

UNIVERSIDAD NACIONAL AUTONOMA DE MÉXICO.



DIVISIÓN DE ESTUDIOS DE POSTGRADO, FACULTAD DE  
INGENIERIA.

ESQUEMA DE SEGMENTACIÓN DE IMÁGENES  
DE PERCEPCIÓN REMOTA: APROXIMACIÓN  
CONTEXTUAL Y DE REGULARIZACIÓN.

TESIS

QUE PARA OBTENER EL TÍTULO DE:  
**MAESTRO EN INGENIERIA EN TELECOMUNICACIONES**  
PRESENTA:  
MARCO ANTONIO ENCINAS CASTRO.

DIRECTOR DE TESIS:  
D.I. MIGUEL MOCTEZUMA FLORES.



México, D.F.

Septiembre 2005





Universidad Nacional  
Autónoma de México



**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 **División de Estudios de Postgrado de la Facultad de Ingeniería, de la Universidad Nacional Autónoma de México**, por darme la oportunidad de ingresar al área de concentración de **Eléctrica** con opción en **Telecomunicaciones**, pudiendo concluir los estudios de postgrado, así mismo por permitirme conocer a la valiosa plantilla de profesores que la conforman, logrando obtener el conocimiento y la formación de la cual me siento orgulloso.

**Al Dr. Miguel Moctezuma Flores:**

Por haber aceptado dirigir este trabajo de tesis, por la paciencia que siempre tuvo con mi persona en cada una de las ocasiones en que le consulté, por el apoyo y la enorme ayuda brindada durante su dirección. Y sobre todo por haber conocido a una persona sencilla y amable a pesar de todo el conocimiento y de sus logros a lo largo de su exitosa carrera.

A los **doctores** que me honraron con su participación como miembros del jurado, a quienes les agradezco la ayuda proporcionada en la mejora de este trabajo. Cuyos nombres son:

**Dr. Jorge Alberto Márquez Flores. Dr. Pablo Roberto Pérez Alcázar.**

**Dra. Ana María Vázquez Vargas**

**Dr. Víctor García Garduño.**

A **Dios:**

Por cambiar radicalmente mi existencia, al darme la oportunidad de conocer su infinita bondad y amor a pesar de mis errores. Gracias Señor por haber hecho posible esta meta de lograr pertenecer a la UNAM, por no dejarme desfallecer en los momentos difíciles, y por reconocer que **Jesucristo** es mi salvador y esperanza en todo momento.

A mis **padres:**

Por todo su sacrificio y preocupación, por el gran apoyo, comprensión y paciencia incondicional durante todo este tiempo, por el ejemplo, honestidad, consejos, valores y amor que infundaron a mi persona. Por haber confiado en mí siempre a pesar de todo, gracias de verdad y perdón por los errores cometidos en el pasado.

A mi **hermana:**

Por ser alguien tan especial, por su cariño, por las palabras de aliento y la confianza infundada, por todo lo vivido desde niños; me siento orgulloso de contar siempre y en todo momento contigo.

A mis **amigos y compañeros:**

Ricardo, Saúl, Ernesto, Alex, Alicia, Cira, Oscar, por su valiosa amistad, amabilidad, honestidad, sinceridad y consejos. De verdad gracias amigos.

# Índice

<b>Capítulo 1. Introducción.....</b>	<b>1</b>
1.1 Objetivo.....	1
1.1.1 Antecedentes del tema.....	1
1.1.2 Definición del problema.....	2
1.1.3 Forma propuesta de solución.....	2
1.1.4 Metodología empleada.....	3
1.1.5 Aportación.....	4
<b>Capítulo 2. Fundamentos de percepción remota.....</b>	<b>5</b>
2.1 Descripción general.....	5
2.1.1 Percepción remota o teledetección.....	6
2.2 Principios y fundamentos.....	7
2.2.1 Espectro electromagnético.....	7
2.2.2 Fuentes de radiación electromagnética.....	10
2.2.3 Efectos atmosféricos en la propagación de la radiación electromagnética.....	10
2.3 Aplicaciones de la percepción remota.....	11
2.3.1 Fotografía aérea.....	11
2.3.1.1 Clasificación de las fotografías aéreas.....	13
2.3.1.2 Capacidades de las cámaras fotográficas.....	14
2.3.2 Imágenes multiespectrales.....	14
2.3.3 Imágenes de radar de apertura sintética S.A.R.....	17
2.3.3.1 El método S.A.R.....	18
2.3.3.2 El sensor radar.....	18
<b>Capítulo 3. Principales satélites de percepción remota.....</b>	<b>20</b>
3.1 Introducción.....	20
3.2 ENVISAT.....	20
3.2.1 Sensores del ENVISAT.....	21
3.3 Los satélites europeos de la detección alejada ERS-1 y ERS-2.....	24
3.3.1 ERS-1.....	24
3.3.2 Parámetros técnicos ERS-1.....	25
3.3.3 ERS-2.....	27
3.3.4 Parámetros técnicos ERS-2.....	28
3.4 LANDSAT.....	28
3.4.1 El satélite LANDSAT 5.....	29
3.4.1.1 Resolución espectral.....	29
3.4.1.2 Resolución espacial.....	29
3.4.1.3 Análisis digital.....	30
3.4.2 LANDSAT 7.....	30
3.4.3 Principales diferencias entre el LANDSAT 7 y el LANDSAT 5.....	31

---

---

## ÍNDICE

---

3.4.3.1 Bandas espectrales y resolución espacial.....	32
3.4.3.2 Niveles de corrección geométrica de las imágenes LANDSAT 7.....	32
3.4.3.2 Formatos y soportes existentes.....	33
3.5 SPOT.....	33
3.5.1 Tres satélites en órbita.....	34
3.5.2 SPOT 4.....	35
3.5.2.1 Equipos auxiliares.....	35
3.5.3 SPOT 5.....	35
3.5.3.1 La cámara HRG.....	35
3.5.3.2 La cámara HRS.....	36
3.5.3.3 Capacidad de almacenamiento.....	36
3.5.3.4 Transmisión a tierra.....	36
3.5.7 Las bandas espectrales.....	37
3.6 IKONOS.....	37
3.6.1 Disposición de los datos de IKONOS.....	37
3.6.2. Principales características del satélite IKONOS.....	38
3.6.3 Tipos de resoluciones obtenidas por IKONOS.....	39
3.6.4 Niveles de procesamiento.....	39
3.6. Aplicaciones y ventajas de las imágenes IKONOS.....	40
<b>Capítulo 4. Segmentación Bayesiana.....</b>	<b>42</b>
4.1 Introducción.....	42
4.2 Esquemas de segmentación.....	43
4.2.1. Etapas de la segmentación.....	43
4.2.1.1 Presegmentación Bayesiana.....	43
4.2.1.2 Presegmentación según el criterio de Bayes.....	44
4.2.1.3 Segmentación Bayesiana.....	45
4.2.2 Formulación del problema por medio de la teoría de diseño estadístico.....	46
4.2.2.1 Función de pérdidas.....	46
4.3 Función discriminante de Bayes.....	48
4.3.1 Probabilidad máxima de decisión.....	49
4.3.2 Probabilidad de error.....	50
4.3.3 Funciones discriminante óptimas para modelos distribuidos normalmente.....	52
4.3.3.1.- Distribución normal.....	52
4.5 Clase de rechazo.....	58
<b>Capítulo 5. Segmentación por campos aleatorios de Markov.....</b>	<b>60</b>
5.1 Introducción.....	60
5.2 Etiquetados en imágenes.....	61
5.3 Sistemas de vecinos y asociación de cliques.....	63
5.4 Topologías. (Asignación de etiquetas en los campos aleatorios de Markov).....	66
5.4.1. Función de energía para la segmentación Markoviana.....	67
5.4.2. Algoritmo de recocido simulado.....	70
5.4.2.1 Criterio de Metrópolis y criterio de Potts.....	71

---

## ÍNDICE

---

5.4.3 El campo aleatorio de Markov y la distribución de Gibbs.....	72
5.4.4 Campos aleatorios de Gibbs (GRF, Gibbs random fields).....	74
5.4.5 Equivalencia Markov-Gibbs. ....	76
5.5 Aproximación máximo a posteriori. ....	78
5.6 Modelos de Markov.....	80
5.6.1.- Automodelos.....	80
5.6.2.- Modelo logístico multinivel (MLL).....	83
5.7 Análisis de clase de rechazo.....	84
<b>Capítulo 6. Esquemas de relajación.....</b>	<b>87</b>
6.1 Introducción.....	87
6.2 Marco teórico. ....	87
6.3 Relajación probabilística. ....	89
6.3.1 Umbralización. ....	90
6.3.2 Etiquetado de la relajación probabilística.....	90
6.4 Relajación determinística. ....	94
6.4.1 Algoritmo modo condicional iterativo (ICM).....	94
6.4.2 Implementación del algoritmo ICM para la relajación determinística.....	95
6.5 Relajación estocástica.....	98
6.5.1 Implementación.....	99
<b>Capítulo 7. Resultados de segmentación. ....</b>	<b>102</b>
7.1 Imágenes a procesar.....	102
7.2 Resultados de la segmentación aplicando la función discriminante de Bayes para la imagen 1.....	104
7.3 Resultados de la segmentación aplicando el criterio de Bayes y los CAM para la imagen 1.....	106
7.3.1 Esquema comparativo mediante un acercamiento (zoom) de los métodos de segmentación para la escena 1.....	108
7.4 Resultados de la segmentación aplicando la función discriminante de Bayes para la imagen 2.....	110
7.5 Resultados de la segmentación aplicando el criterio de Bayes y los CAM para la imagen 2.....	112
7.5.1 Esquema comparativo mediante un acercamiento (zoom) de los métodos de segmentación para la escena 2.....	114
7.6 Resultados del proceso de relajación.....	117
7.6.1 Relajación probabilística para la escena urbana 1 .....	118
7.6.2 Relajación determinística para la escena urbana 1.....	121
7.6.3 Relajación estocástica para la escena urbana 1, con diferentes configuraciones de vecindad y 40 iteraciones cada una.....	125
7.6.4 Relajación probabilística para la escena urbana 2 .....	129
7.6.5 Relajación determinística para la escena urbana 2.....	132
7.6.6 Relajación estocástica para la escena urbana 1, con diferentes configuraciones de vecindad y 40 iteraciones cada una.....	136

---

## ÍNDICE

---

---

*Capítulo 8. Conclusiones.....140*

*Capítulo 9. Bibliografía. ....154*

---

---

# 1 Introducción

## 1.1 Objetivo

El objetivo en la presente tesis es el desarrollo de esquemas de segmentación de imágenes mediante métodos probabilísticos, como lo son la segmentación bayesiana y la segmentación por campos aleatorios de Markov; así como también el análisis de procesos de regularización basados en métodos de relajación probabilística, determinística y estocástica. Estos análisis tienen aplicación en la segmentación de imágenes de fotografía aérea pancromática de alta resolución de escenas urbanas del distrito federal.

## 1.2 Antecedentes del tema

La **Percepción Remota** (*Remote Sensing*) o **Teledetección** puede definirse como la ciencia o la técnica para obtener información de un objeto o un proceso, por análisis de datos obtenidos mediante un instrumento que no está en contacto físico con los objetos [1.txt]. Otra definición de percepción remota puede ser: "Es la utilización de sensores para la adquisición de informaciones sobre objetos o fenómenos sin que haya contacto directo entre el sensor y los objetos"[1.p.e].

La percepción remota es una herramienta del procesamiento digital de imágenes, que nos ayuda a interpretar mejor los diferentes tipos de visualización de imágenes a distancia que se pretende conocer, para describir las diversas áreas que la conforman [1.p.e]. Estas áreas están determinadas por diferentes componentes como son agua, vegetación, así como las diferentes texturas del suelo y diversas construcciones.

Las contribuciones de la percepción remota en la evaluación de zonas urbanas, derrames petroleros, deforestación, así como en la segmentación de escenas polares, entre otras, dependen de la simplicidad o complejidad de la zona bajo estudio y del tipo de medidas o contribuciones que el sensor remoto puede realizar [4.p.e].

En el extremo de la simplicidad pueden encontrarse situaciones que involucran a un solo objeto de fácil detección y evaluación, como lo son la identificación de estructuras urbanas y su crecimiento, la evaluación de áreas con erosiones, inundaciones o las porciones de un bosque devastado por un incendio. Por otro lado, en el extremo de la complejidad, se encuentra el análisis de todo un sistema relacionado con un recurso natural, como lo es la prospección petrolera y minera, o la inferencia de cuerpos de agua subterránea, las que emplean contribuciones de varios tipos de sensores aerotransportados y satelitarios, que utilizan medidas y muestreos en campos, que requieren del modelado físico y matemático del sistema y de técnicas refinadas de la teoría estadística, que caracteriza la presencia de tales recursos. Hasta ahora, la mayoría de las aplicaciones se han concentrado principalmente en las situaciones simples.

La aplicación de la percepción remota como parte de un sistema complejo e integrado de captura extensa de datos acerca de varios parámetros físicos, del modelado matemático, y de la predicción de resultados, constituye una herramienta valiosa que demanda de una alta especialización científica; éste es

el caso del estudio de la urbanización en de la ciudad de México, donde la conjugación de muchos elementos permite la exploración de los recursos naturales tan necesarios. Para atacar problemas complejos se requiere de extensivas investigaciones y del desarrollo de nuevos métodos y técnicas. En países en vías de desarrollo, como lo es el nuestro, un considerable esfuerzo tendrá que ser invertido para la creación de los recursos técnicos y humanos necesarios para la adopción de la percepción remota en las tareas de estudio y análisis de los recursos naturales.

La investigación científica en la observación terrestre mediante imágenes de percepción remota, ha seguido una línea aceptable dentro de los últimos años, haciendo uso de algoritmos sustentados en teorías matemáticas. Uno de sus principales objetivos es la información obtenida a través de la cartografía y la distribución urbana. El procedimiento de obtención de cartografía temática para el análisis territorial a partir de imágenes de satélite se basa en un proceso de segmentación en el que entran en juego técnicas estadísticas de clasificación apoyadas en datos espectrales. Con la disponibilidad de otros tipos de datos, como fotografías aéreas o imágenes de satélite de alta resolución, y dada la heterogeneidad de ciertos paisajes, se desarrollan técnicas de segmentación basadas en el cálculo de características que describan la textura de las imágenes.

### ***1.1.2 Definición del problema***

Debido a la complejidad en el proceso de formación de las imágenes y a la representación parcial en datos numéricos de la escena adquirida, al aplicar técnicas estándares de segmentación, se observa notoriamente resultados sobre-segmentados o sub-segmentados. Por tales razones, en el presente trabajo de tesis, se pretende el estudio de técnicas bayesianas y markovianas para la segmentación de imágenes de percepción remota, así como la aplicación de métodos de relajación, con propósito de obtener escenas segmentadas con menos píxeles aislados y con regiones más homogéneas.

Entendiendo por concepto de **segmentación** una técnica que consiste básicamente en la separación de la imagen en regiones u objetos de interés, el primer paso del análisis propuesto en esta tesis, es que a partir de las imágenes de fotografía aérea pancromática de alta resolución se desea procesarlas a través de los métodos de segmentación bayesianos y markovianos.

Posteriormente se someten estos resultados a tres distintos métodos de relajación, con el fin de regularizar la imagen lo mejor posible, debido a que existen píxeles aislados que se deben erradicar ya que aparecen fuera de contextos locales. El propósito es el de aplicar una limpieza morfológica que elimine los píxeles que quedaron aislados en las regiones de interés, de acuerdo a la probabilidad de ocurrencia respecto a las vecindades locales que componen la escena. Se analizará la funcionalidad de tres métodos de relajación a fin de escoger uno solo.

### ***1.1.3 Forma propuesta de solución***

La forma propuesta de la solución para este trabajo, consistió en obtener primeramente los conocimientos básicos en forma general, lo cual se incluye en el capítulos 2; en el capítulo 3 se muestra la operación sistemática de los principales satélites de percepción remota; específicamente en los capítulos 4 y 5 se mencionan las bases teóricas para la implementación de los algoritmos utilizados en los métodos de segmentación; por último en el capítulo 6 se implementan tres algoritmos de relajación. Más adelante se hace una comparación de cada uno de los métodos de segmentación para acercarnos a un mejor resultado así como de cada uno de los métodos de relajación. Así pues, para poder abordar la

problemática de una manera más eficiente y correcta, métodos probabilísticas, estocásticos y contextuales son abordados en la presente tesis.

### **1.1.4 Metodología empleada**

A efectos de definir el método más adecuado al tipo de datos a procesar, y de forma independiente se aplicaron dos métodos de segmentación sobre la imagen original: un método bayesiano y otro por campos de Markov. Posteriormente, a ambos resultados y también en forma independiente, se les aplica una etapa de relajación a fin de obtener un resultado más homogéneo. Las etapas 1 y 2 de la siguiente sección describen los métodos de segmentación empleados y la etapa 3 describe la relajación.

#### **Etapa 1:**

Para distribuciones gaussianas, la segmentación bayesiana consideró el caso óptimo. Primeramente con una probabilidad a priori de ocurrencia de etiquetas de segmentación y posteriormente bajo la actualización de esta probabilidad. Se consideran los casos:

Caso 1.- Cuando las matrices de covarianza son iguales para diferentes clases. Cuyo significado físico es que la separación de clases o clusters (grupos) son iguales en dimensiones y forma, pero los grupos son centrados sobre diferentes medias.

Caso 2.- Cuando la matriz de covarianza  $C_k$  es de forma diagonal. Cuyo significado físico es que el cluster tiene componentes iguales a lo largo de los ejes principales y la distribución es de forma esférica.

#### **Etapa 2:**

Se utiliza una segmentación por Campos Aleatorios de Markov (CAM), en donde la distribución de etiquetas de segmentación se determina por medio de funciones de energía que hacen uso de las medias obtenidas por medio de la segmentación bayesiana. Se emplea un algoritmo de Recocido Simulado (Simulated Annealing) para la minimización de la función de energía.

