios de Markov

3.1. Problema de etiquetado

Muchos problemas de visión pueden ser considerados como problemas de etiquetado, de manera que la solución al problema es el conjunto de etiquetas asignado a los pixeles de la imagen. El etiquetado es una representación natural para el estudio de los Campos Aleatorios de Markov [3].

Un problema de etiquetado puede ser especificado en términos de un conjunto de estados y un conjunto de etiquetas. Sea S un conjunto discreto con m estados

$$S = \{1, \dots, m\}$$

Un estado representa comúnmente un punto o una región en el espacio euclidiano, tal como un pixel en una imagen o una característica como una línea o una superficie. El conjunto de estados puede ser categorizado en términos de su homogeneidad. Los estados sobre una malla son considerados como espacialmente homogéneos.

Una etiqueta es un evento que tal vez ocurre a un estado. Sea Λ un conjunto de etiquetas, éste puede ser discreto o continuo, para el objetivo

del presente trabajo supongamos en adelante Λ un conjunto discreto, donde una etiqueta asume un valor en un conjunto de M valores.

$$\Lambda = \{l_1, \dots, l_M\}$$

3.1.1. Problemas de etiquetado en visión

El problema de etiquetado consiste en asignar una etiqueta del conjunto de etiquetas Λ a cada estado de S . Por ejemplo, la detección de bordes en una imagen consiste en asignar una etiqueta l_i dentro del conjunto $\Lambda = \{\text{borde}, \text{noborde}\}$ para cada estado $i \in S$, donde S está formado por los pixeles de la imagen.

Definición 3.1.1. El conjunto

$$f = \{f_1, \dots, f_m\}$$

Es llamado un *etiquetado* de los estados de S en términos de las etiquetas en Λ .

Definición 3.1.2. Llamaremos "*mapeo*" a la función

$$f: S \longrightarrow \Lambda$$

que a cada estado de S le asigna una única etiqueta, f_i de Λ .

En la terminología de los Campos Aleatorios de Markov, llamaremos *configuración* al etiquetado.

Cuando todos los estados tienen el mismo conjunto de etiquetas Λ , el conjunto de todas las configuraciones posibles, es el producto cartesiano

$$\mathbb{F} = \underbrace{\Lambda \times \Lambda \times \dots \times \Lambda}_m = \Lambda^m$$

donde m es el tamaño de S .

En ciertas circunstancias, el conjunto de etiquetas puede no ser común a todos los estados. En un caso extremo, cada estado i puede tener su propio conjunto de etiquetas posibles Λ_i y esto nos lleva al siguiente espacio de configuraciones:

$$\mathbb{F} = \Lambda_1 \times \Lambda_2 \times \dots \times \Lambda_m$$

3.1.2. Clasificación del problema del etiquetado en visión.

En términos de homogeneidad y de continuidad, podemos clasificar un problema de etiquetado en visión dentro de una de las cuatro categorías siguientes:

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

Las dos primeras categorías caracterizan el procesado de bajo nivel realizado sobre las imágenes observadas, y las dos últimas el procesado de alto nivel realizado sobre las características extraídas.

La restauración de imágenes en niveles de gris, o el suavizado de imágenes, son ejemplos de PE1. El conjunto de estados S corresponde a los píxeles de la imagen y el conjunto de etiquetas Λ a un intervalo real.

La segmentación de imágenes es un PE2. Se divide una imagen en regiones mutuamente excluyentes, cada una de las cuales tiene algunas propiedades uniformes y homogéneas cuyos valores son significativamente diferentes a los de las regiones vecinas.

El reconocimiento y emparejamiento de objetos basados en características es un PE3. Cada estado indica una característica de la imagen tal como un punto, un segmento de línea o una región, las etiquetas son un conjunto discreto por naturaleza y cada una de ellas indica una característica del modelo.

La estimación de la orientación a partir de un conjunto de correspondencia de puntos puede ser formulada como un PE4. Cada etiqueta puede asumir el valor de una matriz real representando una transformación admisible (ortogonal, afín o en perspectiva).

Para un problema de etiquetado discreto con m estados y M etiquetas existen M^m configuraciones posibles. Entre todas las posibilidades, solamente algunas son óptimas en términos de algún criterio que mida el costo de las soluciones. Esto es lo que llamamos aproximación a la optimización del etiquetado en visión.

3.1.3. Etiquetado con restricciones de contexto

El uso de información de contexto es, en última instancia, imprescindible en la comprensión de una imagen [41]. Abend et al. [1] usan la información de contexto en análisis de imagen y en reconocimiento de patrones. El reconocimiento de caracteres se considera como un problema estadístico de decisión. Las dependencias del sistema de vecinos entre los píxeles de una imagen se obtienen suponiendo que hay independencia estadística. La información sobre el sistema de vecinos se utiliza para el cálculo de las probabilidades condicionales. Ese sistema también incluye la estimación de los parámetros en el muestreo de caracteres; el reconocimiento se hace mediante la estimación de los parámetros. El trabajo de Abend es probablemente el primero que utiliza los modelos de Campos Aleatorios de Markov para el reconocimiento de patrones. Se utiliza un modelo de Campos Aleatorios de Markov para reducir el número de parámetros necesarios en el proceso usan-

do restricciones de contexto. Otro desarrollo de modelos basados en contexto es el llamado “relaxation labeling” (RL). RL es una clase de procedimientos iterativos que utilizan restricciones de contexto para reducir ambigüedades en el análisis de imagen. Haralick [21] desarrolla una teoría para explicar RL desde un punto de vista de Bayes.

En términos de probabilidad, las restricciones de contexto pueden expresarse localmente en término de probabilidades condicionales $P(f_i | \{f_j\})$, donde $\{f_j\}$ denota el conjunto de etiquetas del estado $j \neq i$; y la probabilidad conjunta global se denota por $P(f)$.

En los casos donde las etiquetas son independientes unas de otras, la probabilidad conjunta es el producto

$$P(f) = \prod_{i \in S} P(f_i) \quad (3.1.1)$$

3.2. Campos Aleatorios de Markov

Los Campos Aleatorios de Markov son una parte de la Teoría de la Probabilidad que proporciona una herramienta para analizar dependencias espaciales o contextuales de fenómenos físicos. Se usa en el problema del etiquetado en visión para establecer las distribuciones de probabilidad de las etiquetas. Para definir los Campos Aleatorios de Markov supongamos S una malla regular.

3.2.1. Sistema de vecinos y clique

Definición 3.2.1. Dos estados i y j son vecinos si:

$$dist(i, j)^2 \leq r, \quad j \neq i$$

donde $dist(a, b)$ denota la distancia euclídeana entre a y b , y r es un número entero.

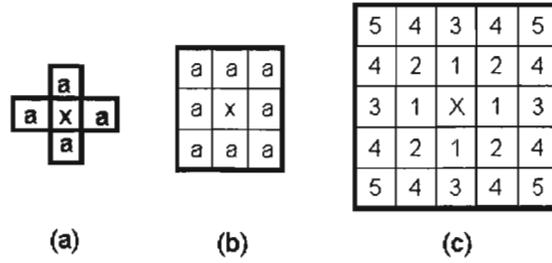


Figura 3.1: Vecinos en una malla regular S

El conjunto de estados que son vecinos al estado i es denotado por N_i , es decir $N_i = \{j \in S \mid dist(i, j)^2 \leq r, \quad j \neq i\}$. Los estados de S están relacionados mediante lo que llamaremos un *sistema de vecinos*.

Definición 3.2.2. Un *sistema de vecinos* para el conjunto de estados S esta definido por

$$N = \{N_i \mid \forall i \in S\}$$

donde N_i es el conjunto de los estados vecinos a i para los que

1. $i \notin N_i$
2. $i \in N_j \iff j \in N_i$

En el sistema de vecinos de primer orden, llamado sistema de 4-vecinos, cada estado (interior) tiene cuatro vecinos, según lo mostrado en la Figura 3.1(a), donde x denota al estado considerado y a sus vecinos. En el sistema de vecinos de segundo orden, también llamado sistema de 8-vecinos, hay ocho vecinos para cada estado (interior), según lo mostrado en la Figura 3.1(b). Los números $n = 1, \dots, 5$ mostrados en la Figura 3.1(c) indican los estados vecinos exteriores en el sistema de vecinos de orden n -ésimo.

Cuando se ha especificado un orden entre los elementos de S , el conjunto de vecinos se puede determinar más explícitamente. Por ejemplo, el conjun-

to de estados de una malla regular rectangular $S = \{(i, j) \mid 1 \leq i, j \leq n\}$ corresponde a los pixeles de una imagen en dos dimensiones de tamaño $n \times n$, cada estado interior, representado por (i, j) , tiene 4 vecinos $N_{i,j} = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}$, los estados de la frontera tienen tres y los estados de las esquinas tienen dos.

3.2.2. Campos Aleatorios de Markov

Definición 3.2.3. Sea $F = \{F_1, \dots, F_m\}$ una familia de variables aleatorias definidas en S , donde cada variable aleatoria F_i toma un valor f_i en Λ . Llamaremos *Campo Aleatorio* a la familia de variables aleatorias F .

Usaremos la siguiente notación: $F_i = f_i$ para denotar que el suceso F_i toma el valor f_i .

Definición 3.2.4. Un suceso conjunto $\{F_1 = f_1, \dots, F_m = f_m\}$, abreviadamente $F = f$, es una *realización* de F . Llamaremos *configuración* del suceso conjunto F , a $f = \{f_1, \dots, f_m\}$.

Para un conjunto discreto de etiquetas Λ , denotamos por $P(F_i = f_i)$ a la probabilidad de que la variable aleatoria F_i tome el valor f_i , y denotamos la probabilidad conjunta por $P(F = f) = P(F_1 = f_1, \dots, F_m = f_m) = P(f)$.

Definición 3.2.5. Una familia de variables aleatorias F se dice que es un *Campo Aleatorio de Markov* sobre S con respecto a N si y sólo si:

1. $P(f) > 0 \quad \forall f \in \mathbb{F}$
2. $P(f_i \mid f_{S-i}) = P(f_i \mid f_{N_i})$ (Propiedad Markoviana)

La condición (1) es para exigir que F sea un Campo Aleatorio. La condición (2) se llama característica local, se dice que la probabilidad de un suceso local a i condicionado a todos los sucesos restantes es equivalente a aquella

condicionada a los sucesos vecinos de i . Puede demostrarse que la probabilidad conjunta $P(F = f)$ de cualquier Campo Aleatorio se determina únicamente por estas probabilidades condicionales locales [3]. Cualquier F cumpliendo estas condiciones es un Campo Aleatorio de Markov con respecto a tal sistema de vecinos.

Un Campo Aleatorio de Markov puede tener otras características tales como homogeneidad e isotropía. Diremos que es homogéneo si $P(f_i | f_{N_i})$ se calcula sin importar la posición relativa del estado i en S .

Para ciertos problemas puede ser necesario definir más de un Campo Aleatorio de Markov, cada uno de ellos definido en uno de los conjuntos espaciales de estados con intersección no vacía.

Hay dos aproximaciones para especificar un Campo Aleatorio de Markov, en términos de probabilidades condicionales $P(f_i | f_{N_i})$ y en términos de probabilidad conjunta $P(f)$. Besag [3] discute la aproximación mediante la probabilidad conjunta en vista de las desventajas de la aproximación mediante la probabilidad condicional. Afortunadamente, un resultado teórico sobre la equivalencia entre los Campos Aleatorios de Markov y la distribución de probabilidad de Gibbs [20, 3] proporciona medios de especificar la probabilidad conjunta de un Campo Aleatorio de Markov matemáticamente manejables.

3.2.3. Campos Aleatorios de Gibbs

El par $\{S, N\}$ es una gráfica en el sentido usual, donde S contiene los nodos y N define las aristas.

Definición 3.2.6. Una *clique* c para $\{S, N\}$ se define como un subconjunto de S tal que c consiste en un único estado $c = \{i\}$ o un par de estados vecinos $c = \{i, j\}$, o tres estados vecinos $c = \{i, j, k\}$ y así sucesivamente.

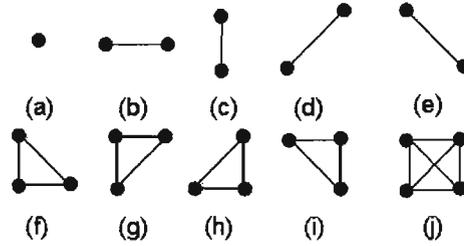


Figura 3.2: Cliques en una malla regular S

Denotamos el conjunto de cliques de un solo estado, el conjunto de cliques de dos estados,... por C_1, C_2, \dots respectivamente, donde

$$C_1 = \{i \mid i \in S\}$$

$$C_2 = \{(i, j) \mid j \in N_i, i \in S\}$$

$$C_3 = \{(i, j, k) \mid i, j, k \in S\} \text{ son vecinos dos a dos}$$

El conjunto de todas las cliques para $\{S, N\}$ será entonces $C = C_1 \cup C_2 \cup \dots$

El tipo de clique para $\{S, N\}$, de una malla regular, viene determinado por su tamaño, dimensión y orientación. La Figura 3.2 muestra distintos tipos de cliques para sistemas de vecinos de primer(a,b,c) y segundo orden(a,b,c,d,e,f,g,h,i,j). A medida que aumenta el orden del sistema de vecinos, el número de cliques crece rápidamente y también el costo computacional.

Definición 3.2.7. Diremos que un conjunto de variables aleatorias F es un *Campo Aleatorio de Markov* en S con respecto a N si y sólo si sus configuraciones siguen una distribución de Gibbs. Una distribución de Gibbs tiene la siguiente función de densidad:

$$P(f) = \frac{1}{Z} e^{-\frac{1}{T}U(f)} \quad (3.2.1)$$

donde:

$U(f)$ =función de energía;

Z =Constante de normalizacion; $Z = \sum_{f \in \mathbb{F}} e^{-\frac{1}{T}U(f)}$

T =Temperatura.

La función de energía $U(f)$ se define por:

$$U(f) = \sum_c V_c(f) \quad (3.2.2)$$

la suma de las funciones “*clique potenciales*” para todas las cliques posibles C , donde V es un potencial. El valor de $V_c(f)$ depende de la configuración local de la clique c .

Diremos que un Campo Aleatorio de Gibbs es homogéneo si $V_c(f)$ es independiente de la posición relativa de la clique c en S . La homogeneidad se asume en la mayoría de los modelos de Campos Aleatorios de Markov, en visión por conveniencia matemática y computacional. La isotropía es una propiedad de regiones independientes de la dirección.