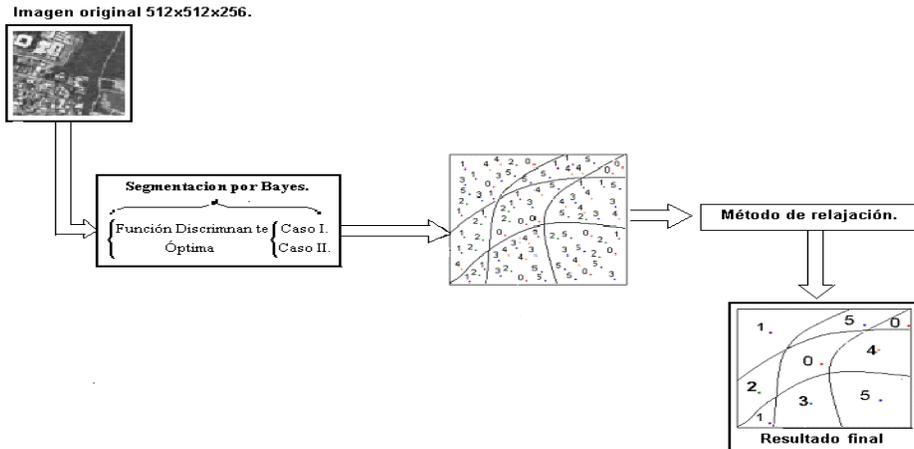
#### **Etapa 3:**

Se hace una comparación de los dos resultados anteriores y se someten ambos al análisis de los métodos de relajación, los cuales son: relajación probabilística, relajación determinística y por último la relajación estocástica. De modo que se puedan comparar los tres métodos y decidir, en base al costo computacional definido en tiempo y a la regularización de la imagen, el método más adecuado.

- En el método de relajación probabilística, las probabilidades de etiquetado son iterativamente actualizadas de forma independiente de acuerdo a la ocurrencia en una vecindad.
- En el método de relajación determinística, se busca una estimación de regularización de una segmentación basada en un modelo de máximo a posteriori (MAP). Para esto nos basamos en los CAM (campos aleatorios de Harkov) para la definición de los criterios de etiquetado de los píxeles, así como el algoritmo ICM (modo condicional iterativo).
- En el método de relajación estocástica, se observa una secuencia estimada de un elemento aleatorio escogido y a su vez la generación de un nuevo valor para el elemento, en concordancia con la distribución de probabilidad propuesta. Entonces el nuevo valor es aceptado o rechazado de acuerdo con algún cálculo estadístico de acuerdo a una cantidad no grande de una secuencia de elementos

vecinos. Podemos asumir diferentes propuestas de conjuntos de configuración de vecindad, tomando como base el conjunto integrado por vecindades en 8-conexidad y cuatro arreglos de vecindad en un número de  $\psi$  píxeles (adicionando 2 renglones y 2 columnas) para adoptar distintas formas de análisis de contextual.

La siguiente figura ilustra el esquema final que se propone en este trabajo de tesis.



### 1.1.5 Aportación

El presente trabajo pretende aportar en primera instancia un estudio de métodos de segmentación desde un enfoque de aplicación a imágenes de percepción remota; así como el análisis y la aplicación de métodos de relajación que buscan proporcionar resultados con píxeles más regulares en el contexto de distribuciones de etiquetas de segmentación.

Además se pretende dejar las bases para el procesamiento por métodos de relajación, imágenes segmentadas provenientes de diversas fuentes remotas o multispectrales, haciendo modificaciones en los programas implementados.

# 2 Fundamentos de Percepción Remota

## 2.1 Descripción general

Por **procesamiento digital de imágenes** (PDI) se entiende la manipulación de una imagen a través de la computadora. Para comparar, en la disciplina de reconocimiento de patrones, la entrada del proceso es una imagen y la salida consiste en una clasificación o una descripción de la misma [1 p.e].

Uno de los objetivos de utilizar el procesamiento digital de imágenes es mejorar el aspecto visual de ciertos elementos estructurales, de manera que permitan tener una percepción adecuada del contenido de la información para una mejor interpretación [1 txt]. Es por ello que esta disciplina ha generado un gran interés, aunado al desarrollo vertiginoso de la tecnología de computación, y al de nuevos algoritmos para tratar señales bidimensionales, y tridimensionales, creándose así una gama de aplicaciones cada vez mayor en el tratamiento de la información.

Como resultado de esa evolución en la tecnología de la computación, el procesamiento digital de imágenes está ampliando sus dominios, que incluyen las más diversas áreas, como por ejemplo: **análisis de recursos naturales y meteorología por medio de imágenes de satélites; transmisión digital de señales de televisión; análisis de imágenes biomédicas, incluyendo el conteo automático de células y examen de cromosomas; análisis de imágenes metalográficas y de fibras vegetales; obtención de imágenes médicas por ultrasonido, radiación nuclear o técnicas de tomografía computarizada; aplicaciones en automatización industrial envolviendo el uso de sensores visuales en robots, entre otras.**

La formación de una **imagen digital**, es la conversión de la función continua  $f(x,y)$  en una serie de valores discretos (generalmente valores enteros), por muestreo tanto del plano espacial como de la intensidad. En otras palabras, una imagen digital se representa como una matriz bidimensional denotada por  $I[m,n]$ , donde cada elemento corresponde al valor de intensidad  $I$  de cada punto de muestreo de la imagen  $[m,n]$ , definido como **píxel**.

Un píxel se designa con la matriz cuyos subíndices indican su pertenencia a una fila y columna determinada, es decir,  $I_{(i,j)} = 128$  indica que el píxel que se encuentra en la fila  $i$  y columna  $j$  en la imagen  $I$  tiene un valor de intensidad o nivel de gris igual a 128.

Como se verá a continuación, las técnicas de PDI son aplicadas siempre con los **niveles de gris (NG)** atribuidos a los píxeles de una imagen. Dependiendo de la técnica utilizada el usuario trabajará con una única imagen o con varias imágenes, siendo esta última conocida como técnicas multiespectrales, por tratar de varias imágenes de la misma escena en regiones diferentes del espectro electromagnético.

Las principales técnicas de percepción remota son:

- Contraste de imagen.
- Lectura de píxeles.
- Filtraje espacial
- Operaciones aritméticas. IHS-RGB.
- Componentes principales.
- Modelo de mezcla.
- Segmentación de imagen.
- Clasificación de Imagen.
- Estadística de imagen.
- Eliminación de ruido.

Ahora definiremos el concepto de percepción remota.

### 2.1.1 Percepción remota o teledetección

La **Percepción remota** (*remote sensing*) es la adquisición y medición de información del mundo físico mediante la detección de radiación, de manera que se pueda identificar y catalogar la materia observada. Es decir la **percepción remota** ó también llamada **teledetección** es la ciencia de obtener información de un objeto analizando los datos adquiridos mediante algún dispositivo que no está en contacto físico con dicho objeto [1 p.e].

Existen artefactos diseñados para captar información del mundo físico en distintas bandas (longitudes de ondas) del espectro simultáneamente y de esa manera traducir esos datos a imágenes en distintas capas, cada una representando una longitud de onda en particular. A esto se le conoce como percepción multispectral.

La **percepción multispectral** es una de las técnicas utilizadas para la detección de materiales. Esta técnica utiliza información de distintas bandas de luz para identificar las propiedades del mundo físico [4 p.e]. Esta técnica provee un registro completo de las respuestas de materiales en las distintas longitudes de ondas en consideración, permitiendo así hacer una gráfica continua del espectro de onda para cada píxel de la imagen, conocida como curva espectral. La información es almacenada en la memoria de una computadora para ser procesada e interpretada.

La Percepción remota se basa en la utilización de sensores para la adquisición de informaciones sobre objetos o fenómenos sin que haya contacto directo entre el sensor y los objetos. Los **sensores** son elementos capaces de coleccionar energía proveniente del objeto, convirtiéndola en una señal posible de ser registrada y presentándola en forma adecuada para la extracción de informaciones utilizando la energía electromagnética o radiación electromagnética.

Así mismo, los sensores remotos son sistemas fotográficos u óptico-electrónicos capaces de detectar y registrar, en forma de imágenes o no, el flujo de energía radiante reflejado o emitido por objetos distantes. Durante la fase de adquisición de datos por los sensores, se pueden distinguir los siguientes elementos básicos: energía radiante, fuente de radiación, objeto (también llamado albedo), trayectoria y sensor (sistema de formación de imágenes óptico y detector) [3 p.e].

Los sensores pueden ser clasificados en función de la fuente de energía o en función del tipo de producto que produce.

**En función de la fuente de energía:**

- 1) **Pasivos:** no poseen una fuente propia de radiación. Miden la radiación solar reflejada o la radiación emitida por los albos. Por ejemplo, los sistemas fotográficos.
- 2) **Activos:** poseen su propia fuente de radiación electromagnética, trabajando en fajas restringidas del espectro. Por ejemplo, los radares.

**En función del tipo de producto:**

1) **Formación de imágenes.-** se obtiene como resultado una imagen de la superficie observada. Suministran informaciones sobre la variación espacial de la respuesta espectral de la superficie observada, clasificados en:

a).-**Sistema de cuadro ("framing systems"):** adquieren la imagen de la totalidad de la escena en un mismo instante, por ejemplo: RBV.

b).-**Sistema de barrido ("scanning systems")** por ejemplo: TM, MSS, SPOT.

c).- **Sistema fotográfico**

2) **No formación de imágenes.-** no generan una imagen de la superficie observada. Por ejemplo, los radiómetros (salida en dígitos o gráficos) y los espectro radiómetros (firma espectral).

Son esenciales para la adquisición de informaciones minuciosas sobre el comportamiento espectral de los objetos de la superficie terrestre.

## **2.2 Principios y fundamentos**

### **2.2.1 Espectro electromagnético**

Las cargas eléctricas estacionarias producen campos eléctricos, las cargas eléctricas en movimiento producen campos eléctricos y magnéticos. Los cambios cíclicos en estos campos producen radiación electromagnética de esta manera la radiación electromagnética consiste en una oscilación perpendicular de un campo eléctrico y magnético.

La radiación electromagnética transporta energía de un punto a otro, esta radiación se mueve a la velocidad de la luz (siendo la luz un tipo de radiación electromagnética).

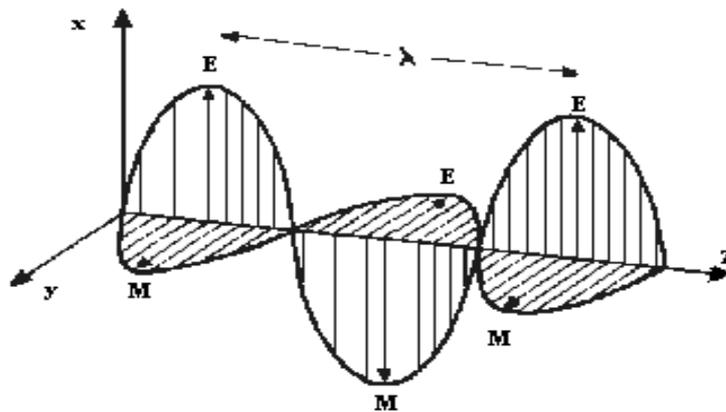
Todo lo que existe en la naturaleza radia energía, en todas las direcciones; radian energía las personas, las nubes, los gases atmosféricos, los vegetales, los objetos metálicos, etc.

La radiación de energía, a diferencia de la conducción o la convección, no necesita de ningún "medio" para la transferencia de energía.

Por ejemplo lo que ocurre entre el Sol y la Tierra, la energía radiactiva viaja por el vacío del espacio, mientras que la convección transporta calor solo a través de cuerpos o a través de fluidos como lo son la atmósfera o los océanos.

La intensidad de la energía que un objeto radia depende básicamente de su temperatura en una relación directa con su longitud de onda.

La radiación electromagnética transporta energía de un punto a otro, esta radiación se mueve a la velocidad de la luz (siendo la luz un tipo de radiación electromagnética). La radiación (ondas) electromagnética está constituida por diversas características físicas (intensidad, longitud de onda, frecuencia, energía, polarización, etc.). Independientemente de esas características, todas las ondas electromagnéticas son esencialmente idénticas, presentando independencia con relación a la existencia o no de un medio de propagación (propiedad importante de este proceso de transferencia de energía). Esta independencia es fácil de entender a través de la figura 2.3 [5 p.e].



E = Campo eléctrico.  
M = Campo magnético.

**Figura 2.3. Representación del campo electromagnético [5 p.e].**

El campo eléctrico y el campo magnético son perpendiculares entre sí y ambos oscilan perpendicularmente a la dirección de propagación de la onda, así el campo eléctrico genera un campo magnético y el campo magnético genera un campo eléctrico.

La velocidad de propagación de la onda electromagnética en el vacío corresponde a la velocidad de la luz ( $3 \times 10^8$  m/s). El número de ondas que pasa por un punto del espacio en un determinado tiempo define la frecuencia (f) de la radiación. La frecuencia de onda es directamente proporcional a la velocidad de propagación de la radiación. Cuanto mayor es la velocidad de propagación de la onda, mayor es el número de ondas que pasarán por un punto en un tiempo dado (t) y mayor será su frecuencia. La velocidad de propagación (v) en un medio dado es constante.

La onda electromagnética también puede ser caracterizada por la longitud de onda (lambda) que puede expresarse por la ecuación:

$$\lambda = v / f \dots\dots\dots(2.1)$$

El rango de longitud de onda o frecuencias en que se puede encontrar la radiación electromagnética es muy amplia. Con la tecnología disponible actualmente, se puede generar o detectar la radiación electromagnética en una extensa faja de frecuencia, que se extiende de 1 a 1024 Hz, o longitudes de onda en la faja de 108 metros a 0,01A [5 p.e].

Este espectro es subdividido en rangos, representando regiones que poseen características peculiares en términos de los procesos físicos, generadores de energía en cada rango, o de los mecanismos físicos de detección de esta energía. Dependiendo de la región del espectro, se trabaja con energía (electro-volts), longitudes de onda (micrómetro), o frecuencia (hertz). Las principales franjas del espectro electromagnético están representadas en las figura 2.4 y 2.5.

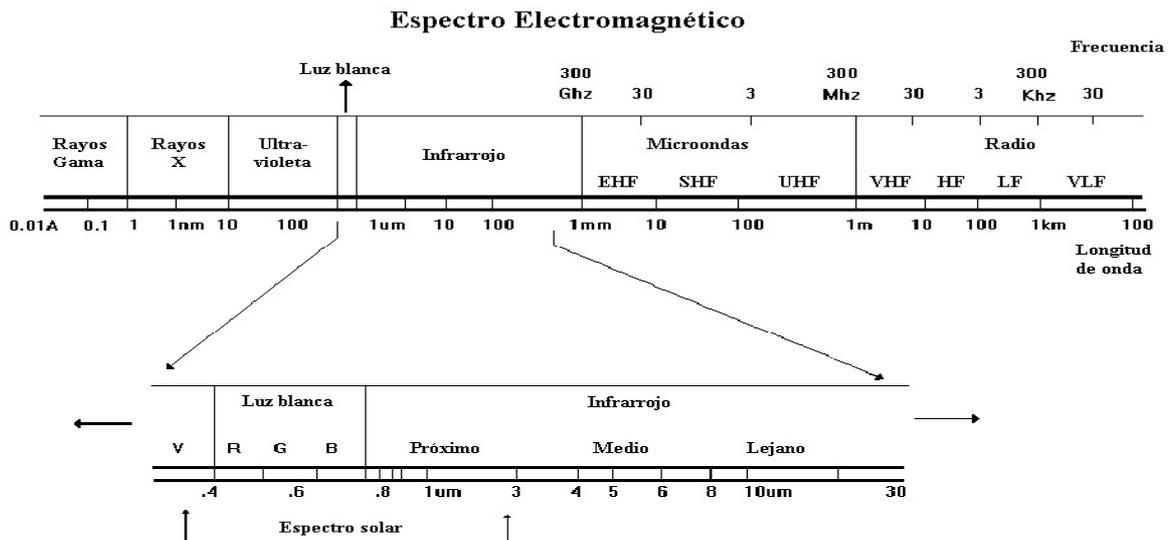


Figura 2.4.-Espectro electromagnético.

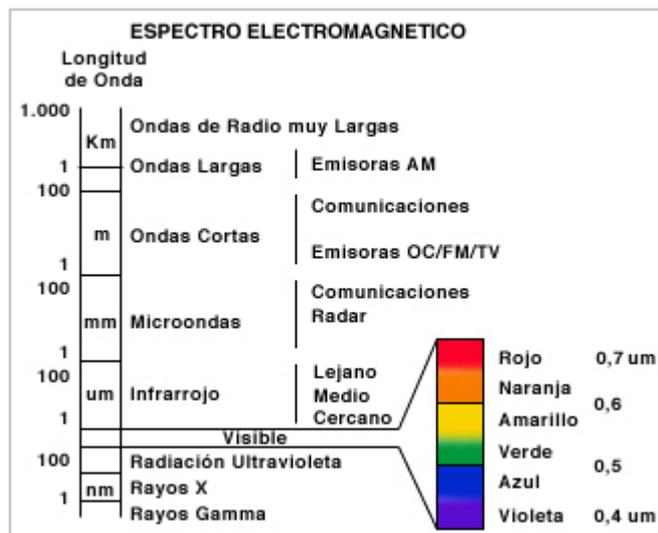


Figura 2.5.- Rangos comunes del espectro electromagnético. La percepción remota opera en varios de ellos [5 p.e].

La energía electromagnética puede ser adquirida mediante técnicas de teledetección. La luz visible, la emisión infrarroja y las microondas son algunas expresiones que forman parte

del espectro electromagnético y se propagan a la velocidad de 300.000 kilómetros por segundo.