$P(f)$ mide la probabilidad de la ocurrencia de configuración determinada, o patrón f . Las configuraciones más probables son aquellas con energías más bajas. La temperatura T controla la forma de la distribución. Cuando la temperatura es alta, todas las configuraciones tienden a estar igualmente distribuidas. Con la temperatura cercana a cero, la distribución se concentra alrededor de los mínimos globales de la energía. Dada T y $U(f)$, podemos generar la “*clase de patrones*” muestreando el espacio de configuraciones \mathbb{F} de acuerdo con $P(f)$.

Para los problemas de etiquetado discretos, una función *clique potencial* $V_c(f)$ se puede especificar por un número de parámetros. Por ejemplo, sea $f_c = (f_i, f_j, f_k)$ la configuración local en una clique triple $c = \{i, j, k\}$, f_c toma un número finito de estados y por lo tanto, toma un número finito de

valores.

A veces, puede ser conveniente expresar la energía de una distribución de Gibbs como la suma de varios términos, cada uno correspondiente a cliques de distinto tamaño, es decir:

$$U(f) = \sum_c V_c(f) = \sum_{\{i\} \in C_1} V_1(f_i) + \sum_{\{i,j\} \in C_2} V_2(f_i, f_j) + \dots \quad (3.2.3)$$

La función anterior implica una distribución homogénea de Gibbs porque V_1, V_2, \dots , son independientes de las localizaciones i, j, k . En distribuciones no homogéneas de Gibbs, las funciones clique se deben escribir como: $V_1(i, f_i)$, $V_2(i, j, f_i, f_j), \dots$

Un caso especialmente importante de (3.2.3) es cuando sólo se consideran las cliques de tamaño dos como máximo. En este caso, la energía puede también ser escrita como

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{j \in N_i} V_2(f_i, f_j) \quad (3.2.4)$$

Considerando el segundo termino de la ecuación anterior, y siendo $\{i, j\}$ y $\{j, i\}$ dos cliques distintas en C_2 (ya que los estados de la clique están ordenados), la probabilidad condicional puede escribirse como

$$P(f_i | f_{N_i}) = \frac{e^{-[V_1(f_i) + \sum_{j \in N_i} V_2(f_i, f_j)]}}{\sum_{f_i \in L} e^{-[V_1(f_i) + \sum_{j \in N_i} V_2(f_i, f_j)]}} \quad (3.2.5)$$

3.2.4. Equivalencia Markov-Gibbs

Un Campo Aleatorio de Markov se caracteriza por sus propiedades locales (propiedades Markovianas) mientras que un Campo Aleatorio de Gibbs se

caracteriza por sus propiedades globales (la distribución de probabilidad de Gibbs). El teorema de Hammersley-Clifford [20] establece la equivalencia entre estos dos tipos de características.

Teorema 3.2.1. *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 .*

Veamos 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 . Consideremos la probabilidad condicional

$$P(f_i | f_{N_i}) = \frac{P(f_i, f_{N_i})}{P(f_{N_i})} = \frac{P(f)}{\sum_{f'_i \in L} P(f')} \quad (3.2.6)$$

donde $f' = f_1, \dots, f_{i-1}, f'_i, \dots, f_m$ es cualquier configuración que contenga todos los estados de f menos, posiblemente, el i .

Sustituyendo $P(f) = \frac{1}{Z} e^{-\sum_{c \in C} V_c(f)}$ lo cual también proporciona una fórmula para calcular la probabilidad condicional de las funciones potenciales, tenemos que:

$$P(f_i | f_{N_i}) = \frac{e^{-\sum_{c \in C} V_c(f)}}{\sum_{f'_i \in L} e^{-\sum_{c \in C} V_c(f')}} \quad (3.2.7)$$

Dividimos L en dos conjuntos A y B de manera que A es el conjunto de cliques que contienen a i y B el conjunto de cliques que no contienen a i .

Entonces podemos escribir:

$$P(f_i | f_{N_i}) = \frac{[e^{-\sum_{c \in A} V_c(f)}] [e^{-\sum_{c \in B} V_c(f)}]}{\sum_{f'_i \in L} [e^{-\sum_{c \in A} V_c(f')}] [e^{-\sum_{c \in B} V_c(f')}]} \quad (3.2.8)$$

Como $V_c(f) = V_c(f')$ para cualquier clique c que no contenga a i , $e^{-\sum_{c \in B} V_c(f)}$ se simplifica en el numerador y en el denominador. Por lo tanto, esta probabilidad depende solamente de los potenciales de las cliques que contienen i ,

$$P(f_i | f_{N_i}) = \frac{[e^{-\sum_{c \in A} V_c(f)}]}{\sum_{f'_i \in L} [e^{-\sum_{c \in A} V_c(f')}]} \quad (3.2.9)$$

es decir, depende de la etiqueta en los vecinos de i . Esto prueba que un campo aleatorio de Gibbs es un campo aleatorio de Markov. La demostración completa se puede encontrar en [3, 37, 28]

El valor práctico del teorema es que proporciona una manera simple de especificar la probabilidad conjunta. Se puede definir la probabilidad conjunta $P(F = f)$ especificando las funciones clique potenciales $V_c(f)$ y eligiendo las funciones potenciales apropiadas para el comportamiento deseado del sistema. De esta manera, se codifica el conocimiento a priori sobre las interacciones entre las etiquetas.

Cómo elegir las formas y los parámetros de las funciones potenciales para una codificación apropiada de restricciones es muy importante a la hora de modelar un Campo Aleatorio de Markov. Las formas de las funciones potenciales determinan la forma de la distribución de Gibbs. La distribución de Gibbs está completamente definida cuando se han especificado todos los parámetros de las funciones potenciales.

Para calcular la probabilidad conjunta de un Campo Aleatorio de Markov, que tiene una distribución de Gibbs, es necesario evaluar Z . Dado que la suma es un número combinatorio de configuraciones, el costo computacional es muy elevado. La evaluación explícita se puede evitar mediante la probabilidad-máxima en modelos de visión basados en Campos Aleatorios de Markov

cuando $U(f)$ no contiene parámetros desconocidos.

Sin embargo, esto no es verdad cuando la estimación de los parámetros es también una parte del problema. En el último caso, la función de energía $U(f) = U(f | \theta)$ es también función de los parámetros θ , así también $Z = Z(\theta)$. Para evitar la gran dificultad planteada, en la práctica se calcula una aproximación de la probabilidad conjunta.

3.2.5. Automodelos

Las restricciones de contexto entre dos etiquetas son un primer nivel a tener en cuenta en la información del contexto. Son muy utilizadas debido a su simplicidad y bajo costo computacional. Se codifican en la función de energía de Gibbs como funciones clique potenciales para cada par de estados. Es decir, la función clique potencial queda por la ecuación (3.2.4) llamada energía de segundo orden y se calcula para cada clique formada por dos estados.

Podemos especificar un Campo Aleatorio de Gibbs o un Campo Aleatorio de Markov mediante la selección apropiada de V_1 y V_2 .

Cuando $V_1(f_i) = f_i G_i(f_i)$ y $V_2(f_i, f_j) = \beta_{i,j} f_i f_j$, donde $G_i(\Delta)$ son funciones arbitrarias y $\beta_{i,j}$ son constantes que reflejan la interacción entre cada par de estados i y j , entonces la función de energía es

$$U(f) = \sum_{\{i\} \in C_1} f_i G_i(f_i) + \sum_{\{i,j\} \in C_2} \beta_{i,j} f_i f_j \quad (3.2.10)$$

Definición 3.2.8. El modelo definido por la ecuación (3.2.10) recibe el nombre de *auto-modelo* [3]. Los *auto-modelo* se pueden clasificar más a fondo según las suposiciones hechas sobre las f_i .

Un *auto-modelo* se dice que es un modelo *auto-logístico*, si las f_i toman

valores en el conjunto de etiquetas discreto $\Lambda = \{0, 1\}$ (o $\Lambda = \{+1, -1\}$). La correspondiente función de energía es de la forma

$$U(f) = \sum_{\{i\} \in C_1} \alpha f_i + \sum_{\{i,j\} \in C_2} \beta_{i,j} f_i f_j \quad (3.2.11)$$

donde $\beta_{i,j}$ puede ser interpretado como la interacción entre los coeficientes.

Cuando el sistema de vecinos N es una malla (4-vecinos en una malla 2D, o 2-vecinos en una malla 1D), entonces el modelo *auto-logístico* se reduce al modelo de *Ising*. La probabilidad condicional para el modelo *auto-logístico* con $\Lambda = \{0, 1\}$ es

$$P(f_i | f_{N_i}) = \frac{e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}}{1 + e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}} \quad (3.2.12)$$

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

Capítulo 4

Modelo de Segmentación para Imágenes Multiespectrales

Como ya se mencionó en el capítulo dos, el problema de segmentación puede ser visto como un problema de etiquetado, donde el problema es dividir a la imagen en regiones mutuamente excluyentes, asignando de acuerdo a cierta característica una etiqueta λ_i de un conjunto Λ' finito discreto de etiquetas.

$$\Lambda' = \{\lambda_1, \dots, \lambda_k\}$$

4.1. Modelo de la Imagen Multiespectral

Sea $S = \{\overline{(i, j)} \mid 1 \leq i \leq n, 1 \leq j \leq m\}$ el conjunto de estados que denota una malla rectangular, donde cada estado representa al vector conformado por los valores de los pixeles (i, j) de una imagen multiespectral de B bandas como lo muestra la Figura 4.1 (a).

Sea $S' = \{(i, j) \mid 1 \leq i \leq n, 1 \leq j \leq m\}$ el conjunto de estados que denota una malla rectangular, y que representa a los pixeles de una imagen de etiquetas de $n \times m$ como lo muestra la Figura 4.1 (b).

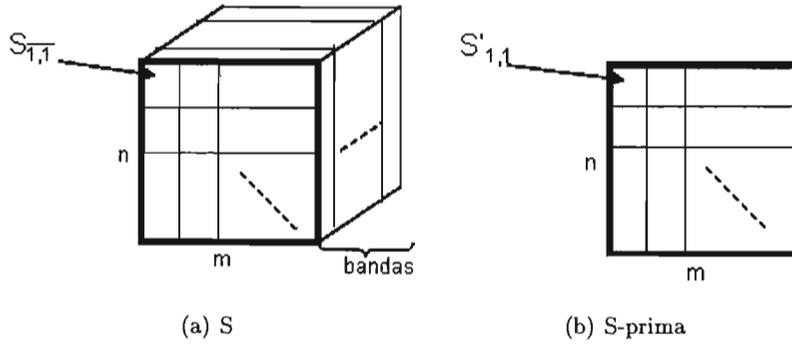


Figura 4.1: Conjuntos de estados

Sean $N_{i,j} \subset S$ y $N'_{i,j} \subset S'$, los conjuntos de pixeles vecinos al los estados $\overline{(i,j)}$ y (i,j) respectivamente y sean $C_{i,j}$ y $C'_{i,j}$ los conjuntos de cliques C de orden uno asociados a $N_{i,j}$ y $N'_{i,j}$ que contienen al los pixeles $\overline{(i,j)}$ y (i,j) .

4.2. Campos Aleatorios de Markov

En nuestro modelo deseamos tomar en consideración la información contextual, es decir la asignación de la etiqueta no sólo dependerá de la respuesta espectral que tenga el pixel, sino también dependerá del valor espectral de los pixeles vecinos

Si S' es una imagen de etiquetas con $M = nxm$ píxeles, la imagen verdadera de etiquetas la denotamos por $l^* = \{l_1^*, l_2^*, \dots, l_M^*\}$. Entonces, el objetivo se puede ver como la estimación de l^* , es decir, el problema es encontrar el estimador óptimo de la imagen verdadera l^* .

Sea $F = \{F_{1,1}, \dots, F_{i,j}, \dots, F_{n,m}\}$, el Campo Aleatorio de Markov definido en S que representa los datos de la imagen (observaciones), con $F_{i,j}$ tomando valores $f_{i,j}$, donde $f_{i,j}$ es el vector de observaciones para el pixel $\overline{(i,j)}$, donde

cada coordenada toma valores en el conjunto Λ .

$$\Lambda = \{0, \dots, 255\}$$

El conjunto de todas las posibles observaciones $f = \{f_{1,1}, \dots, f_{n,m}\}$ está denotado por Φ .

La información contextual puede introducirse en el problema de etiquetado de pixeles a través de un Campo Aleatorio de Markov L , el cual modela la dependencia estadística entre los pixeles vecinos. La imagen verdadera l^* puede verse como una realización de dicho Campo Aleatorio de Markov donde cada variable aleatoria $L_{i,j}$ toma valores en el conjunto de etiquetas Λ' .

Sea $L = \{L_{1,1}, \dots, L_{i,j}, \dots, L_{n,m}\}$, $(i,j) \in S'$ el otro Campo Aleatorio de Markov definido en S' que representa las etiquetas de la imagen de etiquetas, con $L_{i,j}$ tomando valores $l_{i,j}$, donde $l_{i,j}$ es el valor que tiene la imagen segmentada en un conjunto Λ' .

$$\Lambda' = \{\lambda_1, \dots, \lambda_k\}$$

Ω es el conjunto de todas las posibles configuraciones $l = \{l_{1,1}, \dots, l_{n,m}\}$, $l_{i,j} \in \Lambda'$. L sirve como distribución a priori para las etiquetas que se están estimando.

El campo aleatorio F es llamado el campo de observación y L es llamado el campo de etiquetado. Asumamos que las imágenes multispectrales consisten en diversas clases que son realizaciones de una colección de las variables aleatorias (L, F) que interactúan recíprocamente.

Construyamos primero un estimador Bayesiano para el campo de etiquetado

$$P(l|f) = \frac{P(f, l)}{P(f)} = \frac{P(f|l)P(l)}{P(f)} \quad (4.2.1)$$

Dado que la realización del campo de observación es conocida, $P(f)$ es constante lo que implica que podemos escribir

$$P(l|f) \approx P(f|l)P(l) \quad (4.2.2)$$

El estimador es la siguiente función de decisión δ

$$\delta : \Phi \rightarrow \Omega$$

$$f \rightarrow \delta(f) = \hat{l}$$

El estimador que tiene el mínimo riesgo de Bayes está definido por:

$$\hat{l} = \operatorname{argmin}_{l' \in \Omega} \int_{l \in \Omega} R(l, l') P(l|f) dl \quad (4.2.3)$$

El estimador máximo a posteriori (MAP) es el estimador más usado en el procesamiento de imágenes

$$\hat{l}^{MAP} = \operatorname{argmax}_{l \in \Omega} P(l|f) \quad (4.2.4)$$

Si conocemos las probabilidades $P(f|l)$ y $P(l)$ definidas para toda la imagen podemos calcular el valor de L que hace máxima la probabilidad a posteriori (MAP). Sin embargo dado que estamos trabajando con un conjunto de k etiquetas no ordenadas, el número de configuraciones posibles es de $k^{n \times m}$. Resulta evidente que incluso para imágenes pequeñas y para un reducido número de etiquetas no es posible calcular de forma directa el valor óptimo.