Las ondas emitidas por una fuente de energía al pasar a través de cualquier medio, especialmente de la atmósfera, no se transmiten completamente y sufren perturbaciones. En este proceso se manifiesta la Ley de Conservación de Energía, esto es:

$$I = R + A + T \dots\dots\dots(2.2)}$$

en donde:

- I = Energía incidente**
- R = Energía reflejada**
- A = Energía absorbida**
- T = Energía transmitida**

Existen rangos del espectro electromagnético en los cuales se observa que la atmósfera produce mínimas perturbaciones. A estos rangos se les denomina como ventanas atmosféricas. Los elementos involucrados en la obtención de datos, además del medio de transmisión, son: fuente de energía, objeto y sensor. De acuerdo a estos elementos, la percepción remota puede dividirse en: pasiva de energía reflejada, pasiva de energía emitida y activa de energía reflejada.

### **2.2.2 Fuentes de radiación electromagnética**

Las fuentes de radiación electromagnética (REM) pueden ser divididas en naturales (sol, tierra, radioactividad) y artificial (radar, láser, entre otras).Cualquier fuente de energía electromagnética es caracterizada por su espectro de emisión, el cual puede ser continuo o distribuido en fajas discretas. El sol, por ejemplo, emite radiación distribuida continuamente en una faja que va de los rayos X hasta la región de microondas, aunque concentrado en el intervalo de 0,35  $\mu\text{m}$  - 2,5  $\mu\text{m}$  [5 p.e]. .

Toda sustancia con temperatura superior al cero absoluto (0, K ó -273° C) emite radiación electromagnética, como resultado de sus oscilaciones atómicas y moleculares. Esta radiación emitida puede incidir sobre la superficie de otra sustancia pudiendo ser reflejada, absorbida o transmitida. En el caso de la absorción, la energía es generalmente reemitida, con diferentes longitudes de onda.

En la práctica, los cuatro procesos: emisión, absorción, reflexión y transmisión ocurren simultáneamente. Dependiendo de las características físicas y químicas de la sustancia, los cuatro procesos ocurren en diferentes regiones del espectro con intensidades distintas. Este comportamiento espectral de las diversas sustancias es denominado firma espectral y es utilizado en percepción remota, para distinguir diversos materiales entre sí.

### **2.2.3 Efectos atmosféricos en la propagación de la radiación electromagnética**

Cuando se colecta información a través de un sensor remoto, sea a nivel de satélite o aeronave, la mayoría de las veces la señal colectada corresponde a la radiación proveniente del

sol, que interactúa con la atmósfera hasta alcanzar el albedo y retorna al sensor interactuando nuevamente con la atmósfera.

Existen regiones del espectro electromagnético en las cuales la atmósfera es opaca, o sea, no permite el paso de la radiación electromagnética. Estas regiones definen las "bandas de absorción de la atmósfera". Las regiones del espectro electromagnético en las que la atmósfera es transparente a la radiación electromagnética proveniente del sol, son conocidas como "ventanas atmosféricas". De esta manera, los siguientes factores interfieren en la percepción remota y siempre deben ser asociados con la atmósfera: absorción, efectos de masa de aire, dispersiones debido a moléculas gaseosas o partículas en suspensión, refracción, turbulencias, emisión de radiación por los constituyentes atmosféricos, entre otros, propiciando de esta manera una atenuación en la información recolectada.

Entonces, podemos concluir que la atenuación de la radiación es dada por:

$$\text{ATENUACIÓN} = \text{ABSORCIÓN} + \text{DISPERSIÓN} \dots \dots \dots (2.3)$$

Debido a los factores de atenuación es importante planificar bien, antes de la adquisición de los datos y de dar inicio a los procesos de interpretación. La franja más utilizada en Percepción remota está entre  $0,3 \mu\text{m}$  y  $15,0 \mu\text{m}$ , la cual se conoce como espectro óptico. En este rango, los componentes ópticos de reflexión y refracción, tales como lentes, espejos y prismas, son utilizados para coleccionar y reorientar la radiación.

## 2.3 Aplicaciones de la percepción remota

### 2.3.1 Fotografía aérea

Otra de las formas más comunes de la Percepción remota es la **Fotografía aérea**. La cámara fotográfica aérea es un sensor remoto artificial. Es un sistema óptico que permite la impresión de un paisaje en forma permanente y posibilita un análisis detallado de formas, tamaños, texturas y colores de los objetos que allí aparecen.

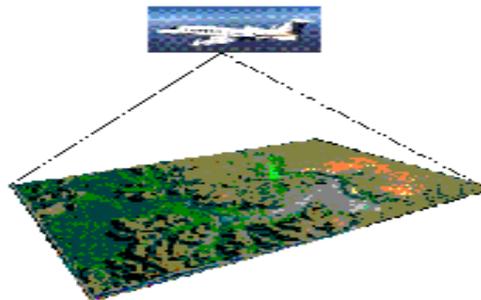


Figura 2.7.- Panorama de una fotografía aérea [3 p.e].

La capacidad de imprimir papel en los rangos del ultravioleta e infrarrojo es una ventaja más del sistema fotográfico respecto de la visión humana. Es posible además, imprimir rangos selectivos de energía al utilizar filtros de diferentes colores. El área captada, el detalle espacial mínimo discernible y la sensibilidad de captación dependen del sistema óptico, del papel que se utilice y de la distancia a la cual se capta el área.

Si la energía captada no se imprime en papel, sino se convierte en energía eléctrica mediante un dispositivo llamado detector, el sensor se denomina radiómetro. La señal eléctrica que se genera puede ser almacenada en algún medio magnético o transmitida a un lugar remoto a través de un sistema de comunicaciones.

El radiómetro barredor de imagen (scanner) es un sistema óptico pendular o rotatorio que en cada ciclo capta energía proveniente de una franja de la superficie. Durante la pendulación o la rotación del espejo principal, esta franja va siendo secuencialmente captada en pequeñas celdas o píxeles. Si el radiómetro se desplaza a una velocidad tal que cada franja que barre es contigua a la anterior, es posible lograr una imagen o cuadro.

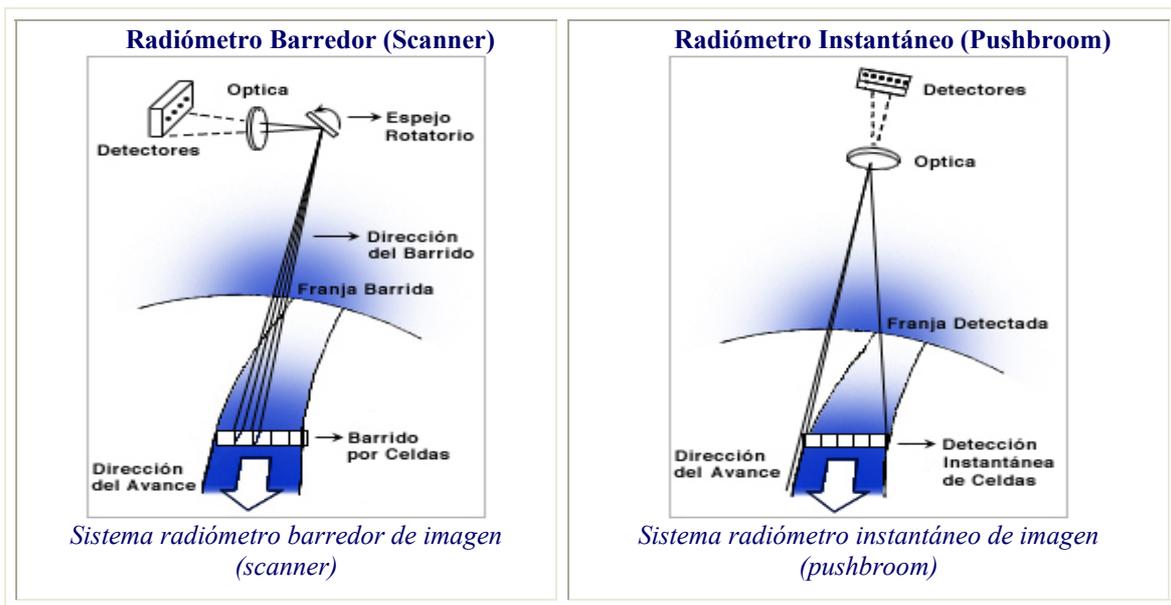


Figura 2.8.-Sistemas radio-metría [3 p.e].

Las ventajas básicas que ofrece la fotografía aérea son:

- \* Permite tener una vista de los detalles de la superficie en estudio en su contexto espacial a escala menor que la imagen satelital. Esto la transforma en una herramienta muy útil para trabajar en áreas pequeñas.

- \* Flexibilidad. Por el hecho de usar plataformas aéreas para embarcar los sensores de imagen, es muy fácil rediseñar el plan de trabajo según las necesidades que vayan surgiendo.

- \* La fotografía nos da una visión instantánea de procesos bajo condiciones dinámicas (inundaciones, derrames de petróleo, tráfico, incendios). Si bien este tipo de trabajo se realiza muy bien con imagen satelital, la fotografía aérea permite desplazarse al lugar de interés sin demoras. De esta forma se dispone de un medio complementario al uso de plataformas satelitales especialmente apto para cubrir aquellos procesos dinámicos con períodos de cambio inferiores a la revisita de los satélites (catástrofes).

Los sensores a bordo del avión serán capaces de cumplir los requerimientos para resoluciones espaciales finas (1-m), flexibles y bandas de espectro angosto, no costosas, frecuente cobertura repetida y rápidos tiempos de rotación. A su vez, se pueden incorporar

filtros para seleccionar rangos espectrales de la energía que se desea registrar (bandas o canales espectrales) [3 p.e].

### **2.3.1.1 Clasificación de las fotografías aéreas**

Según la inclinación del eje óptico las fotografías aéreas se clasifican en:

**La fotografía aérea con enfoque oblicuo.-** El eje de la cámara no es perpendicular al plano horizontal en el momento de la toma. Estas fotografías también se denominan panorámicas; las deformaciones de los objetos son grandes debido al efecto de perspectiva. El área fotografiada es de forma trapezoidal.

Es la que se utiliza con más frecuencia, por darnos una idea precisa de lo que allí se encuentra y establece una relación espacial entre los elementos allí presentes. Las cámaras utilizadas son de formato medio, aproximadamente 3 veces mayor a los formatos de 35mm, lo que permite una calidad sobresaliente, las imágenes obtenidas pueden archivarse en cualquiera de los formatos electrónicos disponibles en la actualidad.



Figura 2.9.- Fotografías aéreas de corte oblicuo [3 p.e].

**La fotografía aérea con enfoque vertical.-** Son aquellas en las que el eje óptico de la cámara es normal o perpendicular a un plano horizontal de referencia situado a la altura del terreno fotografiado. El enfoque vertical es ideal para la confección de planos y mapas así como levantamientos topográficos del terreno, el equipo utilizado es similar al caso anterior, aunque los formatos suelen ser de tipo medio o grande, lo que permite una mayor gama de ampliaciones manteniendo una calidad magnífica.



Figura 2.9 fotografías aéreas de corte vertical [3 p.e].

**La fotografía termográfica.-** Se utiliza en investigaciones científicas sobre evoluciones del terreno, contenidos químicos, enfermedades de fauna, contaminación medioambiental etc. Además, la imagen infrarroja, capaz de detectar diferencias de hasta 0,5° C entre un objeto y

los que le rodean, permite la localización de elementos invisibles para la fotografía normal como podrían ser los escapes en tuberías enterradas. La última aplicación de esta técnica está en la evaluación de los aislantes utilizados en la construcción de viviendas, presentando las diferencias entre las pérdidas de calor y las reales y determinar un patrón de cambio en la temperatura.

### **2.3.1.2 Capacidades de las cámaras fotográficas**

El equipo fotográfico ha de ser específico y se define como scanner de infrarrojos, algunos de los cuales disponen de una excepcional sensibilidad que puede llegar a detectar diferencias de 0,2°C con una resolución espacial de 0,8 m.

La **ortofotografía** es una técnica fotográfica que permite la creación de fotografías llamadas georeferenciadas y orto rectificadas para la producción de mapas de gran precisión y con composición continua, eliminando la tradicional fotocomposición realizada en los procesos de post-producción y que nunca llegaba a eliminar realmente las líneas de separación entre una fotografía y la siguiente, permitiendo su visualización en pantallas normales de ordenador y el tratamiento de las imágenes con una gran variedad de programas informáticos incluyendo fusión con los programas DTM (modelo digital del terreno) lo que permite una perspectiva lateral.

Estas imágenes son imágenes (digitales) electrónicas. La imagen le ayuda al productor a identificar y manejar los problemas rápida y eficientemente para evitar pérdidas serias (CTIC). El rollo de cámara se utiliza para tomar varias semanas, entonces el tiempo de rotación no fue muy bueno para detectar problemas de los cultivos antes de que estos se convirtieran en daños severos. En la actualidad los rollos tipo IR pueden ser procesados con el mismo sistema como un rollo regular de transparencias. Se puede hacer en locales de revelado en una hora, tan pronto como aterrice el avión, y puede estar listo para explorar en un formato digital después de un corto tiempo. Un dispositivo acoplado a un cargador (CCD), llamado de otra manera como registrador (scanner), se utiliza para convertir las fotografías a una imagen digital de tal manera que la computadora pueda trabajar con los datos. A cada elemento imagen (un píxel o punto) se le asigna un número digital dependiendo en su valor de brillantez. Los valores van en un rango de 0 a 255 en un sistema de computadora de 8 bits.

### **2.3.2 Imágenes multiespectrales**

Una imagen multiespectral o multibanda es un conjunto de imágenes, con las mismas propiedades geométricas, cada una de las cuales recoge el valor de reflectancia en un determinado intervalo de longitud de onda del espectro electromagnético [5 p.e].

En una imagen multiespectral digital (figura 2.10) a la sucesión de valores sobre los diferentes planos que componen la imagen digital multiespectral se le llama *patrón espectral*; dicho de otra manera un píxel representa un patrón espectral. El arreglo espacial de píxel con valores forma lo que se conoce precisamente como patrón espacial.

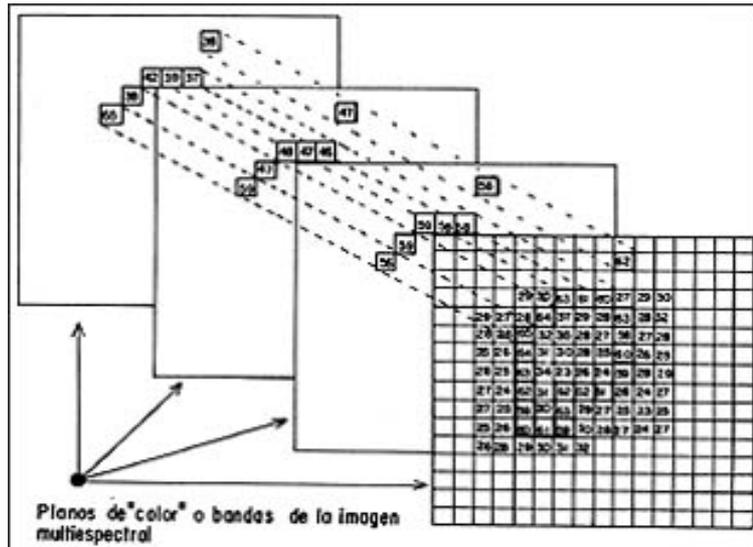


Figura 2.10 Patrones espacial y espectral de una imagen digital multispectral [5 p.e].

En una imagen multispectral digital los patrones espectrales son siempre de tipo lógico y los patrones espaciales son lógicos en la representación numérica y físicos en la representación visual.

La clasificación de una imagen es una operación muy frecuente en la que se emplean técnicas multispectrales. Mediante la clasificación se agrupan píxeles con valores de reflectancia similares, y se separan los píxeles disímiles. Una imagen se puede clasificar de forma visual (p.e. clasificación de una composición en falso color) o bien de forma asistida por ordenador. En el último caso se distingue entre clasificación no supervisada (el ordenador realiza por completo la clasificación sin información adicional del usuario sobre la naturaleza de los píxeles), y clasificación supervisada (para cada clase el usuario especifica un cierto número de píxeles de entrenamiento de los que conoce su naturaleza en el terreno) [1 txt].

El proceso de clasificación mediante información multispectral está basado en la diferenciación de píxeles no sólo teniendo en cuenta sus valores en una determinada longitud de onda, sino considerando la combinación de los valores del píxel en las diferentes bandas empleadas en la clasificación [1 txt]. Así, se pueden distinguir cuerpos o superficies que tienen una misma respuesta espectral en una cierta longitud de onda pero diferente en otro.

Las técnicas de análisis multispectral sirven para evaluar distintos tipos de alteraciones fisiológicas de los cultivos y para detectar el grado de humedad en regiones semiáridas, analizándose conjuntamente con datos meteorológicos auxiliares sobre el terreno y del calendario del crecimiento del cultivo.

Una imagen se puede representar por una matriz de datos, donde las líneas y columnas definen las coordenadas espaciales del "píxel". Para representar la radiancia de la escena para cada "píxel" se utiliza un número finito de bits.

En realidad, la medida de la radiancia, representada en cada "píxel" por su nivel de gris, no es solamente la de la radiación reflejada por la superficie contenida en la escena del "píxel", sino también la radiación debido a la dispersión atmosférica.

La cuantificación de la radiancia continua de una escena está representada por los niveles de gris discretos en la imagen digital, y dada por un número de bits por "píxel" para producir un intervalo de radiancia. Los sensores de la nueva generación obtienen normalmente imágenes en 8 o 10 bits (equivalente a 256 o 1024 niveles digitales).

En realidad el nivel de gris es representado por la radiancia media de un área relativamente pequeña en una escena. Esta área es determinada por la altitud del sistema sensor a bordo del satélite y otros parámetros como el IFOV (Instantaneous field of view), que es el ángulo formado por la proyección geométrica de un único elemento detector sobre la superficie de la Tierra.

En la figura 2.12 se ilustra el sistema de coordenadas usualmente utilizado para representar una imagen digital. El eje x representa el número de columnas y el eje y, el número de líneas.

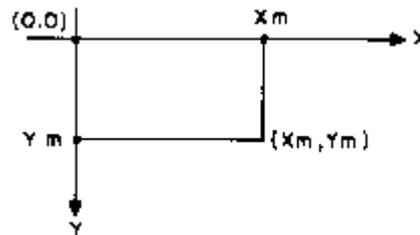


Figura 2.12 [3 p.e].

En el caso de las imágenes multispectrales, la representación digital es más compleja, porque para cada coordenada (x, y), habrá un conjunto de valores de nivel de gris. Cada "píxel" se representa por un vector, con tantas dimensiones como bandas espectrales.

Las imágenes multispectrales son cada vez más utilizadas no sólo en la teledetección, sino también en los demás ámbitos del proceso de imagen, porque además de la información básica de luminancia, contienen información del comportamiento de la escena en función del espectro electromagnético. Esto permite la segmentación en función del color y por tanto permite detectar y aislar elementos que pasarían desapercibidos en una imagen pancromática (imágenes que se encuentran en niveles de gris), como cursos de agua en una imagen de satélite o defectos de una pieza inspeccionada por un sistema de visión artificial.

La imagen multispectral se compone de varias imágenes de la misma escena tomadas en bandas diferentes del espectro. En general se habla de imágenes en color real, cuando la codificación de la imagen contiene información sobre tres o más componentes de color. Ya sea las componentes RGB (rojo, verde y azul), HSI (croma o tono, saturación e intensidad), CMY (cyan, magenta, amarillo), CMYK (cyan, magenta, amarillo y negro) u otras. Una de las características de las componentes de una imagen multispectral es que la mayor parte de la información que contienen es común a todas ellas.

La expresión matemática de la imagen en color, en el espacio RGB, es la combinación de tres funciones bidimensionales, cada una de ellas correspondientes a cada uno de los componentes principales del color (Rojo, Verde y Azul):

$$F(x,y) = f_R(x,y) \cdot f_G(x,y) \cdot f_B(x,y) \dots \dots \dots (2.4)$$

La representación discreta o digitalizada, de una imagen en color real, corresponde a una matriz tridimensional en que cada elemento representa un píxel en una fila y columna con los valores de cada plano de la imagen o componente RGB del color. (RGB: Red, Green, Blue).

Una condición obvia para trabajar con una imagen multiespectral es que sus componentes estén perfectamente alineadas. En caso contrario, al procesar la imagen aparecerán características inexistentes que falsearán el resultado. Normalmente, el error de alineación será de inferior magnitud que el tamaño de los elementos de la imagen, con lo que ésta se podrá procesar sin problemas. Como mucho, en algunos casos podría ser conveniente la aplicación de un filtro paso-bajas.

Como difícilmente será posible solucionar este problema en el propio dispositivo, habrá que corregirlo en la imagen obtenida. El método que actualmente se propone se basa en el hecho de que las diversas componentes de la imagen multiespectral tienen la mayor parte de la información común, lo que supone una alta correlación entre ellas, por lo que es posible calcular su desplazamiento relativo a partir de los resultados de la correlación. Para conseguir una resolución de subpíxel, se aproxima la función correlación de variable real con una función de segundo grado a partir de los resultados de la correlación discreta. El realineamiento se realiza mediante una transformación geométrica, o sea, una transformación de coordenadas (en este caso un desplazamiento de variable real) seguida de interpolación.

Las imágenes multiespectrales, además de la información de luminancia, contienen información del comportamiento de la escena en función del espectro electromagnético, lo que permite la segmentación en función del color. Por ello son ampliamente utilizadas en diversos ámbitos del proceso de imagen. La imagen multiespectral se compone de varias imágenes de la misma escena, tomadas en bandas diferentes del espectro, que tienen que estar perfectamente alineadas. Las figuras 2.13 se muestran ejemplos de imágenes multiespectrales en color real con 256 niveles de intensidad en cada una de las componentes RGB.



Figura 2.13 a).- Imagen con componentes RGB del color (RGB: Red, Green, Blue) [5 p.e].



Figura 2.13 b).- del aeropuerto de Barcelona, captada sensor ETM del satélite LANDAST 5) [14 p.e].

### 2.3.3 Imágenes de radar de apertura sintética SAR

El método de radar de apertura sintética (SAR) para la generación de imágenes es el corazón de la misión. Para observar la superficie de Venus se utilizó un radar trabajando en el rango de frecuencia de las microondas, ya que las nubes que cubren el planeta son opacas a frecuencias visibles pero son transparentes para la alta radiofrecuencia a la que trabaja el radar. La técnica para producir imágenes mediante SAR era conocida desde hacía 40 años. En la mayoría de los casos se usa un radar con una antena de alta directividad en una plataforma móvil, habitualmente un avión o una nave espacial [2 p.e].

**2.3.3.1 El método S.A.R (Radar de apertura sintética)**

El radar de apertura sintética (SAR) obtiene imágenes de la superficie para confeccionar mapas. Las imágenes de un SAR no muestran colores pero proporcionan datos relacionados con la rugosidad de la superficie. Para un sensor SAR una superficie marina es un ejemplo de una superficie rugosa, por el oleaje provocado por el viento, así como una zona boscosa, en cambio un mar en calma representa lo que se llama superficie lisa, así como una carretera o llano con pastizales uniformes.

La palabra "sintética" en la apertura sintética se refiere al tamaño de la antena que recibe los ecos. Esta antena se comporta como la lente de una cámara, será la antena mayor (apertura) la que dé mayor resolución de imagen. Como la antena se mueve mientras recibe los ecos, en la tierra habrá que realizar un procesamiento de los datos recibidos por una antena todavía mayor (una apertura sintética). El tamaño de la antena de apertura sintética lo determina la distancia que recorre el satélite mientras está tomando una imagen radar de superficie. Las antenas de alta ganancia y las del altímetro transmitirán los pulsos radar (que están en una parte no visible del espectro electromagnético) y recibirán los pulsos reflejados (llamados ecos).

El radar transmite energía en pulsos cortos hacia un lado de la plataforma. El SAR forma una imagen lineal, con un ancho determinado por la elevación de la iluminación de la antena determinada por el tiempo en que el radar está en operación continua de captación de imágenes.

Las imágenes producidas por S.A.R. son muy similares a las fotografías aéreas tradicionales aunque tienen algunas diferencias muy significativas, entre ellas es importante mencionar, por ejemplo, que una imagen S.A.R. no presenta vegetación, ya que árboles y rocas reflejan las ondas radar de forma similar, de ahí el inconveniente de utilizarlas para estudios de índole forestal. En la mayoría de los casos, las superficies rugosas determinan brillos en una imagen radar. Las fotografías presentan sus características dependiendo de la situación de la fuente de alimentación, el color y la intensidad de la luz. Las imágenes de radar están formadas a partir de una frecuencia radar simple y se presenta casi siempre en blanco y negro.

**2.3.3.2 El Sensor radar**

El Sensor radar realiza tres funciones distintas:

- Captación de imágenes S.A.R. (para producir imágenes de los rasgos de la superficie).
- Altimetría (para medir la altura de la superficie).
- Radiometría (para detectar las emisiones térmicas naturales del planeta).

*Sistema radar* - Tasa de adquisición de datos: 806 kbps.

Las características técnicas más importantes del sistema radar son las siguientes:

- *Radar de apertura sintética(S.A.R.):*

Frecuencia: 2,385 GHz.

Potencia de pico: 325 W.  
Longitud del pulso: 26,5 ms.  
Perímetro de radiofrecuencia: 4.400-5800 Hz.  
Cuantificación de enlace ascendente: 2 bits.

- *Operando en modos SAR, altímetro y radiómetro:*
  - Resolución SAR: 150 m de elevación/ 150 m. acimut.
  - Resolución del altímetro: 30m.
  - Precisión del radiómetro: 2 ° C.

## 3 Principales satélites de percepción remota

### 3.1 Introducción

El aspecto que más se ha desarrollado de la astronáutica es el de los satélites artificiales debido a las implicaciones económicas y militares que conllevan. Los satélites artificiales de teledetección se aplican a una gran cantidad de campos, desde usos fiscales (cálculo de superficies arboladas, por ejemplo) a la más pura cartografía civil. Su funcionamiento suele basarse en el espectro no visible de radiaciones y son la mayoría los que funcionan con dispositivos infrarrojos, ultravioleta o de radar. Los primeros de esta clase fueron los LANDSAT americanos, lanzados a partir de 1972 de los cuales hoy existen numerosos satélites de esta clase; incluso existen satélites que incorporan todas estas funciones además de las telecomunicaciones.

Otra de las aplicaciones de los satélites artificiales que ha llegado a ser de gran utilidad es la de localización. Los satélites se han empleado en la navegación marítima a modo de faros que permiten a los barcos establecer su posición.

También se cuenta con los satélites puramente científicos destinados a la observación del espacio exterior desde fuera de la atmósfera, evitando así la perturbación que producen las capas atmosféricas en las observaciones.

A continuación describiremos algunos satélites que se han ido desarrollando en los últimos años, definiendo sus principales características técnicas a lo largo de su desarrollo o modificación y el uso al que han sido destinados todos y cada uno de ellos.

### 3.2 ENVISAT

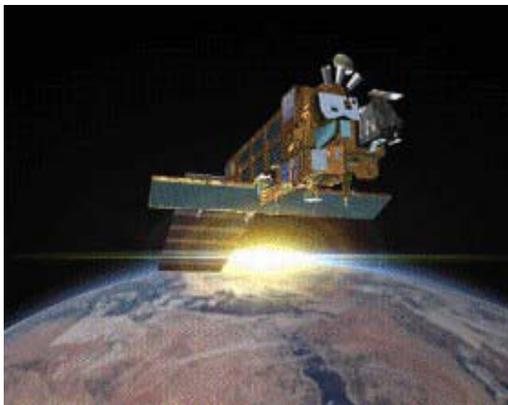


Figura 3.1(a).- Satélite ENVISAT [7 p.e].



Figura 3.1(b).- Satélite ENVISAT[8 p.e].

Las figuras 3.1(a) y 3.1(b) muestran el satélite ENVISAT lanzado en abril de 2002, el cual proporciona una continuidad en las observaciones proporcionadas por el satélite ERS además de agregar nuevas capacidades.

La misión apoyará trabajos de teledetección terrestre a largo plazo, que son esenciales para investigación climatológica y ambiental. ENVISAT supervisará:

- La tierra.
- Océanos.
- Atmósfera.
- Casquetes polares.

Actualmente el satélite ENVISAT es completamente operacional.

#### **3.2.1 Sensores del ENVISAT**

Es el primer radar permanente que incorpora la capacidad de trabajar en polarización dual (dual polaritation), pudiendo transmitir y recibir en la polarización horizontal o vertical. Este modo de polarización alternante (AP) puede mejorar la capacidad de un instrumento del SAR para clasificar diversos tipos de terreno. Las características reflexivas de un material dependen de la polarización de la señal de radar entrante. Los sensores que utiliza son:

**ASAR (advanced synthetic aperture radar).**- Vigila la superficie del planeta generando imágenes de amplia o reducida cobertura geográfica, que supervisa el relieve terrestre.

**MERIS (medium resolution imaging specrometer instrument).**- Toma imágenes de la superficie del planeta y de las nubes capturando la luz de las zonas visible e infrarroja del espectro electromagnético. Es capaz de conocer el color exacto de la superficie oceánica y zonas costeras, reflejar la actividad biológica, observar las nubes, detectar el vapor de agua invisible de la atmósfera e identificar las distintas etapas de crecimiento de las plantas.

**RA-2 Y MWR (radar altimeter 2 & microwave radiometer).**- Miden con una exactitud de 4.5 cm la distancia desde el satélite hasta la superficie del planeta. Su combinación con otros datos permite trazar el perfil de la superficie del mar o hielo, condiciones de las olas y velocidad del viento en el océano.

**GOMOS (global ozone monitoring by occultation of stars).**- Sigue las estrellas y mide sus espectros electromagnéticos. Es capaz de producir tantos datos como una red de 360 estaciones de terrenas.

**MIPAS (michelson interferometer for passive atmospheric sounding).**- Observa la atmósfera en el rango infrarrojo medio y es capaz de rastrear agentes contaminadores industriales y gases de efecto invernadero.

**AATSR (advanced along track scanning radiometer).**- Explora la superficie del océano en varias frecuencias infrarrojas y visibles para medir exactamente su temperatura. La de la superficie del mar la precisa a una exactitud de 0.3 °C.

### Capítulo 3. Principales satélites de percepción remota

**DORIS Y EL LLR (doppler orbitography and radiopositioning integrated by satellite & laser retroreflector).**- Miden su posición orbital con una exactitud de 4.5 cm y su velocidad orbital con una exactitud de 0.4 mm/s. Producen mapas de la superficie del océano y de la topografía terrestre.

**SCIAMACHY (scanning imaging absorption spectrometer for atmospheric chartography).**- Vigila la atmósfera en un amplio rango de longitudes de onda, lo que permite medir cantidades totales y perfiles de gases.

La siguiente figura muestra la ubicación de estos sensores en el satélite ENVISAT.

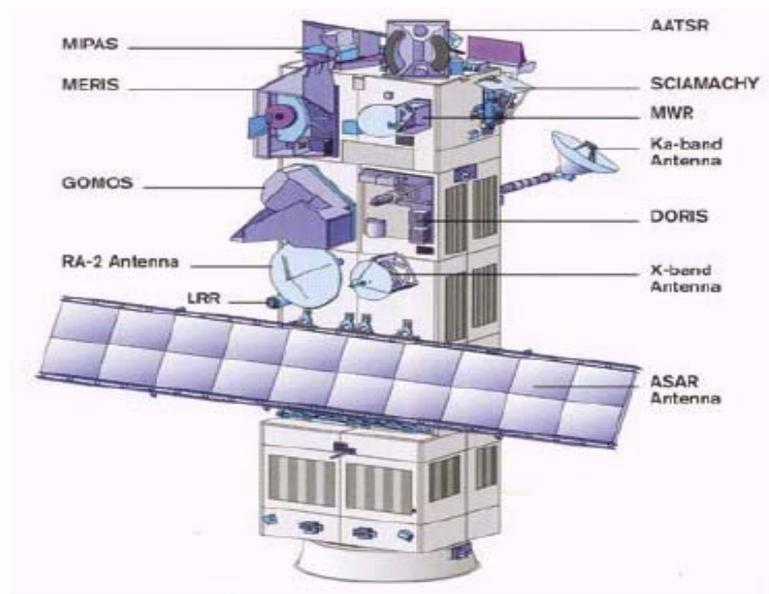


Figura 3.2 Ubicación de los sensores de ENVISAT [7 p.e].

La siguiente tabla muestra un resumen técnico de los sensores del satélite ENVISAT.

INSTRUMENTO	CICLO	RESOLUCIÓN GEOMÉTRICA	LONGITUD DE ONDA	RANGO DE FRECUENCIA
ASAR	Continuo Imagen-30 min	Imagen: 30X30 m Ancho de barrido: 150 X150 Km Global: 1X1 km	Banda C	5.3331 GHz
GOMOS	Continuo	17 Km vertical	UV, Visible IR cercano	250-952 nm
LLR			Visible	532-694 nm
MERIS	Solo de día (43 min)	Completo: 250 X250 m Reducido: 1 X 1 Km	Visible IR Cercano	390-1,040 nm
MIPAS	Continuo	3 Km vertical	IR medio	4.2-14.6 $\mu$ m
MWR	Continuo	20 km	Banda K y Ka	23.8 GHz 36.5 GHz
RA-2	Continuo		S-,Banda Ku	3.2 GHz
AATSR	Continuo	1X1 Km.	Visible IR cercano IR medio	13.57 GHz
DORIS	Continuo	Orbita: 10m radial 25 m otras	Banda S	0.55;0.67;0.865;1.6;3.7;10.85;12 $\mu$ m
SCIAMACHY	Continuo	3 Km vertical	UV, visible, IR cercano	240-2,380 nm

Tabla 3.1 Características principales de los sensores de ENVISAT [7 p.e].

### **Capítulo 3. Principales satélites de percepción remota**

---

**Modo de imagen**, anchura de 100Km, resolución de los 30m (la figura 3.3 es un ejemplo es una toma de Italia).

**Modo amplio de barrido**, anchura de 400Km, resolución 150m (la figura 3.4 es un ejemplo sobre el sur de California).

**Modo de supervisión global**, anchura de 400Km, contando con una resolución de 1Km.

**Viñetas de alta resolución** - resolución de 10m en modo de la onda ASAR, que permite una inspección más avanzada en lo referente a:

- **Clasificación de la tierra**, las combinaciones cruzadas de la polarización ofrecen una discriminación mejor entre los tipos de terreno; tales como vegetación y suelo erosionado, zonas de forestación y urbanas.
- **Supervisión del hielo del mar**, el modo del AP puede ayudar a identificar mejor los límites entre el hielo del mar.

**Oceanografía**, la polarización de HH (horizontal) puede identificar mejor fenómenos del océano tales como ondas internas o efectos debido a la proximidad del suelo marino. La polarización VV (vertical) puede ser más eficaz en ciertas características del océano causadas por condiciones atmosféricas.



Figura 3. 3.- Italia /2003 [8 p.e].

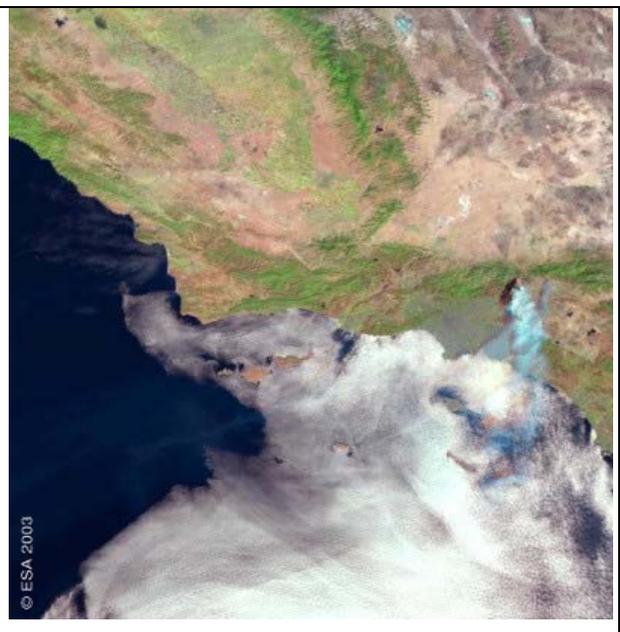


Figura 3.4.- Sur de California EU/2003 [8 p.e].