4.3. Modelo de Segmentación

Dado el gran esfuerzo computacional que acarrea el cálculo del MAP se han desarrollado algoritmos que intentan estimar el MAP.

Las ecuaciones (4.2.2) y (4.2.4) nos permite escribir \hat{l}^{MAP} como:

$$\hat{l}^{MAP} = \operatorname{argmax}_{l \in \Omega} P(f|l)P(l) \quad (4.3.1)$$

Recordemos también que si L es un Campo Aleatorio de Markov en S' sobre N' entonces L tiene una distribución de Gibbs

$$P(l) = \frac{1}{Z} e^{-\frac{1}{T}U'(l)} \quad (4.3.2)$$

$$U'(l) = \sum_{C \in C'} V'_C(l) \quad (4.3.3)$$

donde $U'(l)$ es una función de energía

Las observaciones están relacionadas al proceso de etiquetas a través del modelo de degradación [17], que modela la relación entre el campo de etiquetas L y el proceso de observaciones F . En dicho modelo se observa que la probabilidad a posteriori también tiene una distribución de Gibbs.

Si suponemos que $P(f|l)$ es un Campo Aleatorio de Markov Gausiano, entonces $P(l|f)$ tiene una distribución de Gibbs también y podemos decir entonces que el problema de encontrar el estimador \hat{l} se puede ver como la minimización de la función de energía U

$$P(l|f) \approx \frac{1}{Z} e^{-\frac{1}{T}U(l,f)} \quad (4.3.4)$$

con

$$U(l, f) = U_1(l, f) + U_2(l) \quad (4.3.5)$$

donde

$$U_1 = \frac{B}{2} \ln |2\pi \Sigma_k| + \frac{1}{2} (f_{i,j} - \mu_k)^T \Sigma_k^{-1} (f_{i,j} - \mu_k) \quad (4.3.6)$$

ya que

$$P(f|l) = \frac{1}{(2\pi)^{\frac{B}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2} (f_{i,j} - \mu_k)^T \Sigma_k^{-1} (f_{i,j} - \mu_k)} \quad (4.3.7)$$

donde B es el número de bandas de la imagen multiespectral. Σ_k y μ_k son la matriz de covarianzas y la media (vector) de la clase k ($\lambda_k \in \Lambda'$).

y

$$U_2(l) = \sum_{c \in C} V_2(l_C) \quad (4.3.8)$$

ya que

$$P(l) = \frac{1}{Z} e^{-\frac{1}{T} U_2(l)} \quad (4.3.9)$$

con

$$V_2(l_C) = \begin{cases} \beta & \text{if } i = j, \\ -\beta & \text{if } i \neq j \end{cases} \quad (4.3.10)$$

Empleamos el algoritmo (Simulated Annealing) o Recocido Simulado para la minimización de la energía $U(l, f)$.

4.4. Recocido Simulado

La técnica de Recocido Simulado, es una técnica de búsqueda aleatoria dirigida, introducida en 1983 por Kirkpatrick *et. al* [29]. El término de “Simulated Annealing”, se puede traducir al castellano como “Recocido Simulado”. El término *Recocido* según el Diccionario de la Lengua Española de la Real Academia, hace referencia al “calentamiento de metales para que adquieran nuevamente la ductilidad o temple que suelen perder al trabajarlos”. En seguida damos la razón de ésta analogía cuando tratamos de resolver un problema de optimización.

Como se indica en [29], vemos que existe una relación muy estrecha entre la Mecánica Estadística que estudia el comportamiento de sistemas de múltiples grados de libertad en equilibrio térmico, y los problemas de Optimización Combinatoria, que buscan minimizar una función objetivo que depende de varios parámetros. Dado el elevado número de átomos que constituyen una pequeña cantidad de materia, solamente se observa en los experimentos aquel estado más probable del sistema en equilibrio térmico a una temperatura dada.

Cada configuración del sistema definida por un conjunto de posiciones de los átomos $\{r_i\}$ tiene asociada su factor de probabilidad de Boltzmann, $e^{-\frac{E(\{r_i\})}{k_B T}}$, donde $E(\{r_i\})$ representa la energía de la configuración, k_B es la constante de Boltzmann y T es la temperatura.

La técnica de Recocido Simulado, que de aquí en adelante denominaremos SA (por sus siglas en inglés) difiere de las técnicas tradicionales de mejora iterativa, en su capacidad para escapar de mínimos locales gracias al empleo del criterio Metrópolis para la aceptación de los desplazamientos.

Este algoritmo pertenece a la clase de algoritmos que pretenden resolver un problema de optimización combinatoria.

Un ejemplo de optimización combinatoria, es el problema del agente viajero; Dadas N ciudades y la distancia d_{ij} representando el costo del viaje entre la ciudad i y la ciudad j . El problema es planear la ruta óptima del agente viajero, dado que éste debe visitar cada ciudad exactamente una vez y finalmente regresar al punto de partida.

El ejemplo antes mencionado, es un problema de la clase NP-Completo. Es bien sabido que no hay un método para encontrar la solución exacta a tal problema sin un enorme esfuerzo de cálculo de cómputo. El algoritmo “Recocido Simulado” es una de las soluciones heurísticas propuestas para resolver dicho problema.

4.4.1. Descripción del algoritmo SA

Todo algoritmo SA (Figura 4.2(a)) se puede caracterizar por una serie de parámetros, unos específicos del problema en cuestión y otros genéricos

independientes de la naturaleza del problema.

Como parámetros específicos del problema tenemos:

- Un espacio de soluciones Ω , conjunto finito de todas las soluciones del problema.
- Una función de costo U
- Un mecanismo de generación, que permite el desplazamiento desde una solución a otra de su entorno, entendiéndose como entorno de una solución aquellas soluciones a las que podemos llegar desde una solución dada aplicando el mecanismo de generación.

Como parámetros genéricos, llamados a veces “esquema de enfriamiento”, tenemos:

- Temperatura inicial T_0 .
- Función de decremento de la temperatura o ley de evolución.
- Criterio de equilibrio.
- Criterio de congelación.

Este conjunto de parámetros es independiente del tipo de problema, y controla el funcionamiento del algoritmo. En la Figura 4.2(b) se presenta un diagrama de flujo detallado de lo que sería un algoritmo SA estándar, incluyendo un contador R de los escalones de temperatura realizados.

En nuestro caso agregamos el parámetro β que controla la homogeneidad de las regiones.

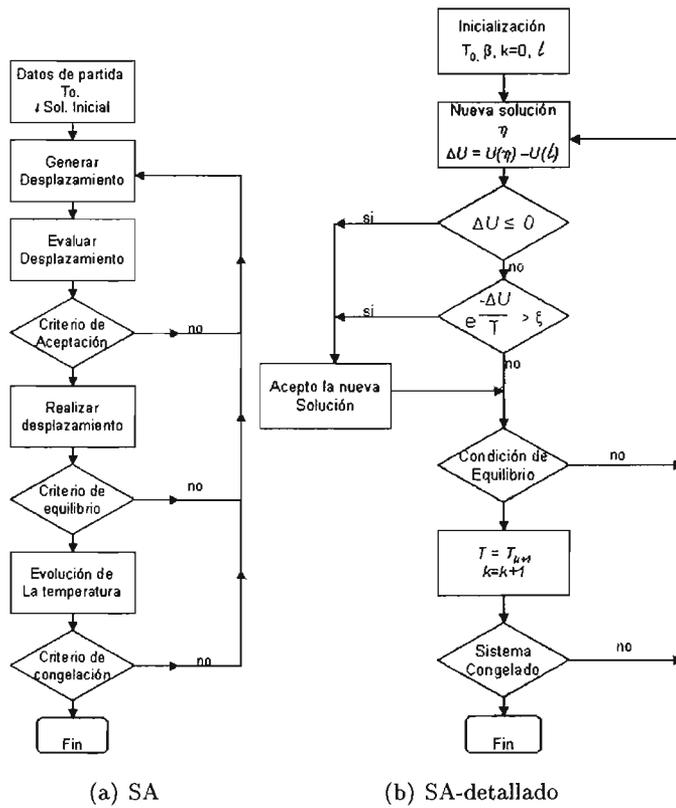


Figura 4.2: Algoritmo Recocido Simulado

4.4.2. Algoritmo SA para Segmentación de Imágenes Multi-espectrales

1. Asignar $k = 0$ e inicializar l de manera aleatoria. Escoger una temperatura inicial lo suficientemente alta $T = T_0$.
2. Construir una perturbación η de l , tal que η difiera en sólo un elemento de l .
3. (Criterio Metrópolis) Calcular $\Delta U = U(\eta) - U(l)$ y aceptar η si $\Delta U < 0$, si no, aceptar η con probabilidad $e^{-\frac{\Delta U}{T}}$.
4. Ir al paso 2 hasta que el criterio de equilibrio sea aceptado.
5. Decrementar la temperatura: $T = T_{k+1}$ e ir al paso 2 con $k = k + 1$ hasta que el sistema esté congelado.

Criterio Metrópolis: Se calcula la variación de energía provocada por el movimiento. Si $\Delta E < 0$ entonces, dicho movimiento es permitido, ya que el movimiento conduce al sistema a un estado con menor nivel energético. Si $\Delta E > 0$ entonces el desplazamiento es permitido con una probabilidad $e^{-\frac{\Delta E}{kT}}$; para ello se genera un número aleatorio ξ entre 0 y 1, de tal forma que $\xi < e^{-\frac{\Delta E}{kT}}$ se desplaza la partícula a su nueva posición, en caso contrario la partícula se regresa a su posición inicial.

4.5. Consideraciones en el esquema de enfriamiento

Varios estudios teóricos, demuestran que si T decrece lo suficientemente lento, el proceso converge a una solución óptima. Como ya se mencionó el *esquema de enfriamiento* controla el funcionamiento del algoritmo, entonces se hacen las siguientes consideraciones y observaciones:

Temperatura inicial.

La temperatura inicial T_0 debe ser lo suficientemente grande para que casi todo (o todo) movimiento sea permitido, es decir, que la probabilidad de pasar del estado i al j (en N_i) sea muy alta. Kirpatrick et. al [29] proponen una regla empírica, escoger un valor grande de T_0 y realizar cierto número de transiciones. Si se define el coeficiente de aceptación χ como el cociente entre el número de transiciones aceptadas y el número de transiciones intentadas, entonces si el valor obtenido de χ es menor de una cierta cantidad, por ejemplo 0.8 entonces se dobla el valor de T_0 , procediendo así sucesivamente hasta verificar la condición impuesta.

En [17] sugieren un valor para la temperatura inicial $T_0 = 4$. En el presente trabajo adoptamos esta última probando los valores 1, 3 y 4 para T_0 .

Temperatura final (criterio de congelación)

El valor final de T es también importante, teóricamente T debería reducirse hasta 0, pero en la práctica la búsqueda converge por lo general a su óptimo local final bastante antes de ese valor nulo de la temperatura. Por lo tanto, si la temperatura de parada se fija muy baja, invertiremos mucho tiempo de búsqueda en las fases finales, que seguramente sería mejor aprovechado en temperaturas superiores. Por el contrario, si la temperatura final se fija en un valor muy alto, es posible que la búsqueda no consiga alcanzar ningún óptimo local.

La condición de parada del algoritmo se puede establecer de diferentes maneras. Una de ellas es fijar un número determinado de valores de T para los cuales se ejecuta el algoritmo. Otro criterio puede ser, establecer un coeficiente de aceptación χ mínimo, de tal forma que en caso de no alcanzarse

se detiene el algoritmo.

En el presente trabajo optamos por comparar la energía de los cambios o movimientos aceptados bajo la siguiente regla:

$$Si \frac{\left(\sum_a^b |\Delta E_k(l) - \Delta E_{k-1}(l)| \right)}{M} < \epsilon \quad a < b, a, b \in \mathbb{N} \text{ y } \beta_c \text{ constante}$$

Se detiene el algoritmo.

Condición de equilibrio (Longitud de la cadena de Markov).

La elección más sencilla para L_k longitud de la k -ésima cadena de Markov, es elegir un valor dependiente (de forma polinómica) del tamaño del problema, así, L_k es independiente de k . Existen propuestas más elaboradas, en las que se establece que para cada valor de T_k se debería realizar un número mínimo de transiciones. Entonces en ese caso se determina L_k de tal forma que el número mínimo de transiciones sean γ_{min} (siendo γ_{min} un número fijo). Sin embargo al aproximarse T a cero, las transiciones son aceptadas cada vez con una probabilidad menor de tal forma que $L_K \rightarrow \infty$ si $T_k \rightarrow 0$.

Por ello L_k debe estar acotado superiormente por una cierta constante \bar{L} , para evitar cadenas de Markov excesivamente largas para bajos valores de T . En esta línea Kirkpatrick [29] toma $\bar{L} = n$ (n el número de variables a resolver), misma opción que se tomó para resolver nuestro problema.

Ley de evolución de la temperatura

El factor más importante en el comportamiento del algoritmo es la regla de enfriamiento o decremento de la temperatura. Existen diferentes leyes de evolución de temperatura que se clasifican en función de su complejidad.

Dentro de las primeras propuestas y que aún son ampliamente usadas tenemos:

$T_k = \alpha T_{k-1}$ (Factor de enfriamiento geométrico, con $\alpha < 1$, y muy cercano, varios autores utilizan $\alpha = 0,95$)

$$T_k = \frac{T_{k-1}}{1+\beta T_{k-1}} \text{ (donde } \beta \approx 0)$$

Capítulo 5

Resultados y Análisis

En el Capítulo 4 se presentó el diseño del algoritmo de segmentación así como algunas consideraciones sobre éste, y en el presente capítulo se muestran los resultados y el análisis de la propuesta.

El trabajo de ésta tesis fue desarrollado utilizando Matlab, implementando primeramente el algoritmo k-means para comparar los resultados de segmentación y posteriormente el diseño del algoritmo de segmentación "Recocido Simulado" para imágenes multiespectrales siguiendo la siguiente metodología:

- Diseño e implementación del algoritmo para imágenes en una banda
- Exploración de los parametros del algoritmo y algunas consideraciones sobre éste.
- Pruebas y selección sobre las consideraciones del algoritmo.
- Diseño, implementación y pruebas del algoritmo para imágenes multi-espectrales.
- Pruebas del algoritmo para imágenes RGB.
- Pruebas del algoritmo para imágenes multiespectrales.