### **3.3 ERS Los satélites europeos para la detección alejada ERS-1 y ERS-2**

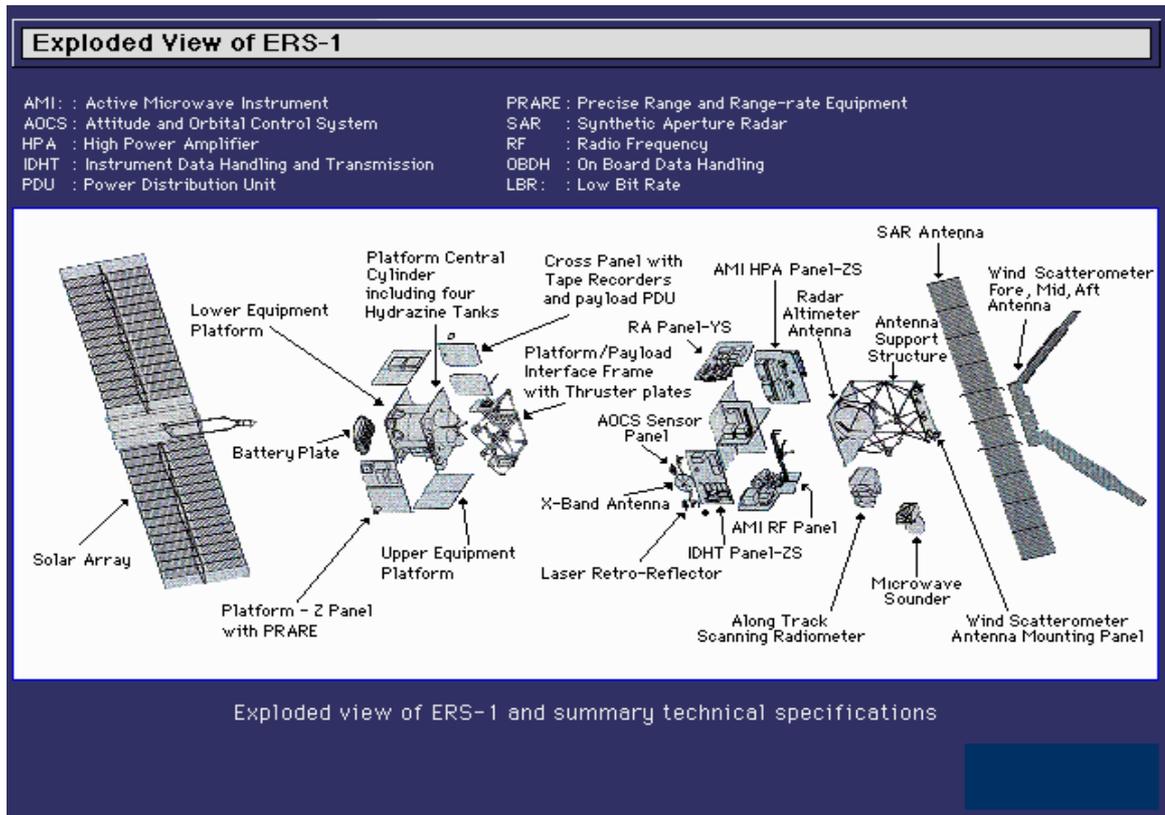
#### **3.3.1 ERS-1**

El satélite ERS-1 fue lanzado el 17 de julio de 1991. ERS-2 fue lanzado el 21 de abril de 1995. Las características de la órbita son iguales que para ERS.-1. Durante el período de la operación paralela de ERS.-1 y de ERS.-2 (misión en tándem) este último está siguiendo a ERS-1 con 24h de retraso [9 p.e].

Ambos satélites se pusieron en una órbita cercano-polar, a una altitud de 780 kilómetros, con una carga útil del instrumento que abarca sensores activos y pasivos de la microonda y un radiómetro infrarrojo termal. Su estructura es la siguiente:

- Instrumento activo microondas (**AMI**) (active microwave instrument) que combina las funciones de un radar de apertura sintética (**SAR**) y un dispersor de viento (wind scatterometer). El SAR funciona en la adquisición de imágenes para cualquiera de los océanos, de regiones polares, de zonas costeras y de la tierra. En SAR produce las imágenes (cerca de 5 kilómetros x 5 kilómetros) en los intervalos regulares de la longitud del océano. El dispersor de viento utiliza tres antenas para la medición de la velocidad del viento de la superficie del mar y su dirección.
- El altímetro del radar (**RA**) (radar altimeter) proporciona medidas exactas de la elevación de la superficie del mar, de las alturas de onda significativas, de los varios parámetros del hielo y una estimación de la velocidad del viento sobre la superficie del mar.
- A lo largo de la pista de exploración del radiómetro (**ATSR**) (along trace scouting radiometer) que combina un radiómetro infrarrojo y un instrumento que sondea la microonda, para la medida de la temperatura superficial del mar, de la temperatura superior de la nube, de la cubierta de la nube y del contenido atmosférico del vapor de agua.
- El equipo de alcance **PRARE** (Precise Range and Range-Equipment) es incluido para la determinación exacta de las características de la posición y de la órbita del satélite, y para la determinación exacta de la posición (fijación geodésica).
- El equipo (**LRR**) (laser retro reflector) permiten la medida de la posición de la órbita del satélite, a través del uso del láser que se encuentran en las estaciones terrestres.

**3.3.2 Parámetros Técnicos ERS-1**



**Figura 3.4.- ERS1 parámetros técnicos [12 p.e].**

Órbita	Tipo: Cercano-circular, polar, sol-síncrono. Altitud: 782 a 785 kilómetros. Inclinação: 98,52 grados. Período: Cerca de 100 minutos. Órbitas por día: 14,3. Ciclo de repetición: tres días, 35-días y 176-días.
Instrumentos	Instrumento activo de microonda que abarca un SAR (modos de la imagen y de la onda) y un wind Scatterometer; un altímetro del radar; un radiómetro de la exploración de la pista; Equipo exacto de alcance (PRARE) y Retro-relectores láser.
Masa	Masa total: 2157,4 kilogramos. Carga útil total: 888,2 kilogramos. Plataforma total: 1257,2 kilogramos Instrumento activo de microondas: 325,8 kilogramos. Altímetro de radar: 96,0 kilogramos. IDHT: 74,0 kilogramos. ATSR: 55,3 kilogramos. PRARE: 12,0 kilogramos. Retro-reflectores del láser: 2,5 kilogramos.
Corriente Eléctrica	Energía máxima de la carga útil: <= 2600 erg. Energía permanente de la carga útil: <= 550 erg. Voltaje de la fuente de alimentación: 23-37 V. Energía a bordo: máximo de 2650 w/h.
Control orbital	Errores absolutos: <= 0.0015deg. sec (sigma 3). Errores máximos: diagonal 0.11deg. (pitch/roll) 0.21deg. (desvío); 0.03deg. armónico y al azar (pitch/roll); 0.07deg. (desvío). Exactitud de la predicción: 30 m (de radial), 15 m (cruz), 1000 m (adelante).

### Capítulo 3. Principales satélites de percepción remota

	Restitución de la órbita: 5 m (de radial), 15 m (cruz) 60 m (adelante).
Tratamiento De Datos	Computadora a bordo: largo de una palabra: 16 bits. Capacidad de memoria de la carga útil: 20 palabras de K máximas. Intercambio de datos de la carga útil: Autobús del tipo de OBDH. Número de los usuarios de la carga útil: redundante.
Comunicaciones	Transponder: velocidad coherente (2 kbit/s). Energía de la transmisión: máximo de 50 a 200 erg. Velocidad de telemetría: 2048 bit/s. Los datos abajo se ligan: - banda X (acoplamiento de la alta tarifa de 105 Mbit/s para el modo de imagen del AMI).

Tabla 3.2.- Parámetros técnicos del ERS1 [12 p.e].

Al realizar su misión el ERS-1 debe moverse en una órbita de modo que sus instrumentos puedan explorar a lo largo de las trayectorias predeterminadas. La órbita elíptica es síncrona de acuerdo al sol, con una altitud de 785 kilómetros y una inclinación de 98.5 grados.

La referencia orbital se obtiene de un sensor infrarrojo situado en la tierra. Se obtiene la referencia del desvío una vez que cada órbita de un sensor del sol se alinea para cuando el satélite está en un punto particular en su órbita. Los medios primarios del control de la altitud son proporcionados por un sistema de ruedas, que pueden ser hechas girar en cualquier dirección, intercambiando el ímpetu angular por el satélite en el proceso.

La misión ERS-1 también incluye maniobras del ajuste para sincronizar el período orbital con requerimientos para la cobertura de tierra. Para los parámetros orbitales se planearon ciclos de repetición de 3, 35 y 168 días.



*ERS-2 en el cuarto limpio en centro técnico de la ESA en Noordwijk, en los Países Bajos. Las antenas del SAR y el generador solar se doblan para arriba y se embalan contra el satélite en la "configuración del lanzamiento".*

Figura 3.5.- El segundo lanzamiento de Ariane de 1995 llevará una carga útil muy especial en espacio: ERS-2, el satélite en segundo lugar ambiental de la Agencia Espacial Europea. Éste será el satélite más complejo construido siempre en Europa para observar la tierra del planeta y su ambiente [12 p.e].

#### 3.3.3 ERS-2

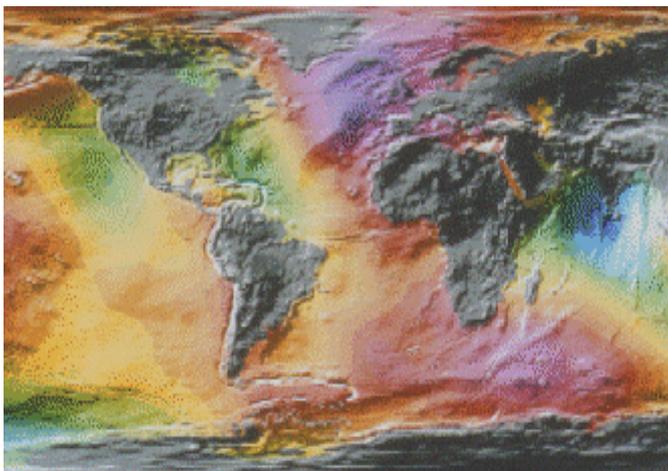
ERS-2 Asume el control donde ERS-1 fue apagado (1995). El segundo satélite ambiental ERS se ocupa de una gama aún más extensa de tareas. Equipado con un instrumento adicional tal como el **GOME** (Ozone Monitoring Equipment); es un espectrómetro que supervisa el contenido de ozono de la atmósfera y monitorea la radiación solar dispersada o transmitida a través de la atmósfera terrestre. El satélite ERS-2 combina un uso más intensivo que su precursor.

La carga de instrumentos del ERS-2 incluye además un radiómetro de exploración equipado de seis canales en las partes infrarrojas y visibles del espectro. Los tres canales para el espectro visible son nuevos y se utilizan para observar la vegetación.



*ERS-2 en el cuarto limpio en centro técnico de la ESA en Noordwijk, en los Países Bajos. Las antenas del SAR y el generador solar se doblan para arriba y se embalan contra el satélite en la "configuración del lanzamiento".*

**Figura 3.5.-** El segundo lanzamiento de Ariane de 1995 llevará una carga útil muy especial en espacio: ERS-2, el satélite en segundo lugar ambiental de la Agencia Espacial Europea. Éste será el satélite más complejo construido siempre en Europa para observar la tierra del planeta y su ambiente.



*Mapa del mundo ERS-2 del nivel del mar.*

**Figura 3.6.-** Mapa del mundo captada por ERS-2.

#### 3.3.4 Parámetros Técnicos ERS-2

Dimensiones:

Altura total: 11,8 m

Generador solar: 11,7 x 2,4 m

Antena del SAR: 10 x 1 m

Masa total: 2516 kilogramos

Las características más notables de ERS-2 se mencionan en la siguiente tabla:

<b>Masa del Satélite</b>	2157 Kg
<b>Potencia generada</b>	2600 watos
<b>Capacidad de batería</b>	2650 watos hora
<b>Altura de la órbita</b>	782 a 785 Km
<b>Inclinación</b>	98.52 °
<b>Periodo de la órbita</b>	Alrededor de 100 minutos
<b>Tiempo de revisita</b>	3, 35 ó 176 según instrumento
<b>Estabilización</b>	Tres ejes
<b>Error de apuntamiento</b>	< 0.21 ° (guiñada)
<b>Determinación de posición</b>	30 m (radial), 15 m (cross), 1000 (along)
<b>Capacidad de almacenamiento</b>	6.5 Gbits
<b>Velocidad de transmisión</b>	105 Mbps
<b>Vida operativa de cada uno</b>	3 años

Tabla 3.3.

#### 3.4.- LANDSAT

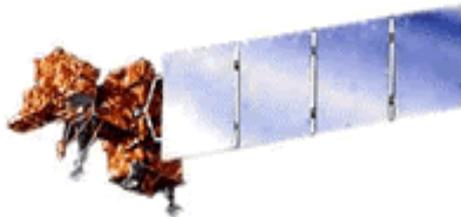


Figura 3.7. Satélite LANDSAT [15 p.e].

El programa **LANDSAT**, fue desarrollado por la **NASA** y originalmente fue denominado Earth Resources Technology Satellite (ERTS). El primer satélite de este programa fue lanzado el 23 de julio de 1972 y se llamó ERTS-1, pero el 14 de enero de 1975 se le cambio el nombre por el de LANDSAT-1 [15 p.e]. Actualmente se han lanzado 7 satélites LANDSAT. Sus fechas de lanzamiento y fin de operación se describen en la siguiente tabla:

<b>SATÉLITE</b>	<b>LANZAMIENTO</b>	<b>FIN OPERACIÓN</b>
<b>ERTS-1</b>	23 de julio de 1972	05 de enero de 1978
<b>LANDSAT-2</b>	22 de enero de 1975	27 de julio de 1983
<b>LANDSAT-3</b>	05 de marzo de 1978	07 de septiembre de 1983
<b>LANDSAT-4</b>	16 de julio de 1982	Finales de 1983
<b>LANDSAT-5</b>	01 de marzo de 1984	En operación
<b>LANDSAT-6</b>	05 de octubre de 1993	Falló la operación
<b>LANDSAT-7</b>	15 de abril de 1993	En operación

Tabla 3.4 Periodo de vida de los satélites LANDSAT [14 p.e].

### Capítulo 3. Principales satélites de percepción remota

Se hará un enfoque principalmente a los dos últimos satélites: LANDSAT 5 y LANDSAT 7.

#### 3.4.1 El Satélite LANDSAT 5

El satélite de observación de la tierra LANDSAT 5 se encuentra orbitando a unos 705 Km. de su superficie. Con él es posible obtener una imagen del mismo lugar cada 16 días. Una imagen comercial completa cubre una superficie de terreno de unos 85 por 85 Km. A medida que el satélite avanza en su órbita (en dirección Norte-Sur, aproximadamente), barre la superficie de la tierra (en sentido Oeste-Este), registrando y transmitiendo a receptores terrestres la información obtenida a partir de la luz solar reflejada por ella.

##### 3.4.1.1 Resolución espectral

Estos sensores están calibrados para "ver" determinados "colores" de luz. Son sensibles sólo determinadas porciones del espectro electromagnético. Cada una de estas porciones o "tipos de luz" se denomina *banda*, discriminando el satélite siete de ellas: tres correspondientes a la luz visible (azul, verde y rojo) y las restantes a luz infrarroja (dos de infrarrojo cercano, una de infrarrojo medio y una de infrarrojo lejano o térmico). Además, puede diferenciar 256 niveles de intensidad en cada una de ellas. De acuerdo a la longitud de onda opera como se indica en la siguiente tabla:

Bandas	Longitud de onda ( $\mu\text{m}$ )	Principales usos
1	0.45 - 0.52	Diseñada para penetración en cuerpos de agua, por lo cual se considera de especial utilidad en el mapeo de aguas costeras. Útil en la diferenciación entre agua y vegetación y entre latifoliadas y coníferas.
2	0.52 - 0.60	Definida para medir el pico de reflectancia y estimar el vigor de la vegetación.
3	0.63 - 0.69	Porción del espectro que es absorbida por la clorofila para la fotosíntesis, importante para la discriminación de los tipos de vegetación.
4	0.76 - 0.90	Utilizada para la determinación de la biomasa vegetal y la delineación de cuerpos de agua.
5	1.55 - 1.75	Por su penetración en el agua, es muy importante en la indicación del contenido de humedad en las hojas y en el suelo. Útil en la diferenciación de nieve y nubes.
6	10.40 - 12.50	Sector del infrarrojo emisivo o termal, utilizado en el análisis de la vegetación, discriminación de contenidos de humedad en el suelo y mapeo termal.
7	2.08 - 2.35	Importante en la discriminación de tipos de rocas y mapeo hidrotérmico.

Tabla 3.5 [16 p.e].

##### 3.4.1.2 Resolución espacial

Cada sensor enfoca a una porción de terreno de unos treinta por treinta metros y todo aquello que refleje luz dentro de esa superficie es "promediado" y registrado como un sólo valor. Esta superficie constituye un píxel y determina la resolución espacial del satélite. No es posible distinguir objetos o características del terreno de tamaño menor a él. Las imágenes de este satélite permiten obtener cartografía a escalas 1:100.000 o

1:80.000. La pérdida de detalle a escalas mayores de 1:50.000 hacen que su uso no sea aconsejable.

#### **3.4.1.3 Análisis digital**

Mediante la aplicación de diversas técnicas en el análisis de los datos de un satélite es posible obtener información útil a diversos campos de la investigación, la industria y la administración. Una de las aplicaciones más comunes de las imágenes satelitales es la representación del terreno para su interpretación visual. La técnica más utilizada por el sistema LANDSAT es la composición en falso color, que consiste en mezclar tres imágenes del mismo sitio y fecha, correspondientes a diversas bandas, representando a cada una de ellas con alguno de los colores rojo, verde o azul. Los 256 niveles digitales de cada banda, representables como una imagen monocromática, se combinan para formar otra imagen en colores.