5.1. Exploración del algoritmo en imágenes de una banda

Con el algoritmo “Recocido Simulado” presentado en el capítulo anterior, nos dimos a la tarea de conocer el comportamiento de los parámetros que lo conforman: T_0 , β y Ley de evolución.

El desarrollo del algoritmo se hizo considerando inicialmente los siguientes parámetros para el esquema de enfriamiento:

- Temperatura inicial. T_0 entre $[1,8]$ datos propuestos por diferentes autores.
- Ley de evolución. $T_k = cT_{k-1}$ con c (en nuestro caso $c=0.95$) una constante positiva y menor a uno.
- Criterio de equilibrio. Revisar cada sitio de S .
- Criterio de congelación. Definir un número fijo de k iteraciones.
- Homogeneidad. β fijo entre $[0.05,0.5]$ propuesto por diferentes autores.

Al algoritmo SA con estos parámetros lo denotaremos como “SA simple”.

Para éste primer análisis se utilizaron las imágenes de la Figura 5.1. Al mismo tiempo realizamos el análisis sobre el criterio de congelación del sistema.

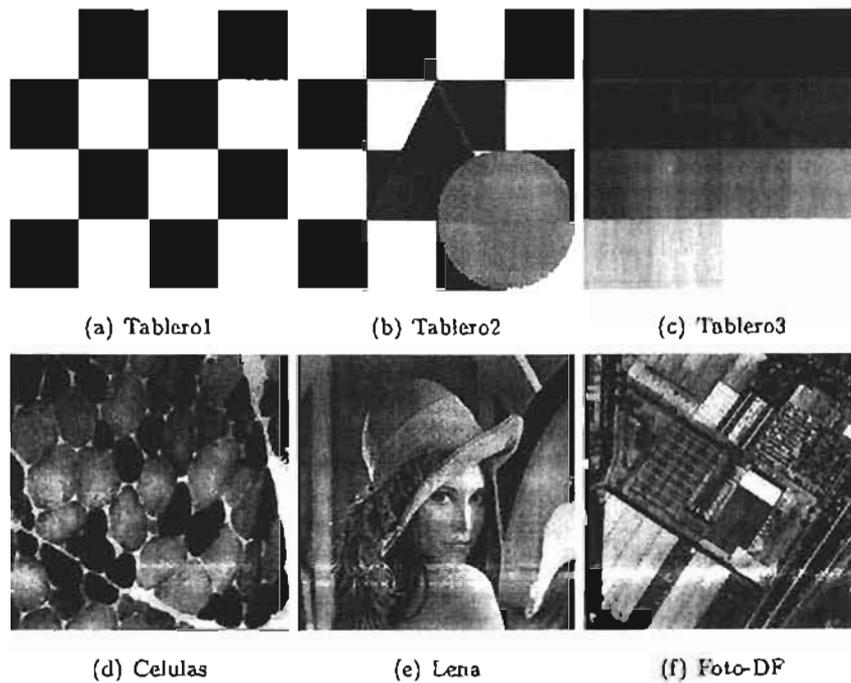


Figura 5.1: Imágenes originales

Recordemos que varios autores utilizan $T_0 = 3$ como temperatura inicial y un número fijo de iteraciones como criterio de congelación, es decir, como criterio para detener el algoritmo.

Comenzaremos explorando el parámetro de la temperatura inicial, variando su valor en 1, 3 y 9 para las diferentes imágenes. Se presentan también los resultados en diferentes iteraciones con la finalidad de ver cualitativamente el comportamiento del algoritmo a través de del tiempo.

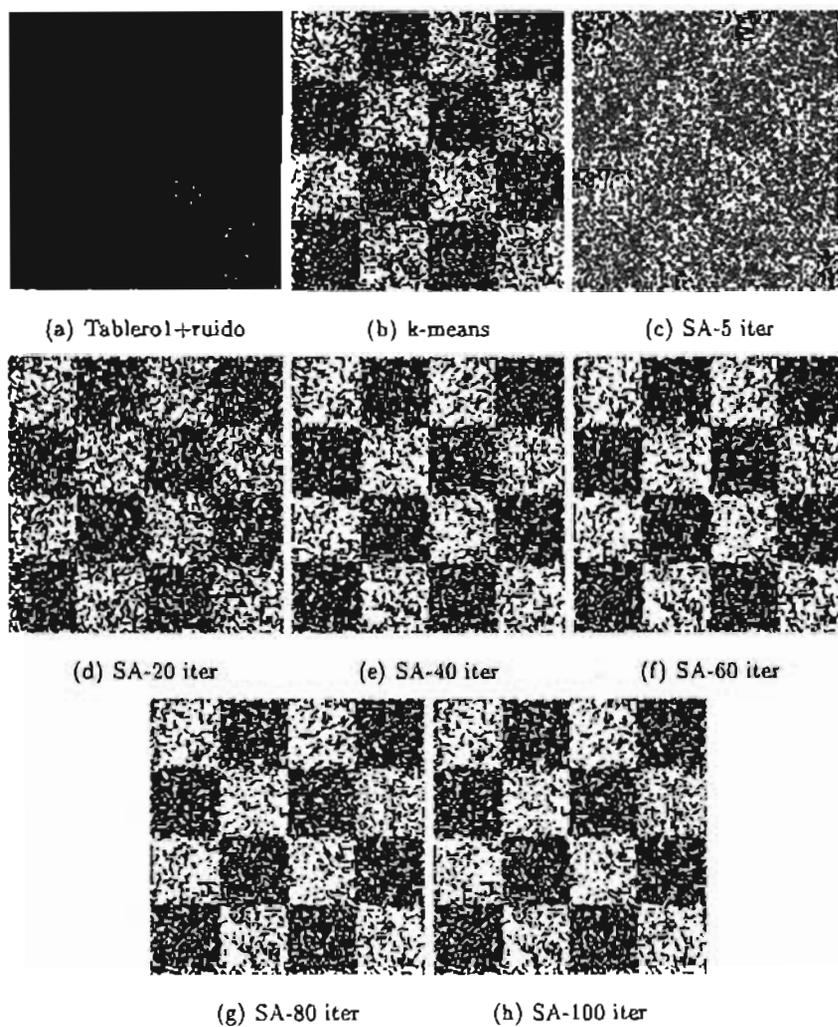


Figura 5.2: Resultados para $T_0 = 1$ y $\beta = .05$

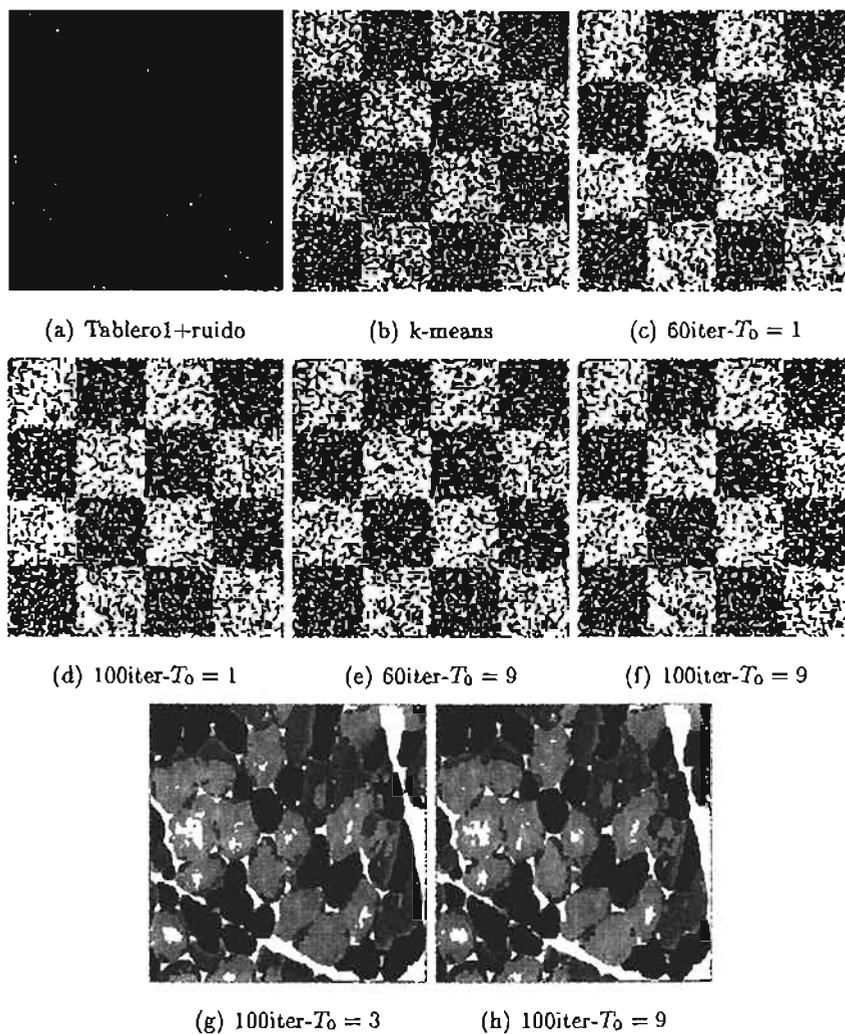


Figura 5.3: Resultados para distintos valores de T_0

En la Figura 5.3 (c,d,e,f,g y h) se muestran los resultados distintos valores de T_0 (1,3 y 9), cualitativamente podemos ver que no hay diferencia entre los resultados aún cuando la diferencia entre los valores de la temperatura son significativos. Con base en estos resultados concluimos que la temperatura inicial no es un parámetro que requiera de un estricto control y decidimos en adelante tomar $T_0 = 3$.

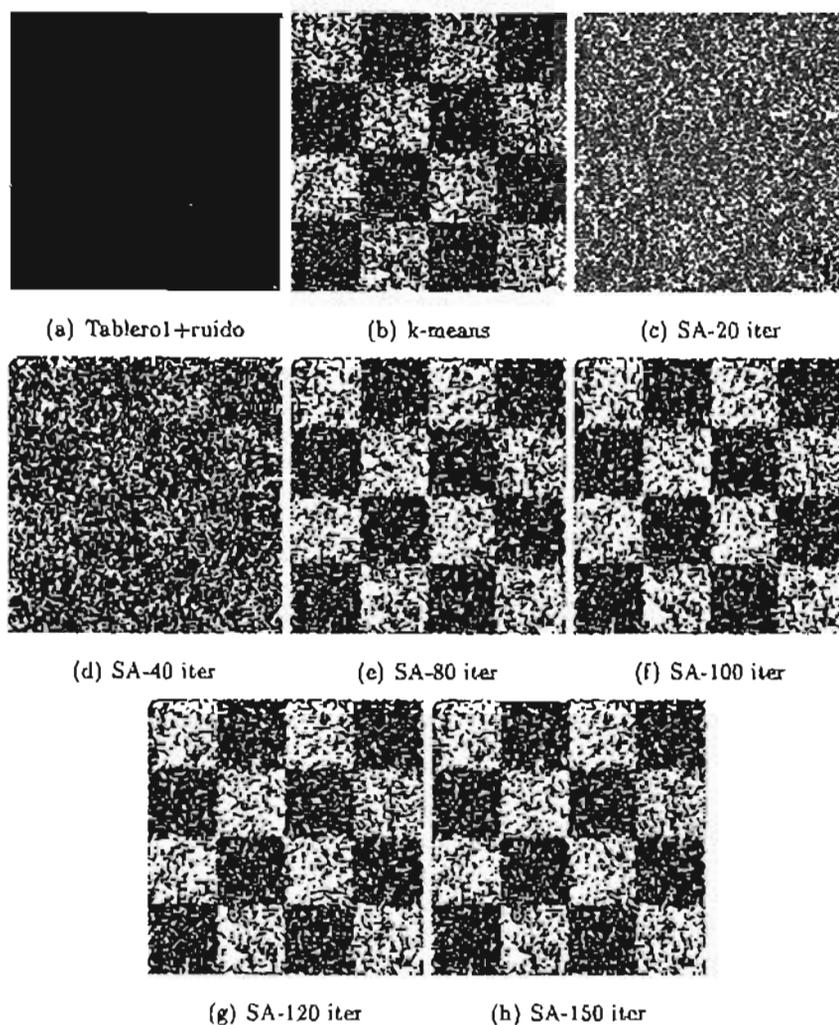


Figura 5.4: Resultados para $T_0 = 3$ y $\beta = .05$

Las Figuras 5.2, 5.3 y 5.4 muestran también los resultados para un número diferente de iteraciones, se observa que no hay cambios entre las iteraciones 80 y 100, o inclusive desde la 60, los que nos hace suponer que los cambios son menores mientras más se avanza en el algoritmo. A continuación graficamos la diferencia consecutiva de la energía y la función de decremento de la temperatura, con la finalidad de ver el comportamiento en ambas.

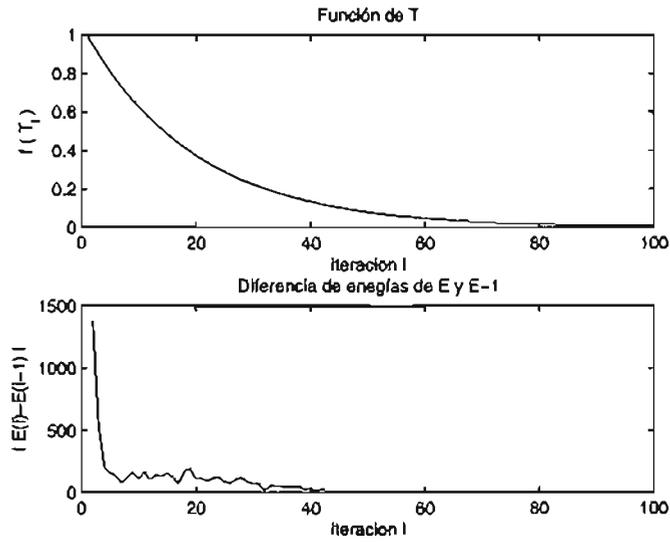


Figura 5.5: Gráfica correspondiente a $T_0 = 1$ y $\beta = .05$

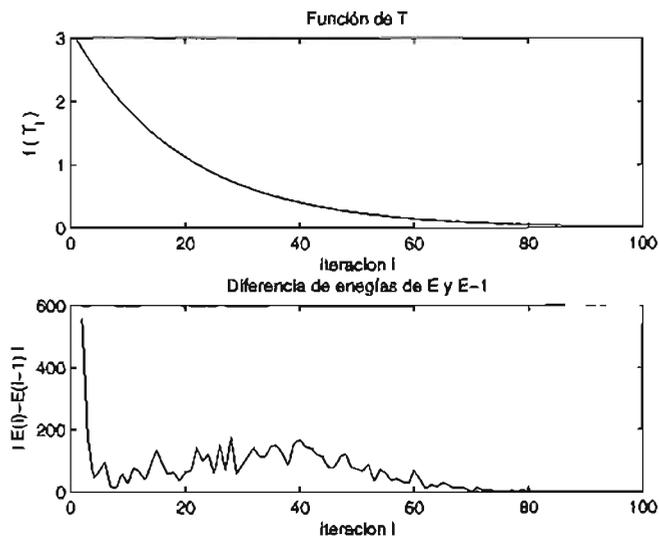


Figura 5.6: Gráfica correspondiente a $T_0 = 3$ y $\beta = .05$

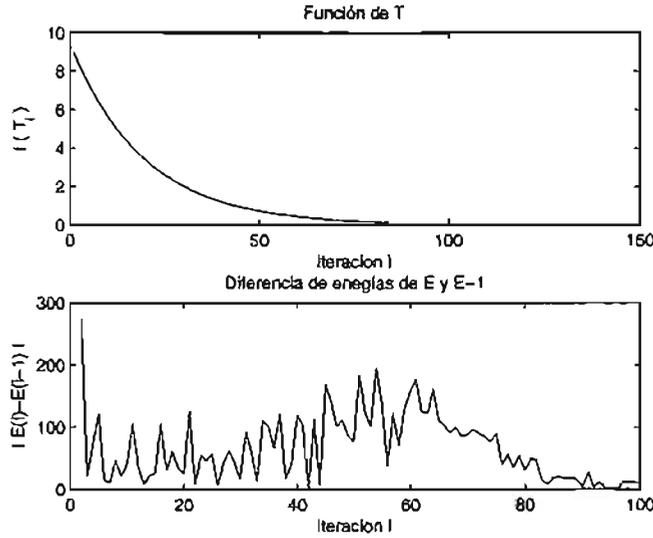


Figura 5.7: Gráfica correspondiente a $T_0 = 9$ y $\beta = .05$

Con los resultados mostrados en las Figuras 5.5, 5.6 y 5.7 podemos suponer que no es necesario que el criterio de congelación sea un número fijo de iteraciones o decrementos de la temperatura pues a partir de cierto número de iteraciones la $P(l_k = l_{k-1}) \approx 1$, es decir, en la práctica la búsqueda converge por lo general a su óptimo local final antes del valor de la temperatura definido como criterio de congelación.

Por ello nuestra primera consideración es sobre el criterio de congelación:

$$\text{Parar si } \frac{\left(\sum_a^b |\Delta E_k(l) - \Delta E_{k-1}(l)| \right)}{M} < \epsilon \quad a < b, a, b \in \mathbb{N} \text{ y } \beta_c \text{ constante}$$

También observamos que el valor de la temperatura inicial no representa un riesgo importante en su control, por lo que supondremos en adelante el mismo criterio que diferentes autores [3, 17], y tomamos a $T_0 = 3$ u otro valor cercano, como el valor para el resto del trabajo.

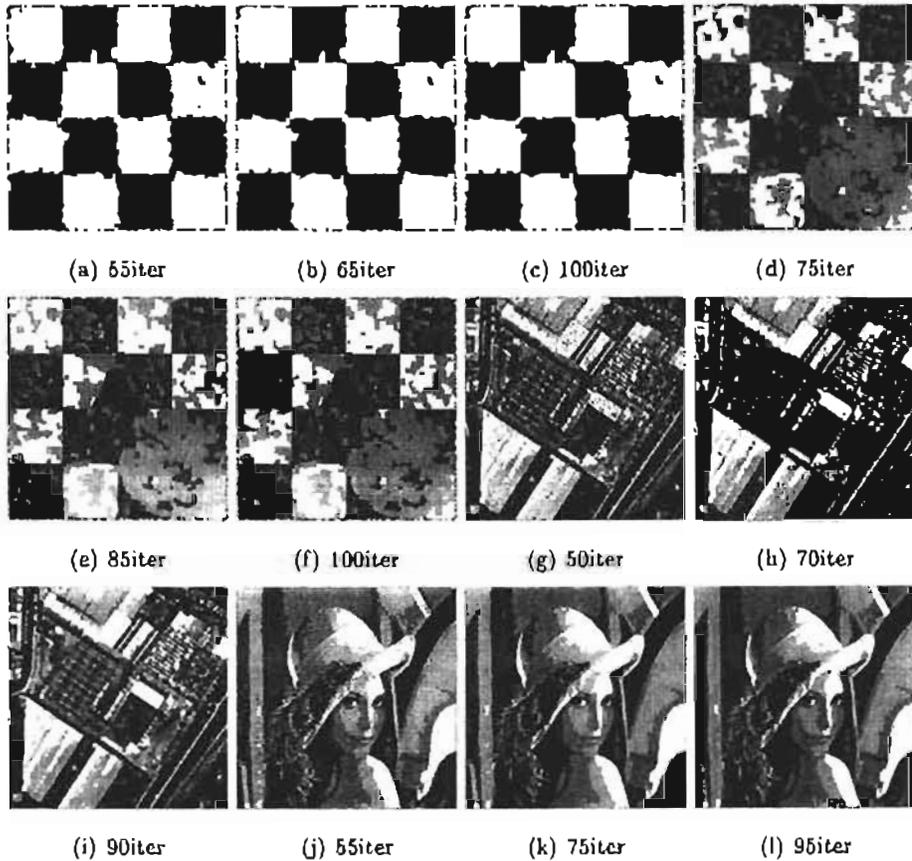


Figura 5.8: Ejemplos de convergencia.

De la Figura 5.8: b, e, h, k, convergen según el criterio de congelación definido.

Ahora realizamos el análisis sobre el parámetro β , la Figura 5.9 muestra el comportamiento del parámetro, éste controla la homogeneidad de las regiones en la que se esta segmentando, mientras más pequeño es el valor, los pixéles vecinos se consideran en menor grado, y mientras mayor es el parámetro los pixéles vecinos ganan importancia.

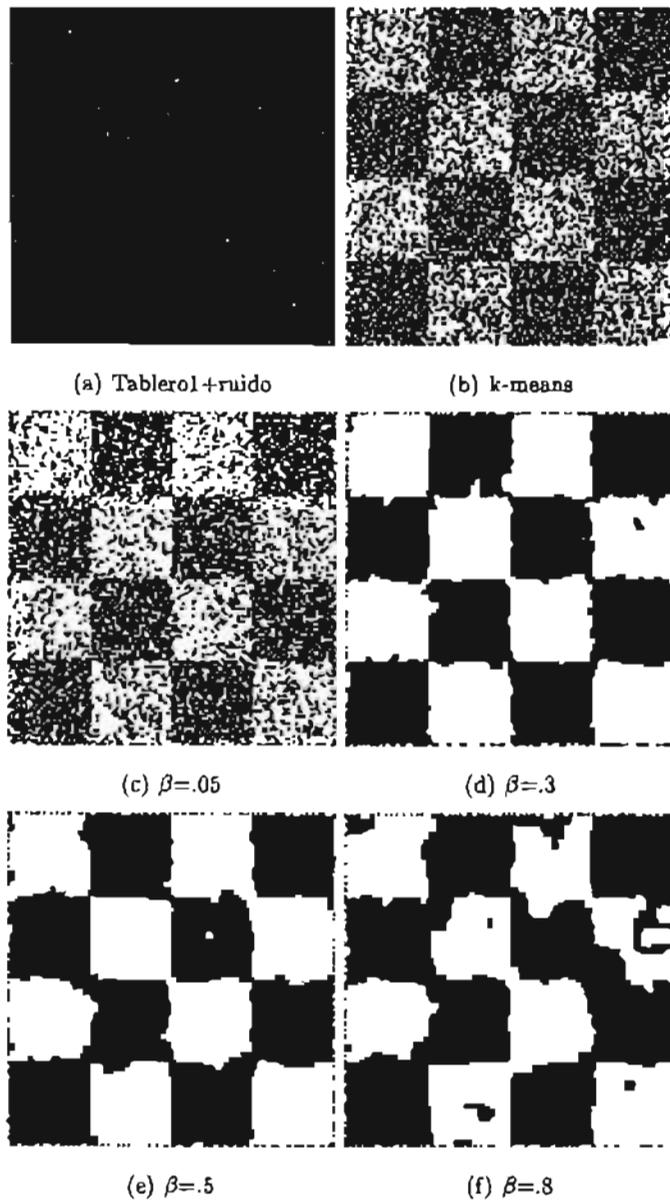


Figura 5.9: Variaciones del parámetro β

Se realizaron pruebas para diferentes parámetros de β en diferentes imá-

genes, oscilando estos valores entre $[0.01, .5]$, observando el comportamiento para los diferentes valores nos percatamos que sería deseable que el parámetro β cambiara conforme decrece el parámetro de la temperatura, algunos autores [35, 17] mencionan esta posibilidad.

Se probaron los siguientes esquemas de enfriamiento:

- Esquema 1 (geométrico)

$$T_i = kT_{i-1}$$

$$\beta = c$$

Con c entre $[.05, .5]$ fue donde se observaron mejores resultados, tomemos $c=.3$ y k una constante cercana a uno ($k=.95$).

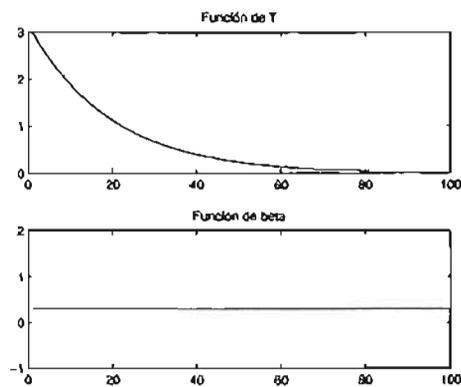


Figura 5.10: Comportamiento del Esquema 1

- Esquema 2

$$\beta_i = \beta_{\min} \left(\frac{\beta_{\max}}{\beta_{\min}} \right)^{\frac{i}{N}}$$
$$T_i = (1/\beta_i) * c$$

La constante c se proporcionó con la finalidad de hacer que la temperatura comenzara con los valores de 3 o 1, ya que de acuerdo a nuestras pruebas fueron los mejores valores.

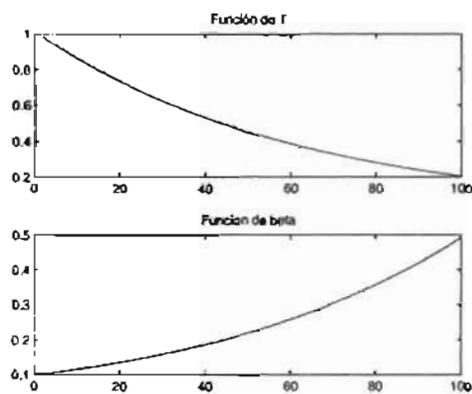


Figura 5.11: Comportamiento del Esquema 2

- Esquema 3

$$\beta_i = \beta_0 * \log(i + 1)$$
$$T_i = \frac{1}{\beta_i}$$

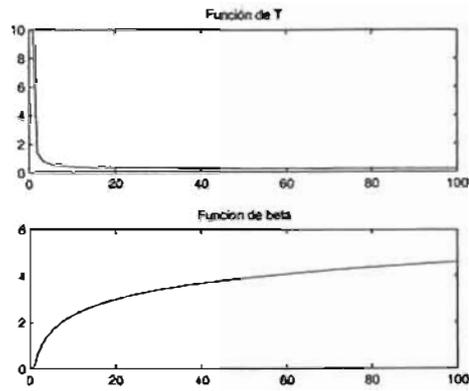


Figura 5.12: Comportamiento del Esquema 3

- Esquema 4

$$T_i = \frac{T_{i-1}}{(1 + .1 * T_{i-1})}$$

$$\beta_i = \frac{1}{T_i}$$

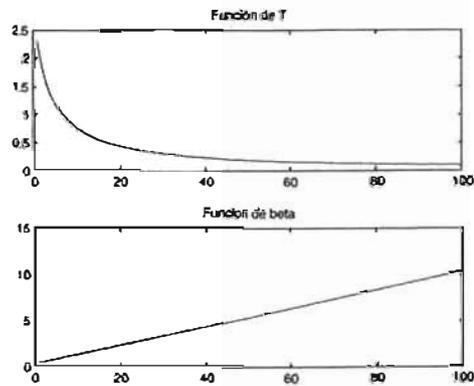


Figura 5.13: Comportamiento del Esquema 4

- Esquema 5

$$T_i = T_{max} - i \frac{(T_{max} - T_{min})}{N}$$
$$\beta = \frac{1}{T_i}$$

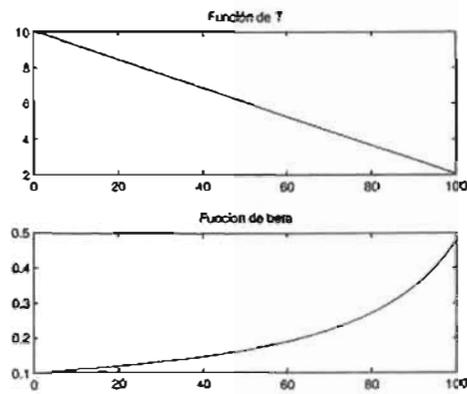


Figura 5.14: Comportamiento del Esquema 5

La Figura 5.15 muestra los resultados de los esquemas de enfriamiento antes mencionados, se puede percibir una similitud entre los resultados de los esquemas 1 y 2, por lo que se continua el análisis sólo con los dos primeros esquemas de enfriamiento.

Para los esquemas 3 y 4 se observa que si se escoge una función de decremento de la temperatura que con pocas iteraciones se aproxime a cero, muy rápido se van definiendo las regiones de la segmentación, sin embargo estas no son correctas (figuras e y g respectivamente).