#### **3.4.2 LANDSAT 7**

Actualmente se tiene un último satélite de esta serie, el denominado LANDSAT 7. Este satélite tiene un peso de 1,969 Kg, perigeo de 701 Km, apogeo de 703 Km e inclinación de 98.2°.

LANDSAT 7 consta, al igual que sus predecesores de los instrumentos anteriores a bordo, aunando además:

**Scanner multiespectral (MSS):** Tiene esta cámara una resolución de 30m para todas sus bandas, excepto para el infrarrojo térmico que consigue resoluciones espaciales de 60m. Este instrumento brinda la posibilidad de realizar mapas temáticos.

**Cámara pancromática (Pan):** Con una resolución de hasta 15 metros, la utilización de esta cámara por los científicos se desarrolla dentro del campo de la cartografía.

Cuenta además con un nuevo sensor denominado ETM+ (Enhanced Thematic Mapper Plus). Su operación es administrada por la NASA (National Space and Space Administration) y la producción y comercialización de imágenes depende del USGS (United States Geological Survey).

Una imagen LANDSAT 7 ETM+ está compuesta por 8 bandas espectrales que pueden ser combinadas de distintas formas para obtener variadas composiciones de color u opciones de procesamiento. Entre las principales mejoras técnicas respecto de su antecesor, el satélite LANDSAT 5, se destaca la adición de una banda espectral (Banda Pancromática) con resolución de 15 metros. También, cuenta con mejoras en las características geométricas y radio-métricas y una mayor resolución espacial de la banda térmica para 60 m. Estos avances permiten emplear datos del LANDSAT 7 en aplicaciones directas hasta una escala de 1:25.000, principalmente, en áreas rurales o territorios de grandes extensiones.

El intervalo espectral y espacial del Satélite LANDSAT 7 se muestra en la siguiente tabla:

### Capítulo 3. Principales satélites de percepción remota

Banda.	Número de rango espectral (micras).	Resolución espacial (m).
1	0.45 a 0.515	30
2	0.525 a 0.605	30
3	0.63 a 0.690	30
4	0.75 a 0.90	30
5	1.55 a 1.75	30
6	10.40 a 12.5	60
7	2.09 a 2.35	30
8 (Pan)	0.52 a 0.90	15

Tabla 3.6 [17 p.e].

**Ancho de barrido** 185 Kilómetros.

**Tiempo de revisita** 16 días (233 órbitas).

**Altitud** 705 Kilómetros.

**Digitalización** 9 bits tomando los 8 mejores.

**Capacidad de almacenamiento** 385 Gb.

**Inclinación** heliosíncrona, 98.2 grados.

**Cruce con el ecuador** Descendente a 10:00 AM + 15'.

**Lanzador** Delta II.

**Día de lanzamiento** Abril de 1999.

El LANDSAT 7 puede adquirir imágenes en un área que se extiende desde los 81° de latitud norte hasta los 81° de latitud sur y, obviamente, en todas las longitudes del globo terrestre. Una órbita del LANDSAT 7 es realizada en aproximadamente 99 minutos, permitiendo al satélite dar 14 vueltas a la Tierra por día, y cubrir la totalidad del planeta en 16 días. La órbita es descendente, o sea de norte a sur, el satélite cruza la línea del Ecuador entre las 10:00 y 10:15 (hora local) en cada pasaje. El LANDSAT 7 está "helio-sincronizado", o sea que siempre pasa a la misma hora por un determinado lugar.

Un factor importante es que el período de revolución del LANDSAT 7 es igual que el del LANDSAT 5 (16 días), y una imagen cubre igual área (185 x 185 km por escena). La conservación de estos parámetros técnicos facilita que el proceso de captura de imágenes se pueda realizar con la misma grilla de referencia (WRS2) lo que permite una perfecta integración entre el procesamiento de las imágenes del LANDSAT 7 con datos históricos del LANDSAT 5 existentes desde 1984. Esto es especialmente útil cuando es necesario utilizar los dos tipos de datos de un mismo lugar en forma simultánea, por ejemplo, para un estudio multitemporal.

#### 3.4.3 Principales diferencias entre el LANDSAT 7 y el LANDSAT 5

- Adición al LANDSAT 7 de una banda pancromática con resolución espacial de 15m.
- Perfeccionamiento del sistema de calibración radio-métrica de los sensores, lo que garantiza una precisión radio-métrica absoluta de 5%.

- Perfeccionamiento de la geometría de captura, lo que brinda una mayor precisión en imágenes corregidas sólo a partir de datos de efemérides de satélite generadas por el GPS de abordo, muy próxima a la precisión obtenida con imágenes georeferenciadas con puntos de control cartográficos.

#### 3.4.3.1 Bandas espectrales y resolución espacial

Las bandas del espectro visible y del infrarrojo mantienen la resolución espacial de 30m del LANDSAT 5 (canales 1, 2, 3, 4,5 y 7). Las bandas del infrarrojo térmico (canales 6L e 6H) pasan a ser adquiridas con resolución de 60 m, contra 120 m del LANDSAT 5. La nueva banda pancromática (canal 8) tiene 15 m de resolución espacial.

El siguiente cuadro comparativo ilustra las diferencias de resolución espectral entre el sensor TM del Landsat5 y el sensor ETM+ del LANDSAT 7. Los valores, expresados en micras, representan los límites de longitudes de onda a los que es sensible cada banda espectral.

Sensor	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 6	Banda 7	Banda 8
TM	0.45 -0.52	0.52 -0.60	0.63 -0.69	0.76 -0.90	1.55 -1.75	10.4 -12.5	2.08 -2.35	
ETM+	0.45 -0.52	0.53 -0.61	0.63 -0.69	0.78 -0.90	1.55 -1.75	10.4 -12.5	2.09 -2.35	0.52 -0.90

Tabla 3.7 [18 p.e].

#### 3.4.3.2 Niveles de corrección geométrica de las imágenes LANDSAT 7

En todas las imágenes las correcciones del sistema son algoritmos de rectificación de la imagen cruda aplicadas automáticamente en la estación de recepción, usando parámetros espaciales contenidos en los archivos descriptores de imagen (datos de posicionamiento y efemérides del satélite), que consiguen minimizar las variaciones espaciales internas presentes en la imagen en su estado bruto, correcciones del ángulo de curvatura terrestre, variaciones de velocidad, altura y actitud del satélite, desplazamientos orbitales, etc. Las imágenes LANDSAT 7 están disponibles en 3 niveles de corrección geométrica:

**Nivel 4:** es un nivel de corrección básica donde la imagen es radio-métrica y geoméricamente corregida de forma sistemática.

**Nivel 5:** también consiste en una imagen con correcciones sistemáticas, pero que se georeferencia utilizando las efemérides del satélite. Los algoritmos de corrección modelan en base a la posición del satélite y la geometría del sensor a través de datos que una computadora de abordo graba sobre la captura. Altitud, efemérides y parámetros de latitud del satélite, descritos en el archivo payload correction data (PCD) y en el archivo calibration parameter file (CPF) son los componentes fundamentales usados para la generación de productos nivel 5.

**Nivel 6:** No es un proceso automático ni una corrección sistemática. Este nivel de procesamiento exige intervención adicional de un operador. La imagen nivel 5 antes descripta es ajustada con puntos de control cartográficos o con puntos de control medidos especialmente con tecnología GPS. Se obtiene una imagen rectificadas a una determinada proyección cartográfica.

**Ortoimagen:** Este nivel de procesamiento exige la intervención de un operador sobre una imagen corregida con puntos de control utilizando también un modelo digital de elevaciones (DEM) para corregir todas las distorsiones. El producto final consiste en una ortoimagen georeferenciada a la proyección cartográfica deseada.

**Imagen de Fusión (Merge):** Es una combinación de buena resolución espacial de la banda Pancromática (15 metros) con la buena resolución espectral de LANDSAT 7. Una correcta interpretación de la información proveniente de imágenes satelitales depende principalmente de los atributos de textura y color presentes en dicha imagen, fundamentalmente para discriminar áreas que presentan variaciones relevantes, como tipos de vegetación y especies, patrones específicos de uso y ocupación del suelo e interpretaciones ligadas a la morfología. Un ejemplo de esto es la figura 3.10.



Figura3.8.- Imagen "Merge" de LANDSAT 7 sobre la zona centro de la ciudad de Buenos Aires [14 p.e].

Las imágenes generadas por el proceso de fusión espectral reúnen en una única imagen diferentes texturas (calidad geométrica) provenientes del canal de mejor resolución espacial (15 m de la banda PAN) y la información temática de color, resultante de la combinación de los diversos canales espectrales.

#### **3.4.3.2 Formatos y soportes existentes**

Las imágenes LANDSAT 7 crudas o derivadas del proceso de fusión están disponibles en formato digital e impreso a requerimiento de los clientes. Cada imagen cubre 185 x 185 Km. (escena completa) pero también es posible fraccionarla en cuadrantes (1/4 de escena, 90 x 90 Km.) o mini escenas de 50 x 50 Km.

### **3.5 Satélite SPOT**

SPOT es el Sistema de observación terrestre diseñado en Francia bajo control del CNES (Centro nacional de estudios espaciales - Francia) y desarrollado con la participación de Bélgica y Suecia. Desde 1986 el CNES ha lanzado 5 satélites SPOT. SPOT Image es líder a nivel mundial en venta de imágenes satelitales y productos derivados de la observación satelital terrestre, con aplicación en agricultura, cartografía, catastro rural, estudios ambientales, planeación urbana, forestería, geología, exploración e ingeniería civil.

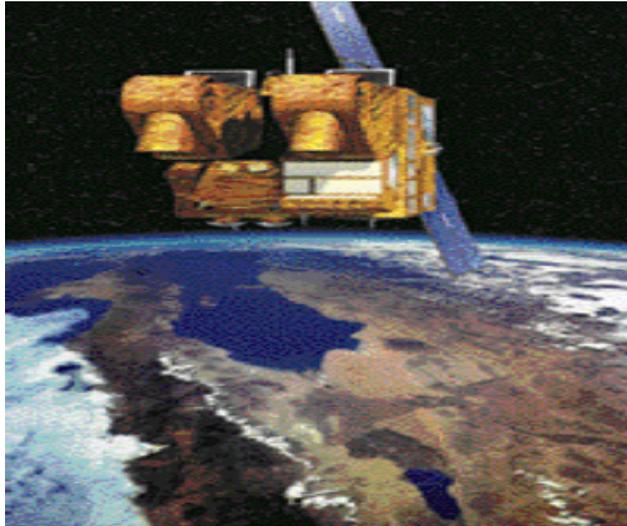


Figura 3.9.- Satélite SPOT [19 p.e].

### 3.5.1 Tres satélites en órbita

El sistema SPOT opera dentro de una constelación de tres satélites que se encuentran en observación, los cuales son: SPOT 1, SPOT 2 SPOT 3.

Los satélites SPOT-1, 2 y 3 usaron el sensor **HVR (High resolution visible)**, el cual es un equipo de exploración por barrido que permite obtener imágenes en dos modalidades: pancromático y multi-banda (verde, rojo e infrarrojo cercano), con una resolución espacial de 10 y 20 m, respectivamente.

El área cubierta en cada escena es de 60 Km., de lado, por lo que estos sensores necesitan incorporar 3000 detectores en modo multi-espectral y 6000 en modo pancromática. A continuación se describen las características de estas 2 modalidades.

	<b>Multiespectral</b>	<b>Pancromática</b>
Bandas espectrales ( $\mu\text{m}$ )	0.50 – 0.59 0.61 - 0.68 0.79 – 0.89	0.51 - 0.73
Tamaño del píxel	20 x 20 m	10 x 10 m
No. de píxeles / línea	3,000	6,000
Ancho de barrido (vista vertical )	60 Km	60 Km

Tabla 3.8 Características de las bandas Multiespectral y Pancromática del sensor HVR [74] [19 p.e].

Se añade a estas características la capacidad de cambiar el campo de visión gracias a un dispositivo móvil instalado en el equipo óptico, que facilita observaciones no verticales de hasta 27° a ambos lados del nadir. Esto permite observar la misma zona en órbitas sucesivas ampliando la resolución temporal de las imágenes.

### **3.5.2 SPOT 4**

El SPOT 4, ofrece una mejor resolución: 2.5 m a 5 m en modo pancromático y 10 metros en modo multi-bandas. El SPOT 4 introduce un novedoso equipo el **HVRI**, se trata de un sensor de alta resolución, sucesor del HVR, que introduce una nueva banda en el infrarrojo medio. Además se sustituye la banda pancromática (0.51-0.73 $\mu$ m) por una nueva banda, la B2 (0.61-0.68 $\mu$ m), capaz de operar en dos modos de resolución 10 y 20 m respectivamente. Esta modificación permite grabar a bordo datos en todas las bandas de frecuencia, para imágenes de lata y baja resolución.

#### **3.5.2.1 Equipos auxiliares**

**PASTE.C.**- Equipo experimental para el estudio del entorno orbital.

**Doris.**- Calcula la posición del satélite en tiempo real mediante un software.

**SPOT** (laser communications passenger).- Láser de estado sólido que permite la comunicación entre diversos satélites en régimen binario.

**ESBT** (Experimental s-band terminal).- Establece un enlace a 4kbits-s, utilizando técnicas de espectro en banda ancha.

**POAM** (Polar ozone and aerosol measurement).- Mide el nivel de ozono y aerosoles sobre la región polar.

**Instrumento de vegetación.**- Orientado al inventario global de vegetación, posee un ángulo de observación de 2,000 Km. y una resolución espacial de 1 Km., además de una alta resolución radiométrica.

### **3.5.3 SPOT 5**

El último satélite de su flota es el SPOT 5. Su capacidad de visión oblicua permite la observación repetida de una misma región sin tener que pasar por la vertical. Las tomas de las vistas estereoscópicas a gran escala son una realidad con el SOT 5. La plataforma SPOT 5 ha derivado de la del SPOT 4. El funcionamiento de 2.5 m es posible gracias a la utilización del nuevo concepto de toma de vista denominado "supermode". El SPOT 5 embarca dos nuevos instrumentos de toma de vista las cámaras HRS y HRG, funcionando en modo pancromático y que apuntan a la vez hacia delante y hacia la parte trasera del satélite. Además de un nuevo sensor estelar ha sido introducido para incrementar el funcionamiento de localización de las imágenes (50 m en lugar de 350 m en los satélites anteriores).

#### **3.5.3.1 La cámara HRG**

El principio de toma de vista se basa en una óptica de campo amplio, que forma la imagen de tierra en una línea de detectores situados en el plano focal. El barrido de línea se obtiene por la lectura seriada de los elementos sensibles de la línea de detección. El barrido del paisaje en la dirección perpendicular resulta del alineamiento del satélite en su órbita. Un separador de espectro permite, además, realizar esta observación en

diferentes ventanas espectrales: posibilita aclarar simultáneamente seis líneas de detectores.

#### **3.5.3.2 La cámara HRS**

El objetivo del instrumento HRS es adquirir las imágenes en la banda pancromática siguiendo dos ejes de vista, respectivamente, a +/- 20, en la parte delantera y trasera del satélite, en relación a su velocidad orbital; también permite adquirir rápidamente un par de imágenes estereoscópicas para elaborar, después del tratamiento en tierra, un modelo digital del terreno. Los flujos de vídeo son comprimidos por un algoritmo cuyos parámetros son adaptados a cada modo de toma de la vista.

#### **3.5.3.3 Capacidad de almacenamiento**

El papel que juega la memoria electrónica es almacenar las imágenes de los instrumentos HRG y HRS para, a continuación, releerlas cuando el satélite está visible desde una estación de recepción en tierra. Con los progresos en la miniaturización de los componentes microelectrónicos las grabaciones magnéticas se han suprimido.

La capacidad de almacenamiento a bordo de las imágenes se ha mejorado a 550 imágenes diarias, en lugar de 400 del SPOT 4, gracias a la utilización de una memoria de masa de 90 Gbits, asociada a una tasa de compresión de datos de 2.6 Gbits. Su gestión por archivos ofrece una gran flexibilidad de uso y permite optimizar la programación. La velocidad de transmisión de datos de imagen hacia tierra es ahora el doble (2x50 Mbits/s).

#### **3.5.3.4 Transmisión a tierra**

Se dispone de una gestión por archivos: las grabaciones pueden ser leídas una o varias veces en las estaciones, en función de la demanda del usuario, independientemente del orden en que han sido almacenados. La velocidad de transmisión de los datos de imagen hacia tierra se ha multiplicado por dos (2x50 Mbits/s), conservando una modulación QPSK. La capacidad de tratamiento de las imágenes (5 canales máximo) se ha superado de forma importante desde el SPOT 4 (2 canales de tratamiento) y se ha introducido una gran flexibilidad debido a que los canales de tratamiento se usan de forma independiente.



Figura 3.10.- Imagen proveniente del satélite SPOT 5 cortesía del CNES [20 p.e].

La carga útil está constituida de los siguientes componentes: un instrumento de toma de vista con cuatro cámaras independientes; un sistema de tratamiento de la imagen que asegura la sincronización de los ritmos y del formateo; un sistema de gestión a bordo encargada de la puesta a punto y de la vigilancia de los equipos; un grabador que autoriza la memorización de 97 minutos de toma de vista; un tele-medidor de la banda X para la transmisión de los datos grabados; un tele-medidor de banda L para la transmisión directa de las tomas de vista y un control de los mecanismos y de la regulación térmica. Estos componentes están integrados en una estructura, fijada en el pie de los HRG al lado del instrumento HRS.

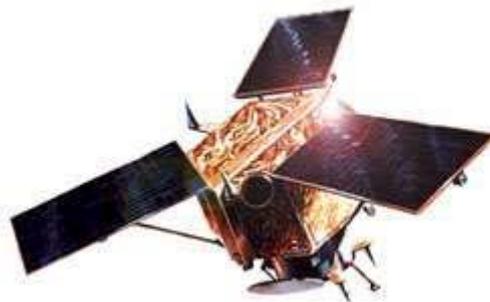
#### **3.5.7 Las bandas espectrales**

La elección de las bandas espectrales de observación de la tierra es resultado de dos criterios: la especificación de la señal recibida en la longitud de onda para analizar la naturaleza de los terrenos, vegetación, desierto, nieve, aguas, ciudades, etc. y la transparencia de la atmósfera y la estabilidad de la transmisión.

### **3.6.- IKONOS**

#### **3.6.1 Disposición de datos de IKONOS**

El satélite IKONOS-2 fue lanzado en septiembre de 1999 y ha estado entregando datos comerciales desde principios de 2000. IKONOS es el primero de la generación de satélites espaciales de alta resolución. Esto significa que el satélite IKONOS es el primer satélite de tipo comercial que posibilita la captación de imágenes con un metro de resolución espacial. El término "IKONOS" proviene del griego y significa "imagen".



**Figura 3.11 Satélite IKONOS [21 p.e].**

IKONOS colecta información de cualquier área en promedio dos veces al día, cubriendo áreas de 20,000 km<sup>2</sup> en una misma pasada y produciendo como resultado imágenes de 1 metro de resolución cada tres días y de 2 m de resolución todos los días.

El satélite IKONOS pesa unos 720 Kg. y órbita la Tierra cada 98 minutos a una altitud de casi 680 Km. en forma sincronizada con el Sol, pasando sobre un determinado lugar aproximadamente a las 10:30 a.m. hora local. La órbita cubierta por el satélite se concentra lejos del área directamente debajo del recorrido del mismo, y los datos de un lugar determinado puede ser captados casi diariamente, si bien no en todos los casos con

1 m de resolución. El sensor digital de imágenes del satélite está diseñado para producir imágenes con elevado contraste, resolución espacial y precisión.

#### 3.6.2. Principales características del satélite IKONOS

- Fecha de lanzamiento del satélite: 24/09/99
- Lugar de lanzamiento: Vandenberg Air Force Base, California /USA
- Altitud: 681 km
- Inclinación: 98.1°
- Velocidad: 7km/s
- Sentido de la órbita: descendente
- Duración de la órbita: 98 minutos
- Tipo de órbita: sincrónica con el sol
- Angulo de vista: rápida alternancia entre diferentes ángulos
- Tiempo de revista: 1 a 3 días
- Resolución en el terreno de cada banda:
  - Pancromática: 1m (considerando posición nominal de 26° para el nadir)
  - Multiespectral: 4m (considerando posición nominal de 26° para el nadir)
- Bandas espectrales:
  - Pan: 0.45 - 0.90  $\mu\text{m}$
  - Azul: 0.45 - 0.52  $\mu\text{m}$
  - Verde: 0.52 - 0.60  $\mu\text{m}$
  - Rojo: 0.63 - 0.69  $\mu\text{m}$
  - Infrarrojo próximo: 0.76 - 0.90  $\mu\text{m}$
- Rango Dinámico: posibilita que la información sea almacenada en 11 bits por píxel, con lo cual redonda en un mayor rango dinámico que facilita el contraste y discriminación de la información. No obstante, los productos pueden ser entregados al usuario en 8 bits por píxel.
- La operación de sus bandas se resume en la tabla 3.9:

	<u>Ancho de banda</u>	Resolución espacial
Pancromático	0,45 - los 0.90 $\mu\text{m}$	1 metro
Banda 1	0,45 - los 0.53 $\mu\text{m}$ (azul)	4 metros
Banda 2	0,52 - los 0.61 $\mu\text{m}$ (verde)	4 metros
Banda 3	0,64 - los 0.72 $\mu\text{m}$ (rojo)	4 metros
Banda 4	0,77 - los 0.88 $\mu\text{m}$ (cerca de infrarrojo)	4 metros

Tabla 3.9 [19 p.e].

La tabla 3.10 muestra distintas imágenes captadas por este satélite:

Ejemplo de las imágenes Sydney Australia.



Ejemplo de la ciudad de Berlín.



Ejemplo del aeropuerto de Francfort, Alemania.



Tabla 3.10 [20 p.e].

Resolución radio-métrica: los datos de IKONOS se recogen como si fuesen pedacitos o segmentos (2048 tonos de grises). Esto significa que hay más definición en los valores de escala grises. Para beneficiarse de esta información adicional, se necesitará un software del proceso de imagen del especialista.

### **3.6.3 Tipos de resoluciones obtenidas por IKONOS**

- **Pancromática (1m PAN):** posibilita a los usuarios distinguir rasgos con dimensiones tan pequeñas como 1 m.
- **Color multiespectral (4m MS):** posibilita a los usuarios distinguir rasgos con dimensiones de 4 m, a partir de información tomada en tres bandas del espectro visible (azul, verde, rojo) e infrarrojo.
- **Pan-sharpened (1m PS):** este producto combina digitalmente las imágenes pancromáticas de 1 metro de resolución con las imágenes multiespectrales de 4 metros, resultando un nuevo producto que presenta la ventaja de contar con una resolución de 1 m y con la alta resolución espectral (mayor discriminación) de las bandas del visible y/o infrarrojo, lo cual la convierte en un producto muy superior a cualquier imagen de los satélites actualmente operacionales.

### **3.6.4 Niveles de procesamiento**

- **IKONOS Geo:** es un producto que define el futuro de las imágenes de alta resolución; está disponible tanto para imágenes pancromáticas (1 m de resolución) como para imágenes color o multiespectrales de 4 m de resolución, lo cual resulta excelente para una variedad de análisis y aplicaciones que requieren el uso de las bandas multiespectrales, para diversas aplicaciones.
- **IKONOS orto-rectificadas:** se trata de imágenes con precisiones métricas del área o zona para la construcción de mapas . Comparativamente, son productos menos costosos y que a su vez permiten la obtención más rápida y eficaz que las ya conocidas y tradicionales fotografías aéreas. Los procesos de orto-

rectificación remueven las distorsiones introducidas por la variabilidad y geometría de relieve y re-muestran las imágenes en una proyección de mapas escogida por el usuario. Incluye el proceso de mosaicos de diferentes imágenes.



Figura 3.12.- Imagen IKONOS 1m de un sector de la hidroeléctrica Itaipú Binacional [22 p.e].

#### 3.6.5 Aplicaciones y ventajas de las imágenes IKONOS

Por el hecho de ser IKONOS el primer satélite comercial que brinda un producto de 1 m de resolución espacial, se tiene la posibilidad de definir los tipos de estándares para imágenes de alta resolución. A través de estas imágenes, ofrece más que simples datos, ya que brinda valiosa información que ayudará a los usuarios a tomar decisiones sobre la base de un cubrimiento diario.

A través de la utilización de productos elaborados por IKONOS y otros satélites tales como LANDSAT, actualmente es posible que:

- Los agricultores pueden monitorear con mayor precisión la condición y vitalidad de sus cultivos y predecir con mayor acierto sus volúmenes de cosecha; además, pueden prevenir problemas y ahorrar importantes sumas de dinero a través de su determinación en estados tempranos.
- Los científicos ambientalistas pueden predecir tendencias en áreas de elevada fragilidad ambiental.
- Los planeadores urbanísticos pueden evaluar los avances de planes comunales de viviendas y las compañías de seguros pueden medir y mapear daños a propiedades luego de desastres naturales.
- Los geólogos ya no necesitarán recurrir a vuelos fotográficos para interpretaciones geológico-estructurales. A partir de la posibilidad que ofrece IKONOS de generar productos con elevado nivel de resolución y recubrimiento estereoscópico, estas imágenes resultan comparativamente más ventajosas que las fotografías aéreas.

Entre las ventajas merecen citarse:

### **Capítulo 3. Principales satélites de percepción remota**

---

- 1.- Mejor calidad métrica y geométrica. Con IKONOS es posible obtener una ortoimagen que cubra totalmente el área de interés, a partir de sólo unos pocos puntos de control terrestre, reduciendo así considerablemente el tiempo de procesamiento.
- 2.- No requieren procesos de mosaicos.
- 3.- Posibilitan la visualización e interpretación estereoscópica tradicional, más la posibilidad de realizar estéreo análisis digital
- 4.- La interpretación estereoscópica digital facilita el y la integración de la información en una base de datos georeferenciada, posibilitando la superposición y el modelado espacial.
- 5.- Permite la utilización de diferentes combinaciones de bandas espectrales, incluyendo productos en infrarrojo, lo que aumenta considerablemente las capacidades de diferenciación y discriminación.
- 6.- Posibilitan la obtención de cartografía de alta calidad en diferentes escalas y combinaciones de bandas, a partir de composiciones de mapas que incluyen diferentes capas temáticas superpuestas, símbolos cartográficos, leyendas, etc.

# 4 Segmentación Bayesiana

## 4.1 Introducción

En el análisis de una imagen es fundamental realizar un reconocimiento de patrones de un objeto o de una entidad de interés en una imagen. Los métodos de decisión teórica para el reconocimiento de patrones se basan en la utilización de *funciones de decisión* también llamadas *funciones discriminantes de Bayes*; basándose en la representación de patrones en forma vectorial y en la búsqueda posterior de aproximaciones que permitan agrupar y asignar estos patrones a las diferentes clases [1 txt].

El problema básico en el reconocimiento de patrones por decisión teórica es el de encontrar las funciones de decisión que tengan la propiedad de pertenencia de un patrón a determinada clase. Una clase de patrones es una familia de patrones que comparten algunas propiedades comunes.

Trataremos de discutir la función del discriminante estadístico, usada para la clasificación de clases, en las cuales la categoría de patrones es conocida en una prioridad, de modo que sea caracterizable por un conjunto de parámetros. Primeramente, analizaremos la formulación del problema por medio de la decisión estadística y funciones de pérdidas, en base a discriminantes de Bayes, así como la decisión de máxima probabilidad. También se muestran algunos análisis de probabilidad de error.

La función discriminante óptima considera un modelo probabilístico con distribución normal, el cuál será discutido a detalle a lo largo de este capítulo. La función óptima será implementada para propósitos de segmentación.

El uso de la función discriminante estadística para segmentación, presenta varias ventajas, puesto que se consideran aspectos que a continuación se discuten:

1).- Existe un considerable conocimiento alrededor de áreas, como comunicación estadística, teoría de la información, teoría de la decisión, teoría de la detección, y así sucesivamente, y en este conocimiento es directamente aplicable el reconocimiento de patrones.

2).- La formulación estadística es particularmente conveniente para el problema de reconocimiento de patrones, puesto que varios procesos son modelados de manera estadística.

En un sistema de clasificación estadística se asume tanto una función de distribución de densidad, así como una distribución gaussiana ó alguna otra función de distribución.

### 4.2 Esquemas de segmentación

**Teorema 1** (Regla de Bayes)

Sea  $S$  un espacio muestral,  $P$  una medida de probabilidad en  $S$  y  $B$  un evento en  $S$ . Sea  $A_1, A_2, \dots, A_n$  una partición de  $S$ , entonces para cada  $i = 1, 2, 3, \dots, n$  tenemos

$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{\sum_{j=1}^n P(A_j)P(B | A_j)} \dots\dots\dots(4.1)$$

**Prueba**

De la definición de probabilidad condicional tenemos  $P(A_i | B) = \frac{P(B \cap A_i)}{P(B)}$ . Para calcular el numerador usamos la definición de nuevo y así obtenemos que  $P(B \cap A_i) = P(B | A_i)P(A_i)$ . El denominador se va a obtener aplicando la fórmula de probabilidad total, se puede obtener el resultado deseado.

$$P(A_i | B) = \frac{P(B \cap A_i)}{P(B)} = \frac{P(A_i)P(B | A_i)}{P(B)} = \frac{P(A_i)P(B | A_i)}{\sum_{j=1}^n P(A_j)P(B | A_j)} \dots\dots\dots(4.2)$$

La expresión (4.2) es el teorema de Bayes, que como se observa es una simple probabilidad condicional, pero que será de gran ayuda para la decisión a la hora de segmentar la imagen [24 p.e].

#### 4.2.1. Etapas de la segmentación

La segmentación es una de las tareas más difíciles del procesamiento de imágenes, la implementación de esta tarea determina el eventual éxito o fracaso del análisis. De hecho la segmentación rara vez llega a alcanzar una solución satisfactoria. Por esta razón se debe poner un considerable cuidado en aumentar la probabilidad de tener una segmentación robusta y adecuada, de acuerdo a las características permitidas que proporciona la imagen a tratar; para que de este modo sea posible el hecho de dividirla en regiones lo más exacto que sea posible.

Las etapas de segmentación tienen que ver con el hecho de tomar una probabilidad establecida o dada para el número de clases denominada como pre-segmentación y a su vez con el hecho de actualizar todas y cada una de las probabilidades; lo cual se refiere a la clasificación de la imagen, por medio de la función discriminante óptima.

La actualización de las probabilidades se lleva a cabo dividiendo el número de píxeles de la ventana de entrenamiento que especifica una clase, entre el número de píxeles total de la imagen.

##### 4.2.1.1 Presegmentación Bayesiana

a).- Sobre la imagen a trabajar, se van a definir  $k$  número de ventanas, las cuales identificarán a  $k$  clases; de este modo se podrá facilitar distinguir las estructuras o etiquetas de la imagen a trabajar (vegetación, agua, pavimento, construcción, etc.).

b).- Sobre cada una de las ventanas definidas se van a calcular la media, varianza, desviación estándar, covarianza, y la correspondiente función discriminante de cada etiqueta.

c).- Se aproximan las probabilidades conjuntas del numerador de la regla de Bayes mediante las funciones mostradas en la figura 4.2, las cuales indican la probabilidad condicional de que un píxel dado pertenezca a la clase  $\omega_k$  en niveles de gris.

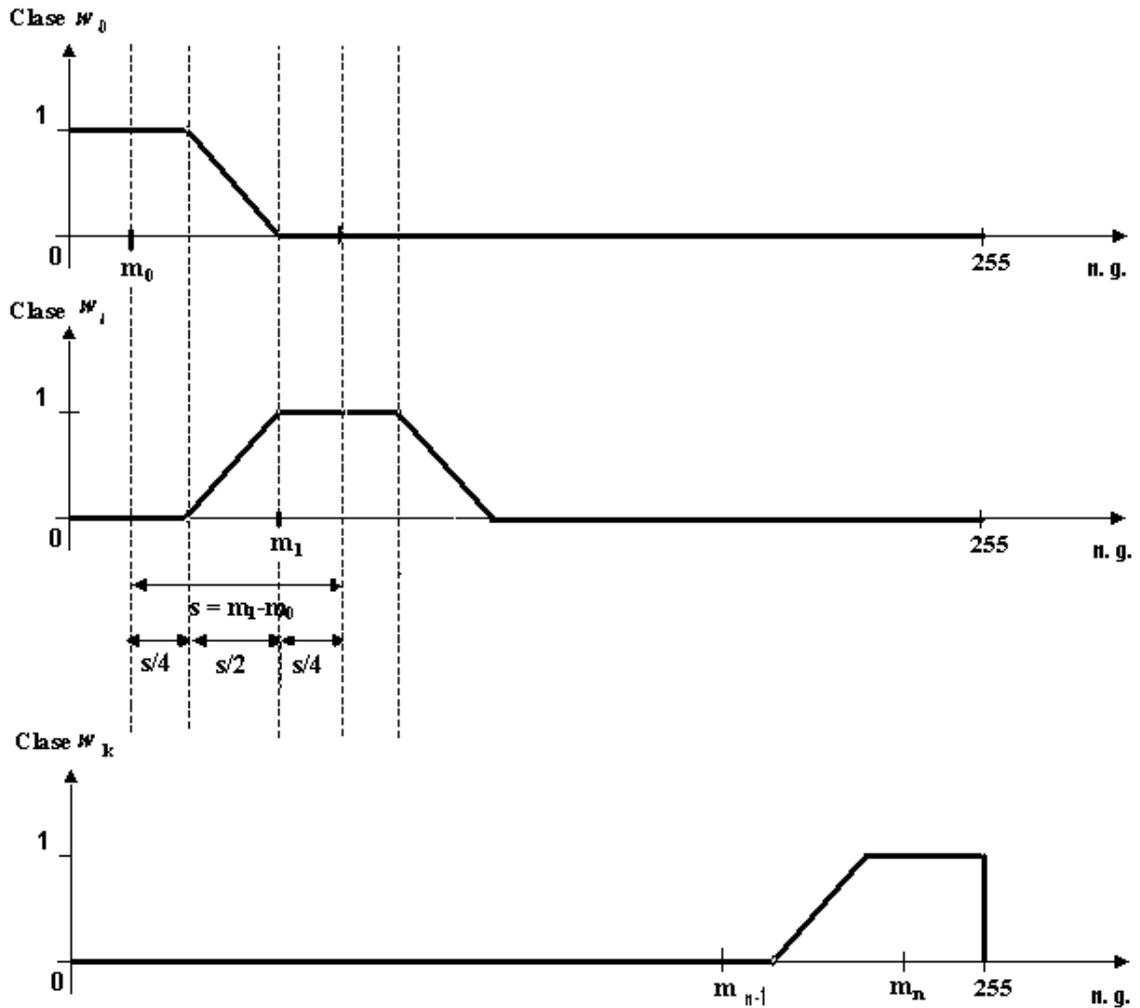


Figura 4.2.- Funciones de probabilidad condicional  $P(X|W_k)$  [8 txt].

#### 4.2.1.2 Presegmentación según el criterio de Bayes

1. Empleando la función discriminante de Bayes:

$$d_k(x) = p(x | w_k)p(w_k)$$

donde:

$x$  es un píxel de la imagen original

$w_k$  las clases con  $k=0,1,2,\dots$

$p(w_k)$  es la probabilidad a priori de la clase  $w_k$

$p(x | w_k)$  es la probabilidad condicional de  $x$  dado  $w_k$  y  $d_k(x) = p(w_k | x)$  representa la probabilidad a posteriori

2. Como no se conoce la probabilidad a priori  $p(w_k)$ , se consideran equiprobables a las probabilidades de ocurrencia de las etiquetas  $p(w_k)$ .