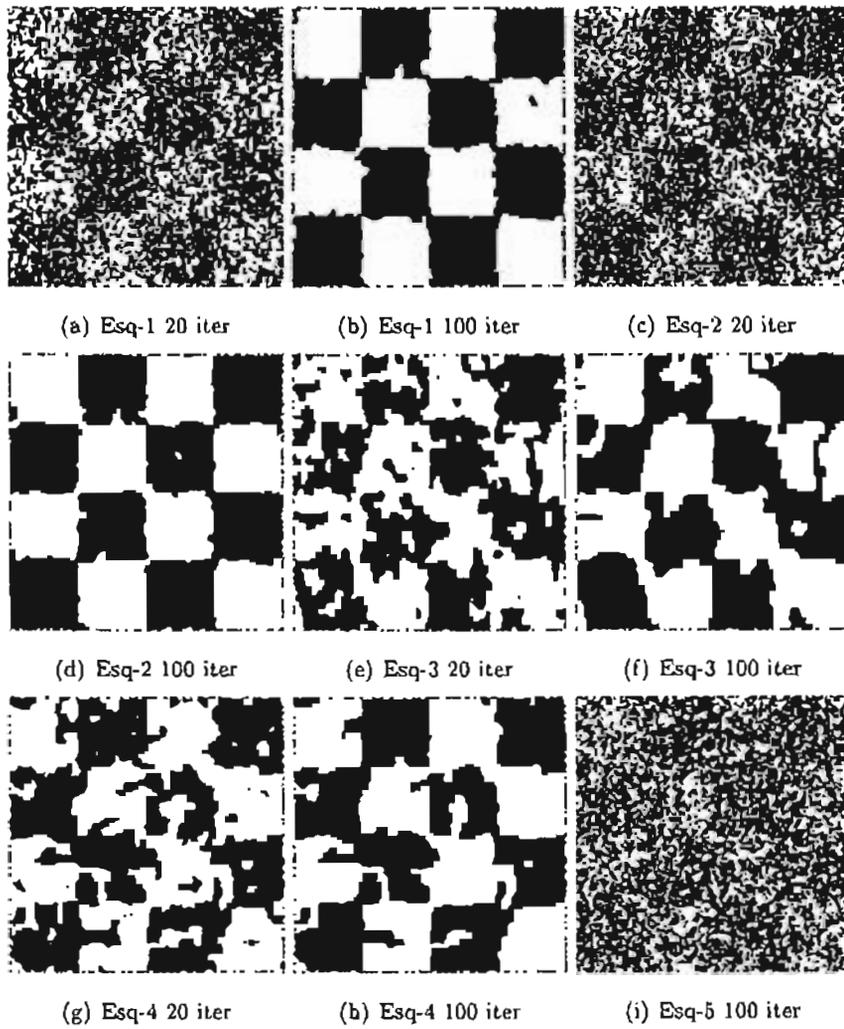
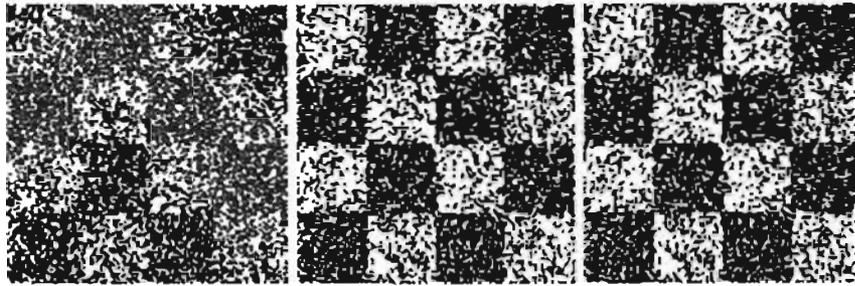


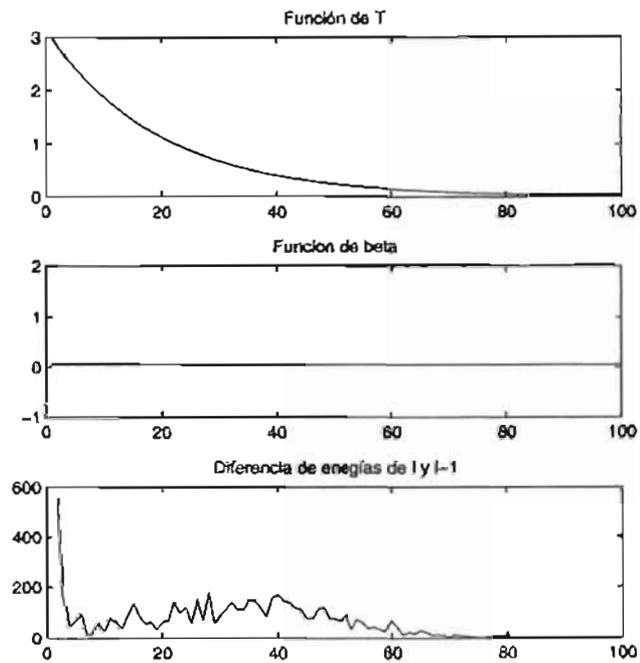
Figura 5.15: Resultados de los esquemas de enfriamiento



(a) 35 iter

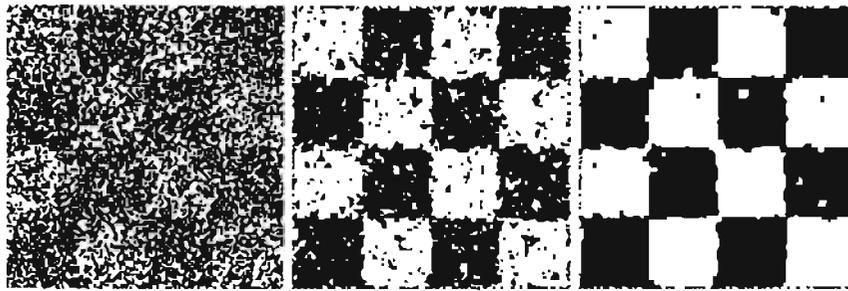
(b) 70 iter

(c) 100 iter



(d) Funciones

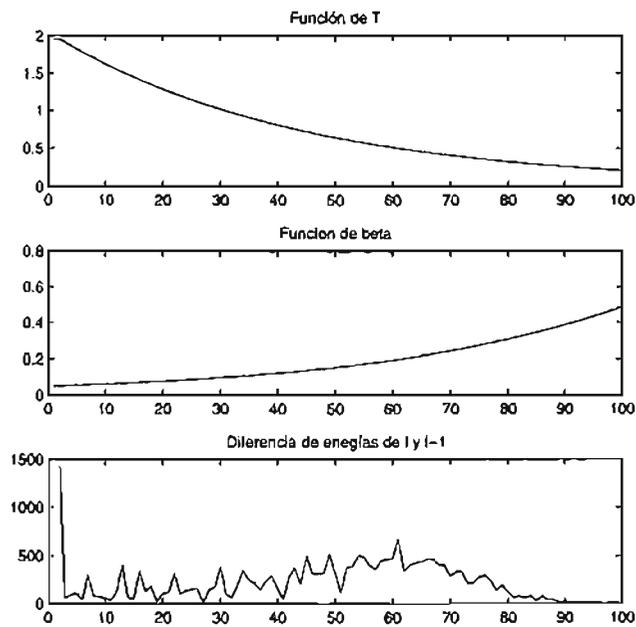
Figura 5.16: Esquema 1 $\beta=.05$



(a) 35 iter

(b) 70 iter

(c) 100 iter



(d) Funciones

Figura 5.17: Resultados para el tablero en el esquema 2

Claramente en las Figuras 5.16 (c) y 5.17 (c) vemos que el Esquema 2 proporciona un mejor resultado.

5.2. Análisis del algoritmo en imágenes a color

Ahora aplicamos el algoritmo del Capítulo 4 en imágenes a color para corroborar su buen comportamiento y la elección del esquema de enfriamiento, es decir, que el presente algoritmo es una buena opción para la segmentación de imágenes multiespectrales.

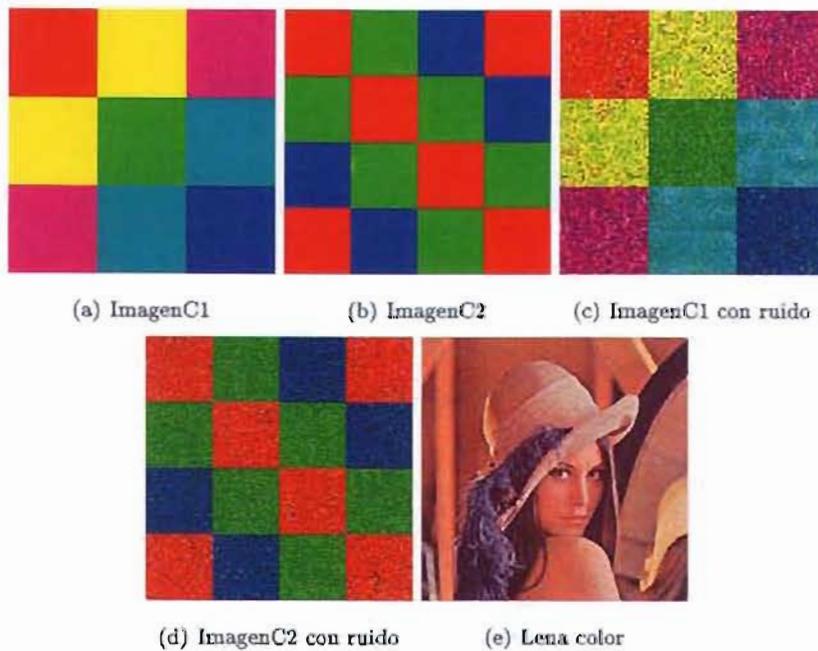


Figura 5.18: Imágenes originales

La Figura 5.18, muestra las imágenes originales empleadas en esta sección así como las imágenes con ruido, a éstas se les agregó un ruido gaussiano al 30 %.

Utilizamos la matriz de confusión y el coeficiente kapa $\tilde{\kappa}$ como indicador para evaluar objetivamente el algoritmo y el algoritmo k-means como el algoritmo a comparar.

La matriz de confusión, es una matriz cuadrada con dimensión igual al número de etiquetas o regiones en las que hay que segmentar. El elemento C_{ij} indica el número de pixeles clasificados de la clase i que fueron clasificados erróneamente en la clase j .

En una segmentación perfecta, los elementos de la diagonal serán diferentes de 0. El coeficiente kappa, es un indicador global de la calidad de la segmentación, dicho coeficiente varía en el intervalo $[-1,1]$.

La segmentación perfecta sería cuando $\tilde{\kappa} = 1$ y una totalmente errónea cuando $\tilde{\kappa} = -1$. El coeficiente kappa se define como:

$$\tilde{\kappa} = \frac{d - q}{N - q} \quad (5.2.1)$$

d : suma de los elementos de la diagonal de la matriz de confusión

N : número de pixeles de la imagen

n : número de etiquetas

q : $q = \frac{\sum_{i=1}^n \sum_{j=1}^n C_{ij}}{N}$

Resultados para la ImagenC1 con ruido.

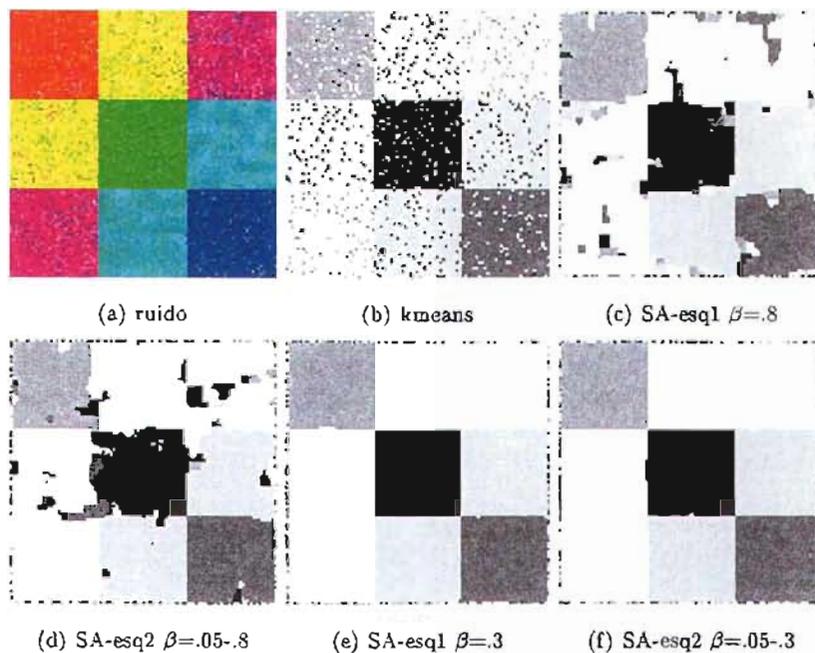


Figura 5.19: ImagenC1

$\kappa = 0,8814$	clase1	clase2	clase3	clase4	clase5	clase6
clase1	923	7	2	52	0	40
clase2	5	946	5	37	30	1
clase3	8	4	930	0	44	38
clase4	82	93	1	1853	13	6
clase5	1	79	98	9	1847	14
clase6	109	0	100	8	9	1822

Cuadro 5.1: Matriz de confusión para la Figura 5.19 (b)

$\kappa = 0,8569$	clase1	clase2	clase3	clase4	clase5	clase6
clase1	922	18	32	16	1	35
clase2	21	847	41	75	30	10
clase3	7	0	978	0	0	39
clase4	22	6	29	1905	49	37
clase5	34	145	39	32	1720	78
clase6	35	69	57	83	38	1766

Cuadro 5.2: Matriz de confusión para la Figura 5.19 (c)

$\kappa = 0,8837$	clase1	clase2	clase3	clase4	clase5	clase6
clase1	920	45	0	4	26	29
clase2	13	991	4	12	0	4
clase3	38	14	880	0	16	76
clase4	59	3	13	1887	58	28
clase5	50	74	22	24	1843	35
clase6	69	95	5	15	45	1819

Cuadro 5.3: Matriz de confusión para la Figura 5.19 (d)

$\kappa = 0,9984$	clase1	clase2	clase3	clase4	clase5	clase6
clase1	1024	0	0	0	0	0
clase2	0	1022	0	1	1	0
clase3	0	0	1018	0	0	6
clase4	0	1	0	2046	0	1
clase5	0	0	0	2	2046	0
clase6	0	0	0	0	0	2048

Cuadro 5.4: Matriz de confusión para la Figura 5.19 (e)

$\kappa = 0,9979$	clase1	clase2	clase3	clase4	clase5	clase6
clase1	1022	0	0	2	0	0
clase2	0	1024	0	0	0	0
clase3	0	0	1024	0	0	0
clase4	0	0	0	2045	3	0
clase5	0	0	0	0	2048	0
clase6	8	0	0	0	3	2037

Cuadro 5.5: Matriz de confusión para la Figura 5.19 (f)

kmeans	SA-esq1 $\beta = .8$	SA-esq2 $\beta = .05-.8$	SA-esq1 $\beta = .3$	SA-esq2 $\beta = .05-.3$
0.8814	0.8569	0.8837	0.9984	0.9979

Cuadro 5.6: Índices kapa de los resultados de ImagenC1

Aunque podemos observar que para $\beta = .3$ el coeficiente $\tilde{\kappa}$ señala que el resultado dado por SA-esq1 es mejor que el dado por SA-esq2, observamos también que para $\beta = .8$, SA-esq2 es mejor que SA-esq1 y que la diferencia entre el esquema 1 y 2 para $\beta = .3$ es mucho menor que para $\beta = .8$.

En este ejercicio comprobamos que se aprovecha en los comienzos del esquema de enfriamiento los valores pequeños de β e ir incrementando β conforme el sistema evoluciona.

Resultados para la ImagenC2 con ruido.

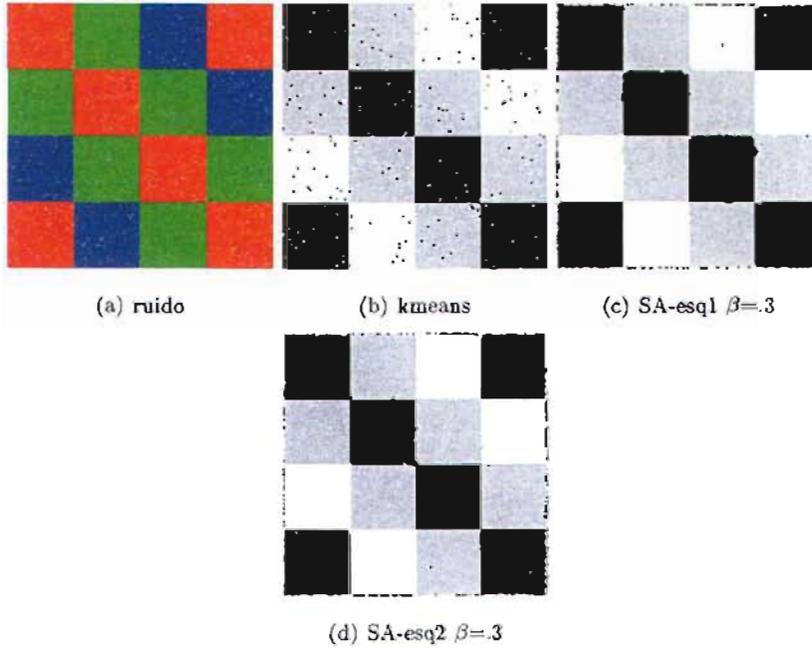


Figura 5.20: ImagenCI

$\kappa = 0,9755$	clase1	clase2	clase3
clase1	6060	45	39
clase2	60	6033	51
clase3	36	33	4027

Cuadro 5.7: Matriz de confusión para la Figura 5.19 (b)

$\kappa = 0,9940$	clase1	clase2	clase3
clase1	6121	18	5
clase2	20	6116	8
clase3	3	11	4082

Cuadro 5.8: Matriz de confusión para la Figura 5.19 (c)

$\kappa = 0,9975$	clase1	clase2	clase3
clase1	6133	5	6
clase2	12	6131	1
clase3	1	2	4093

Cuadro 5.9: Matriz de confusión para la Figura 5.19 (d)

kmeans	SA-esq1 $\beta = .3$	SA-esq2 $\beta = .3$
0.9755	0.9940	0.9975

Cuadro 5.10: Índices kapa de los resultados de ImagenC1

Corroboramos que el esquema de enfriamiento propuesto, ayuda a obtener resultados más estables sin la necesidad de probar valores de β y T_0 como lo señalan algunos autores que lo han empleado en imágenes monocromáticas.

5.3. Comprobación y aplicación en imágenes Landsat

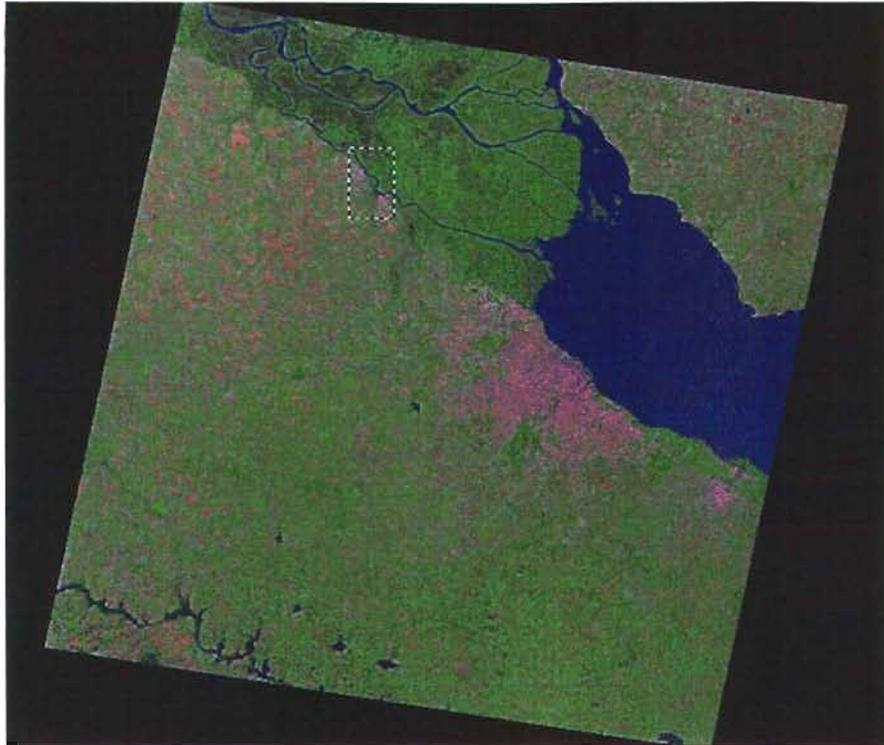


Figura 5.21: Imagen Landsat 7 ETM. Orbita: 225/084, bandas 7-4-2

La Figura 5.21 es una imagen Landsat 7, tomada en el 2000, la zona de color violeta más grande cerca del agua es Buenos Aires Argentina, se tomó sólo un segmento de la imagen para realizar las pruebas del algoritmo que corresponde a una región de la provincia de Buenos Aires. La Figura 5.22 muestra el área seleccionada. En esta imagen observamos claramente dos



Figura 5.22: Área seleccionada de la Figura 5.21

zonas urbanas y el rio Paraná.

Los resultados para la imagen de satélite que se muestran a continuación sólo los podemos evaluar de manera cualitativa ya que se desconoce la segmentación verdadera, este es uno de los grandes problemas del proceso de segmentación.

En ocasiones, para validar o apoyar el resultado de un segmentación se recurre a verificación en sitio o al conocimiento de algún experto que apoye en la interpretación del mapa temático generado.

Resultados

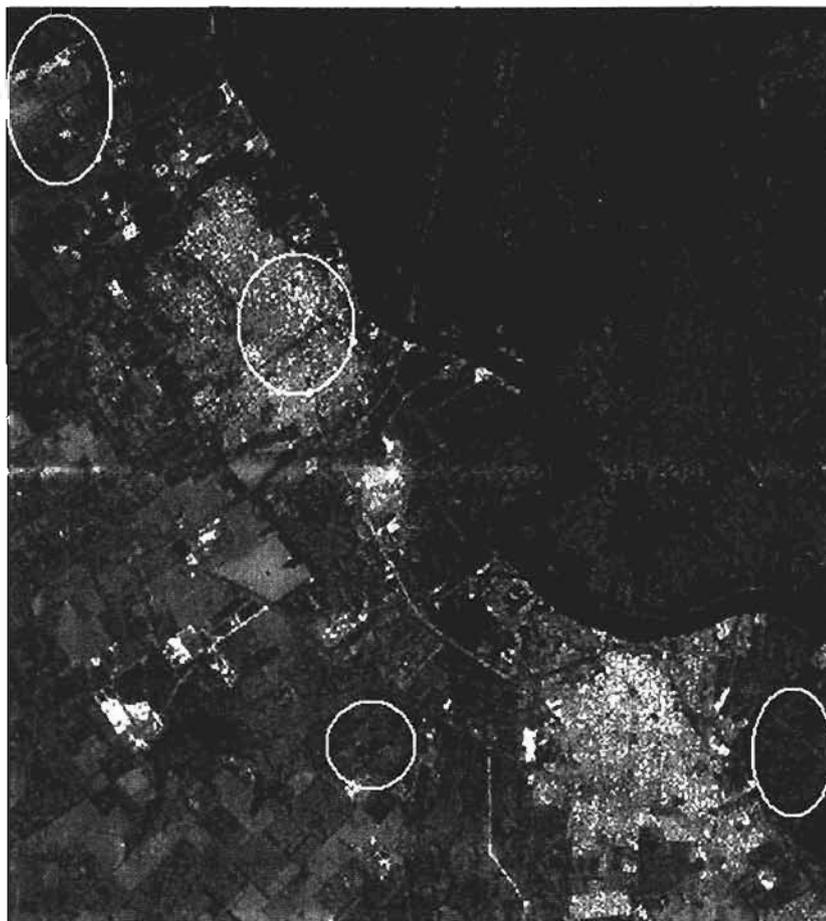


Figura 5.23: k-means - 8 clases.

Los círculos señalan regiones donde el algoritmo k-means no genera regiones homogéneas, como es el caso de la mancha urbana.

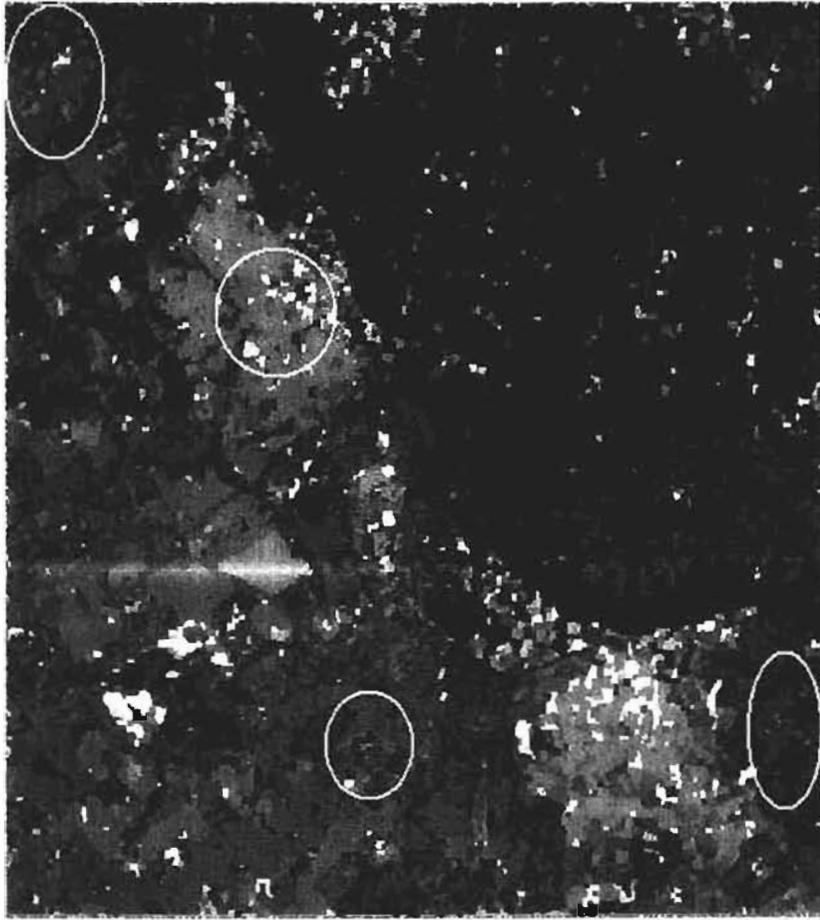


Figura 5.24: SA esquema 1 -8 clases- $\beta=.2$, $T_0=3$

En el circulo superior izquierdo observamos que la regiones generados por el esquema 1 no quedan bien definidas.

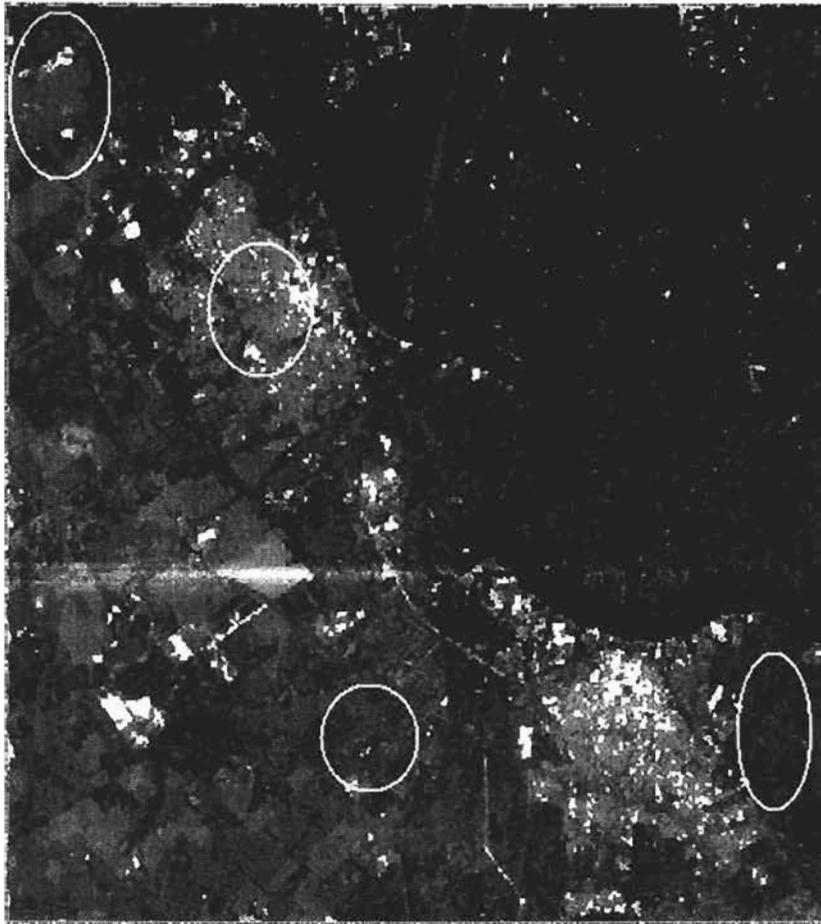


Figura 5.25: SA esquema 2 -8 clases- $\beta=[.01, .2]$, $T_0=3$

Para el esquema 2 observamos que las regiones quedan mejor definidas, en el circulo superior izquierdo y en el del centro es más claro.

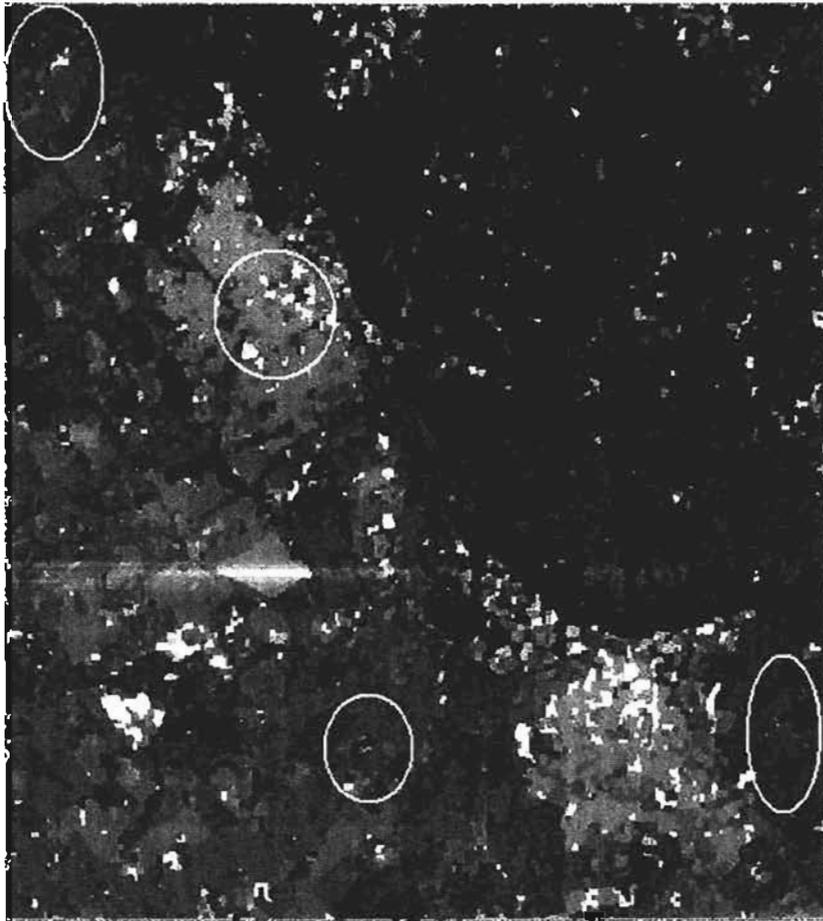


Figura 5.26: SA esquema 1 -8 clases- $\beta=.3$, $T_0=3$

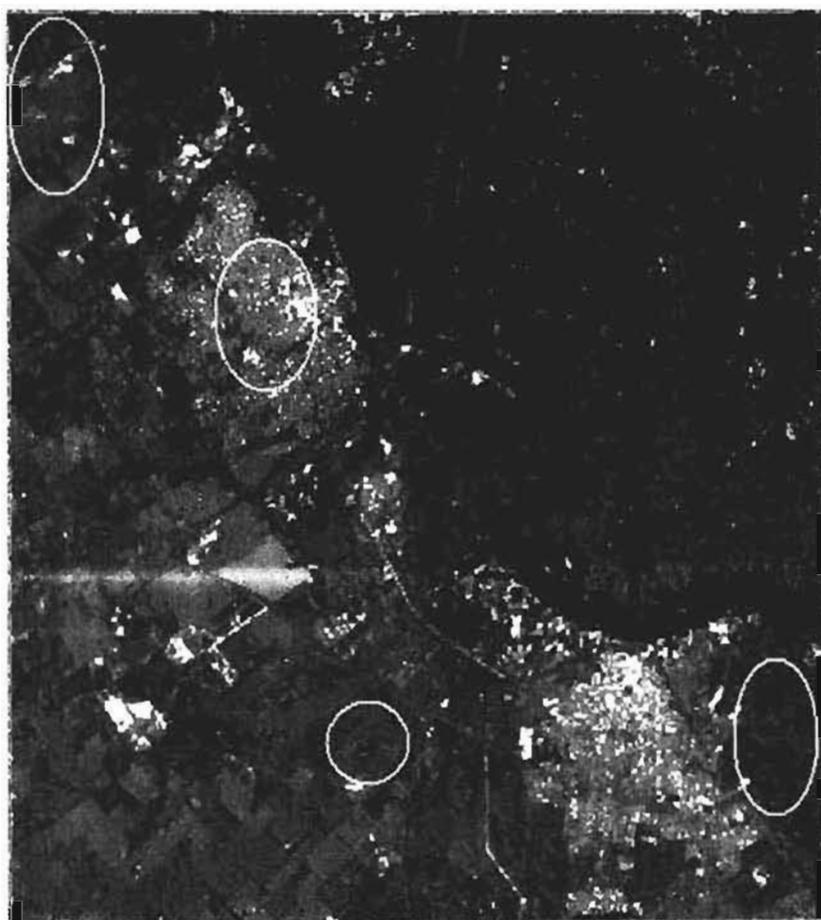


Figura 5.27: SA esquema 2 -8 clases- $\beta=[.01,.3]$, $T_0=3$

Las Figuras 5.26 y 5.27 son resultados incrementando el valor de β , es decir, tratando de crear regiones más homogéneas, bajo el esquema 1, podemos observar que los contornos de lo que pudiera ser regiones empeoran, y que bajo el esquema 2 esto no sucede, por el contrario, genera regiones más homogéneas afectando muy poco el contorno de dichas regiones. Por lo que concluimos que el Esquema-2 es un mejor esquema y que no requiere de estar probando los parámetros del algoritmo.

Capítulo 6

Conclusiones

Con el avance tecnológico en la adquisición de imágenes multiespectrales y el creciente número de bandas que componen a éstas, surge la necesidad de desarrollar métodos de segmentación que utilicen la información espectral y espacial de algunas o todas las bandas.

Los resultados muestran que:

1. La Teoría de los Campos Aleatorios de Markov son una buena elección para modelar dependencias contextuales y que la aplicación de éstos en los problemas de segmentación de imágenes multiespectrales, resulta una buena opción.
2. El modelo de segmentación aquí presentado, comprende la posibilidad de trabajar con varias bandas de la imagen multiespectral. El algoritmo “Recocido Simulado” teóricamente, converge a un óptimo global por lo que lo hace un algoritmo muy atractivo, ya que aunque dicha convergencia no puede ser garantizada el algoritmo tiene la capacidad de escapar de mínimos locales gracias al criterio Metrópolis.

3. Las consideraciones hechas sobre el algoritmo "Recocido Simulado", minimizan la dificultad en la selección de parámetros (ley de evolución, criterio de congelación, etc.), y facilitan la implementación del algoritmo. En este sentido dejamos sin explorar varias consideraciones, tanto en los parámetros específicos como en los genéricos, en Marc et al. [35] se mencionan diferentes consideraciones al respecto, algunas opciones de paralelismo se exponen en [26], donde la selección de cliques de menor orden facilitan dicha tarea.

Por último algunas consideraciones importantes sobre el algoritmo son:

- La segmentación considerando la información contextual, puede llevar a la pérdida de detalle.
- La medición objetiva depende de el nivel de agregación de información que se desea en la segmentación. Dado que, en la práctica la segmentación verdadera de la imagen es desconocida, la medición se aplica en imágenes sintéticas donde esta segmentación verdadera es conocida.
- El tiempo computacional empleado en este algoritmo es alto.
- Aunque teóricamente el algoritmo converge a un óptimo global, es evidente que el número de iteraciones sólo puede ser finito, por lo que en la práctica dicha convergencia no puede ser garantizada.

Bibliografía

- [1] Abend K., Harley T.J and Kanal L. N. Classification of binary random patterns. IEEE Transactions on Information Theory.11(4).1965.
- [2] Ballard D.H.. Reference frames for animate vision. International joint conference on artificial intelligence: Proceedings, Morgan Kauffmann Publishers. 1989.
- [3] Besag, J.. Spatial interaction and the statistical analysis of lattice systems (with discussions). Journal of the Royal Statistical Society, Series B,36:192-236. 1974.
- [4] Besag, J.. Statistical analysis of non-lattice data. The Statistician. 1975.
- [5] Besag, J.. Towards bayesian image analysis. Journal of Applied Statistics. 1989.
- [6] Bouman C. A. and Shapiro, A multiscale random field model for Bayesian Image Segmentation. IEEE Trasaction on Image Processing, Vol. 3, No. 2, pp. 162-177, 1994.
- [7] Bow S.T. Pattern recognition and image preprocessing. Marcel Dekker Inc. New York. 1992.
- [8] Campadelli P.,Medici D., Schettini. Color image segmentation using Hopfield networks. Image and Vision Computing Image. Vol 5 No 3. 1997.

- [9] Cerny V.. A Thermodynamical approach to the travelling salesman problem: An efficient simulation algorithm. Optimization theory and applications. 1985.
- [10] Chellappa, R.. Two-dimensional discrete gaussian Markov random field models for image processing. In Kanal, L. N. and Rosenfeld, A., editors, Progress in Pattern Recognition 2. 1985.
- [11] Chellappa, R. and Jain, A.. Markov Random Fields: Theory and Applications. Academic Press. 1993.
- [12] Crespo J., Serra, J., Schafer, R.W. Graph-based morphological filtering and segmentation. Preprints of the VI Spanish Symposium on pattern Recognition And Image Analysis. Cordoba, España. 1995.
- [13] Dzung L. Pham, Chenyang Xu and Jerry L. Prince. A survey of current methods in medical image segmentation. Technical report JHU/ECE. 1998.
- [14] Elliott, H., Derin, H., Cristi, R., and Geman, D.. Application of the Gibbs distribution to image segmentation. In Proceedings of the International Conference on Acoustic, Speech and Signal Processing. 1984.
- [15] Elvira Fernández. Información por todo lo alto. <http://www.cienciadigital.net/mayo2001/teledeteccion.html> . 2001.
- [16] Fu K.S., Mui J.K. A Survey on Image Segmentation. Pattern recognition. Vol 13. 1981
- [17] Geman D. and Geman S. Stochastic Reflexation, Gibbs distributions and the bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machines Intelligence. 1984.
- [18] Gidas, B.. A renormalization group approach to image processing problems. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989.

- [19] Gonzalez R.C., Wintz P. Digital Image Processing. 2nd Ed. Addison-Wesley. 1987.
- [20] Hammersley, J. M. and Clifford, P.. Markov field on finite graphs and lattices. 1971.
- [21] Haralick, R. M.. Decision making in context. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1983.
- [22] Haralick R.M., Shapiro L.G. Image Segmentation Techniques. Computer Vision, Graphics, and Image Processing. Vol 29 No.1. 1985.
- [23] Jain A. and Dubes R., Algorithms for Clustering Data. Prentice Hall. 1988.
- [24] Jain A.K., Kailath, T. Fundamentals of digital image processing. Prentice-Hall information and systems sciences series. Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [25] Kato Z., Zerubia J. and Berthod M., Bayesian image classification using Markov Random Field. 1993.
- [26] Kato Z., Modelesations Markoviennes Multiresolutions en Vision par ordinateur. Application a la segmentation d images SPOT, PhD. Thesis, University of Nice, 1994.
- [27] Keneth R. Castleman. Digital Image Processing. Prentice Hall. 1996.
- [28] Kindermann, R. and Snell, J. L. . Markov Random Fields and Their Applications. American Mathematical Society, Providence, R.I. (1980)
- [29] Kirkpatrick, S., Gelatt C. D. Jr., Vecchi M. P.. Optimization by Simulated Annealing. Science. 1983.
- [30] Knill D.C. Perception of surface contours and surface shape: from computation to psychophysics. J. Opt. Soc. Am. A, Vol. 9, No. 9. 1992.

- [31] Li, S. Z.. A Markov random field model for object matching under contextual constraints. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1994.
- [32] Li Stan Z.. Markov random field modeling in computer vision. Springer-Verlag. 1995.
- [33] Lira Chavez Jorge. La percepción remota: nuestros ojos desde el espacio. Colección La Ciencia desde México, vol. 33, Fondo de Cultura Económica, México, 1987.
- [34] Lira Chavez Jorge. Introducción al tratamiento digital de imágenes. Fondo de Cultura Económica, México, 2002.
- [35] Marc C. Robini, Simulated annealing, Acceleration techniques, and image restoration, IEEE Transaction on Image Processing, Vol. 8 No. 10 1999.
- [36] Marroquin, J., Mitter, S., and Poggio, T.. Probabilistic solution of ill-posed problems in computational vision. Journal of the American Statistical Association. 1987.
- [37] Moussouris, J. Gibbs and Markov systems with constraints. Journal of statistical physics, 10:11-33.(1974)
- [38] Murray, D. and Buxton, B.. Scene segmentation from visual motion using global optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1987.
- [39] Nikolaidis N. and Pitas I. 3-D Image Processing Algorithms. John Wiley and Sons Inc. 2001.
- [40] Pal N.R., Pal S.K. A Review on image segmentation techniques. Pattern Recognition. Vol 26 No 9. 1983.
- [41] Pavlidis T. A critical survey of image analysis methods. In ICPR, pages 502-511. (1986)

- [42] Pitas I. Digital Image Processing Algorithms. Prentice Hall. 1993.
- [43] Rosenfeld and A. Kak. Digital Picture Processing. Vol 2. 2nd Ed. Academic Press. 1982
- [44] Sánchez Martín Nidia. Panorama actual de las técnicas mixtas de clasificación de imágenes mediante segmentación espectral y por texturas. Aplicación a las imágenes de alta resolución espacial. http://www.mappingiteractivo.com/plantilla-ante.asp?id_articulo=288, 2003.
- [45] Shi J., Belongie S., Leung T. and Malik J. Image and Video Segmentation: The Normalized Cut Framework. Proc. of 1998 Int'l Conf. on Image Processing (ICIP'98), Vol 1. 1998.
- [46] Sonka M., Hlavá V. and Boyle R. Image Processing, Analysis and Machine Vision. Chapman and Hall Computing. 1998.
- [47] Suetens P., Fua, P., Hanson, A.J. Computational Strategies for Object Recognition. ACM Computing Surveys. Vol. 24, No. 1. 1992.
- [48] Won, C. S. and Derin, H. Unsupervised segmentation of noisy and textured images y using Markov random fields. CVGIP: Graphics Model and Image Processing. 1992.
- [49] Wu J., Yan H. and Chalmers A.N. Color Segmentation Using Fuzzy Clustering and Supervised Learning. Journal of Electronic Imaging. 1994.