3. Se analiza cada uno de los píxeles de la imagen, etiquetando a cada píxel de acuerdo con los valores máximos de la probabilidad a:  $\arg \max[p(w | x)]$ .

#### 4.2.1.3 Segmentación Bayesiana

En esta etapa, se utilizan las probabilidades a priori  $p(w_k)$  actualizadas. Ahora, teniendo la actualización de las probabilidades, se clasifica cada uno de los píxeles de la imagen utilizando la misma función bayesiana [8 txt].

Durante el proceso de clasificación, el nivel de gris de cada píxel se proyecta sobre cada una de las funciones para determinar la probabilidad condicional de que el píxel pertenezca a la clase  $w_i$ . El número de funciones está determinado por el número de clases, definidas en las ventanas de entrenamiento.

De acuerdo a lo establecido en la Figura 4.2, dado que  $w_0 = \{x | 0 \leq x \leq m_0\}$  si el nivel de gris del píxel analizado, esta en  $0 \leq x \leq m_0$ , la probabilidad condicional para la clase  $w_0$  es de 1. Para las demás clases, la probabilidad será cero.  $m_0$

Si el nivel de gris del píxel analizado se encuentra en el intervalo  $w_0 \geq m_n$ , la probabilidad condicional para la clase  $w_n$  será de 1. Para las demás clases la probabilidad será de 0.

Si el nivel de gris del píxel analizado no se encuentra en ninguno de los 2 intervalos anteriores, se realiza la diferencia de las medias de las clases:  $S = m_s - m_a$ , donde  $m_s$  es la media de la clase siguiente y  $m_a$  es la media de la clase actual. El resultado de esta diferencia será dividido entre 4 sea  $S/4$ .

Si el nivel de gris del píxel analizado se encuentra en el rango  $m_a \leq w_0 \leq m_a + \frac{S}{4}$ , la probabilidad condicional para la clase  $w_a$  será de 1. Para las demás clases la probabilidad será de 0.

Si el nivel de gris del píxel analizado está en el rango  $m_0 + s/4$  y  $m_0 + 3*s/4$ , entonces existen 2 valores de probabilidad condicional para este píxel, ya que como se observa en la figura 4.2, este píxel puede pertenecer a la clase  $w_0$  o a la clase  $w_1$ . Para obtener la probabilidad condicional de que el píxel analizado pertenezca a la clase  $w_0$ , se utiliza la siguiente función:

$$w_0 = \frac{-2}{S}x + \frac{2}{S}m_0 + \frac{3}{2}$$

que será la pendiente negativa y

$$W_0 = \frac{2}{S}x - \frac{2}{S}m_a - \frac{1}{2} \text{ que será la pendiente positiva.}$$

donde:

- $W_0$  es la probabilidad condicional de la clase  $w_0$
- $x$  es el píxel analizado.
- $m_0$  es la media de  $W_0$
- $S = m_1 - m_0$

Después de calcular las dos funciones anteriores, se determinará que el píxel analizado pertenecerá a la clase que tenga una probabilidad condicional mayor.

Si el nivel de gris del píxel analizado es mayor o igual a  $m_1 - s/4$  y menor o igual a  $m_0 + s/4$ , entonces la probabilidad condicional será de 1.

Para las clases  $w_i$ , se realiza un análisis similar.

Después de obtener todas las probabilidades condicionales  $P(x | w_k)$ , se multiplica cada una de ellas por la probabilidad de ocurrencia, que en la primera etapa se consideran como equiprobables  $p(w_k) = \frac{1}{k}$ , y en la segunda se actualizan. El píxel analizado pertenecerá a la clase cuya probabilidad a posteriori sea mayor.

#### **4.2.2 Formulación del problema por medio de la teoría de diseño estadístico**

##### **4.2.2.1 Función de pérdidas**

- 1).-  $p(w_i)$  es conocida o puede ser estimada.
- 2).-  $p(x|w_i)$  es conocida o puede ser estimada directamente por el conjunto.
- 3).-  $p(w_i|x)$  es generalmente desconocida.

Donde  $p(w_i)$  es la probabilidad a priori de clases  $w_i$ , y  $p(x|w_i)$  es la función de probabilidad de **la etiqueta  $w_i$** , o el estado de la función de densidad de probabilidad para  $x$ , dado que el estado natural es  $w_i$ . Esto es lo que se llama una probabilidad posterior [2 txt (cap. 3)].

Cuando se define un píxel dado a una clase cuando este pertenece a otra, se origina una función de pérdida  $L_{ij}$  puede ser definida como la pérdida, riesgo o penalización para la decisión de que  $x \in w_i$ . Así se minimizar el promedio de pérdidas. Similarmente el promedio de pérdida condicional o el riesgo promedio condicional  $r_k(x)$  puede ser definido como:

$$r_k(x) = \sum_{i=1}^M L_{ik} p(w_i|x) \dots \dots \dots (4.3)$$

la expresión (4.3) es el promedio o pérdida esperada de error de clasificación  $x$  dado en  $w_k$ , pero de hecho puede ser en algunas otras clases  $w_i, i = 1, 2, \dots, M$  y  $i \neq k$ .

La tarea de la clasificación es entonces, encontrar una decisión óptima que pueda minimizar el riesgo promedio o costo. La regla de la decisión consiste en los siguientes pasos:

- 1).- Calcular las pérdidas esperadas,  $r_i(x)$  de decisión tal que  $x \in w_i \forall i = 1, 2, \dots, M$ .
- 2).- Decidir que  $x \in w_k$  si  $r_k(x) \leq r_i(x) \forall i, i \neq k$ .

La función discriminante correspondiente es entonces:

$$d_k(x) = -r_k(x) \dots \dots \dots (4.4)$$

El signo (-) de  $r_k(x)$  es elegido tal que  $d_k(x)$  representa la mayor clase probable. El más insignificante  $r_k(x)$  representa la mayor probabilidad tal que  $x \in w_k$ .

Una matriz de pérdida puede ser expresada como:

$$L = \begin{vmatrix} 0 & & & \\ & 0 & 1 & \\ & & . & \\ 1 & & & 0 \end{vmatrix}$$

Donde  $L_{i,j} = 0, i = 1, \dots, M$ , ya que no ocurre error de clasificación en tales casos; mientras que para  $L_{i,k} = 1$ , hay una penalización en la clasificación  $x \in w_k$ , debido a que  $x \in w_i, i = 1, \dots, M, i \neq k$ . Esto que esta es una función de pérdida simétrica

$$L_{i,k} = 1 - \delta(k-i)$$

donde  $\delta(k-i)$  es la función delta de Kronecker:

$$\delta(k-i) = \begin{cases} 1 & \text{si } i = k \\ 0 & \text{si otro caso} \end{cases}$$

$$\text{si el valor de } L_{ik} = \begin{cases} -h_i & i = k \\ 0 & i \neq k \end{cases}$$

La matriz de pérdida conviene expresarla en matriz negativa de una función de pérdida:

$$L = \begin{vmatrix} -h_1 & & & \\ & -h_2 & & 0 \\ & & -h_3 & \\ 0 & & & -h_M \end{vmatrix}$$

Lo significativo de esta matriz de pérdida negativa es que una pérdida (-) (i.e. una ganancia positiva) es que las pérdidas asignadas para una decisión serán más grandes para un error de decisión que para uno correcto.

### 4.3 Función discriminante de Bayes

Por regla de Bayes podemos escribir

$$p(w_i|x) = \frac{p(x|w_i)p(w_i)}{p(x)} \dots\dots\dots(4.5)$$

donde  $p(x) = \sum_i p(x|w_i)p(w_i)$ , con  $i = 1,2,\dots,M$ , es la probabilidad que  $x$  ocurra sin hacer caso de la categoría en la que pertenece.  $p(w_i)$  es la probabilidad a priori de clase  $w_i$ , y  $p(x|w_i)$  es la función de probabilidad de clase  $w_i$  con respecto a  $x$ ; esto es la función de densidad de probabilidad para  $x$  dado que el estado natural es  $w_i$ , (i.e. es un modelo perteneciente a la clase  $w_i$ ).

Sustituyendo ec. 4.5 en 4.3 para  $r_k(x)$ , tenemos:

$$r_k(x) = \frac{1}{p(x)} \sum_{i=1}^M L_{ik} p(w_i|x) p(w_i) \dots\dots\dots(4.6)$$

puesto que  $p(x)$  es común para toda  $r_j(x)$ ,  $j = 1,2,\dots,M$ , podemos eliminarla de la ecuación de riesgo promedio condicional y buscar solamente el mínimo

$$\min_k r_k(x) = \min_k \sum_{i=1}^M L_{ik} p(w_i|x) p(w_i) \dots\dots\dots(4.7)$$

para poder obtener el mejor resultado entre todas las posibles decisiones, alternativamente podemos decir que:

$$d_k(x) = - r_k(x) \dots\dots\dots(4.8)$$

definida como la *función del discriminante de Bayes*, que provee el funcionamiento óptimo para el punto de vista estadístico.

4.3.1 Probabilidad máxima de decisión

La expresión para el promedio o pérdida esperada de decisión  $x \in w_k$  es

$$r_k(x) = \sum_{i=1}^M L_{i,k} p(w_i|x) p(w_i) \dots \dots \dots (4.9)$$

la cuál puede ser usada para la minimización, obteniendo la máxima probabilidad para  $x \in w_k$ . Para un problema de 2 clases el promedio o pérdida esperado tanto para  $x \in w_1$  y similarmente  $x \in w_2$ , la pérdida de decisión puede ser:

$$r_1(x) = L_{11} p(x|w_1) p(w_1) + L_{21} p(x|w_2) p(w_2) \dots \dots \dots (4.10)$$

$$r_2(x) = L_{12} p(x|w_1) p(w_1) + L_{22} p(x|w_2) p(w_2)$$

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} L_{11} & L_{21} \\ L_{12} & L_{22} \end{bmatrix} \begin{bmatrix} p(x|w_1)p(w_1) \\ p(x|w_2)p(w_2) \end{bmatrix} \dots \dots \dots (4.11)$$

La decisión que  $x \in w_1$  puede hacerse si

$$(L_{21} - L_{22}) p(x|w_2) p(w_2) < (L_{12} - L_{11}) p(x|w_1) p(w_1) \dots \dots \dots (4.12)$$

La desigualdad de arriba puede ser puesta de otra forma:

$$\theta_{12} = \frac{p(x|w_1)}{p(x|w_2)} > \frac{(L_{21} - L_{22}) p(w_2)}{(L_{12} - L_{11}) p(w_1)} = \gamma_{ki} \dots \dots \dots (4.13)$$

donde  $\theta_{12}$  es el cociente de probabilidad que refleja la *regla de máxima verosimilitud* y  $\gamma_{ki}$  es el valor umbral, entonces el criterio para la decisión se convierte en

$$x \in w_1 \text{ si } \gamma_{ki}(x) > \theta_{12} \dots \dots \dots (4.14)$$

Lo anterior puede generalizarse para problemas multiclases, ( $M > 2$ ), entonces:

$$\gamma_{ki} = \frac{p(x|w_k)}{p(x|w_i)} \dots \dots \dots (4.15)$$

y

$$\theta_{ki} = \frac{(L_{ik} - L_{ii}) p(w_i)}{(L_{ki} - L_{kk}) p(w_k)} \dots \dots \dots (4.16)$$

definiendo la consideración para la decisión de la forma:

$$x \in w_k \text{ si } \gamma_{ki}(x) > \theta_{ki} \quad \forall i, \dots \dots \dots (4.17)$$

La expresión (4.17) es lo que se llama regla de máxima probabilidad.

Para una función L de pérdida simétrica, tenemos  $x \in w_l$  si  $\gamma_{li} > \theta_{ki} \forall i, i = 1, \dots, M$ . La regla de máxima probabilidad se convierte en:

$$\frac{p(x|w_k)}{p(x|w_i)} > \frac{p(w_i)}{p(w_k)} \dots\dots\dots(4.18)$$

es decir que  $L_{ik} = 1$  y  $L_{ii} = 0 \forall i, e i \neq k ; i, k = 1, \dots, M$ . Si  $p(w_i) = p(w_k) \forall i, k$ , por lo tanto la regla de máxima probabilidad llega a ser:

$$\text{Asignando que } x \in w_k \text{ si } \gamma_{ki} > 1 \dots\dots\dots(4.19)$$

En un caso más general, cuando  $p(w_i) \neq p(w_k)$ , se debe construir la función discriminante para el caso de una función de pérdida simétrica

$$L_{ik} = \frac{p(x|w_k)}{p(x|w_i)} > \frac{p(w_i)}{p(w_k)} \dots\dots\dots(4.20)$$

entonces

$$p(x|w_k) p(w_k) > p(x|w_i) p(w_i) \forall i, \text{ de esta forma la función discriminante es:}$$

$$d_k(x) = p(x|w_k) p(w_k) \dots\dots\dots(4.21)$$

o de otra forma

$$d_k(x) = \log p(x|w_k) + \log p(w_k) \dots\dots\dots(4.22)$$

La regla de máxima probabilidad es entonces:

$$x \in w_i \text{ si } r_i(x) < r_j(x) \dots\dots\dots(4.23)$$

o

$$\sum_{k=1}^M L_{ki} p(x|w_k) p(w_k) < \sum_{q=1}^M L_{qi} p(x|w_q) p(w_q) \dots\dots\dots(4.24)$$

$$\forall j, j \neq i; j=1, \dots, M.$$

donde  $\sum_{k=1}^M L_{ki} p(x|w_k) p(w_k)$  representa la pérdida promedio de decisión si  $x \in w_i$  y

$\sum_{q=1}^M L_{qi} p(x|w_q) p(w_q)$  representa la pérdida promedio de decisión si  $x \in w_j$ .

### 4.3.2 Probabilidad de error

Tomando en consideración el problema de las 2 clases, se pretende hacer una **clasificación** cuyo objetivo es dividir el espacio en 2 regiones,  $R_1$  y  $R_2$ . La decisión de que  $x \in w_1$  puede ser hecha cuando el patrón  $x$  cae dentro de la región  $R_1$ ; y  $x \in w_2$  cuando el patrón  $x$  cae dentro de la región  $R_2$ . Bajo tales circunstancias se tendrán 2 tipos de valores posibles:

1).-  $x$  cae dentro de la región  $R_1$ , pero realmente  $x \in w_2$ . Esto da la probabilidad de error  $E_1$ , el cual puede ser denotado por  $Prob(x \in R_1, w_2)$ .

2).-  $x$  cae dentro de la región  $R_2$ , pero realmente  $x \in w_1$ . Esto da la probabilidad de error  $E_2$ , el cual puede ser denotado por  $Prob(x \in R_2, w_1)$ . Así la probabilidad total de error es:

$$P_{error} = Prob(x \in R_1|w_2)p(w_2) + Prob(x \in R_2|w_1)p(w_1)$$

$$= \int_{R_1} p(x|w_2)p(w_2)dx + \int_{R_2} p(x|w_1)p(w_1)dx \dots\dots\dots(4.25)$$

La expresión anterior nos permite minimizar para una buena clasificación. Para dar una expresión analítica de la probabilidad de error, vamos a asumir funciones de densidad normal (FDN) multivariada para los modelos de vectores con  $C_1 = C_2 = C$ ; así

$$p(x|w_1) = \frac{1}{[(2\pi)^{n/2} |C|^{1/2}]} \exp\left[-\frac{1}{2}(x-m_1)^T C^{-1}(x-m_1)\right]$$

y .....(4.26)

$$p(x|w_2) = \frac{1}{[(2\pi)^{n/2} |C|^{1/2}]} \exp\left[-\frac{1}{2}(x-m_2)^T C^{-1}(x-m_2)\right]$$

para el caso en que las funciones son simétricas tenemos:

si  $x \in w_2 \Rightarrow \frac{p(x|w_1)}{p(x|w_2)} > \frac{p(w_2)}{p(w_1)} = \theta'_{12}$ ;

similarmente si  $x \in w_1 \Rightarrow \frac{p(x|w_2)}{p(x|w_1)} > \frac{p(w_1)}{p(w_2)} = \theta'_{21}$ .

sustituyendo las FDN multivariadas se obtiene:

$$\frac{p(x|w_1)}{p(x|w_2)} = \frac{\exp\left[-\frac{1}{2}(x-m_1)^T C^{-1}(x-m_1)\right]}{\exp\left[-\frac{1}{2}(x-m_2)^T C^{-1}(x-m_2)\right]} \dots\dots\dots(4.27)$$

Rescribiendo la ecuación

$$p_{12} = \frac{p(x|w_1)}{p(x|w_2)} = -\frac{1}{2}(x-m_1)^T C^{-1}(x-m_1) + \frac{1}{2}(x-m_2)^T C^{-1}(x-m_2)$$

$$= x^T C^{-1}(m_1-m_2) - \frac{1}{2}(m_1+m_2)^T C^{-1}(m_1-m_2) \dots\dots\dots(4.28)$$

Entonces la probabilidad será asignada como:

$$Prob(x \in R_1, w_2) = p[p_{12} > \log \theta'_{12}|w_2]$$

$$\dots\dots\dots(4.29)$$

Pr ob( $x \in R_2, w_1$ ) =  $p[p_{12} > \log \theta'_{12} | w_1]$

El valor esperado de  $p_{12}$  para la clase 1 resulta

$$E_1[p_{12}] = \frac{1}{2} \left[ (m_1 - m_2)^T C^{-1} (m_1 - m_2) \right] \dots\dots\dots(4.30)$$

La varianza de  $p_{12}$  para la clase 1, es definida como:

$$\begin{aligned} Var_1[p_{12}] &= E_1[(p_{12} - \overline{p_{12}})^2] \\ &= E_1[(x - m_1)^T C^{-1} (m_1 - m_2)]^2 \\ &= E_1[(m_1 - m_2)^T C^{-1} (x - m_2) (x - m_1)^T C^{-1} (m_1 - m_2)]^2 \\ &= (m_1 - m_2)^T C^{-1} E_1[(x - m_2)(x - m_1)^T C^{-1} (m_1 - m_2)] \\ &= (m_1 - m_2)^T C^{-1} E_1[(x - m_2)(x - m_1)] C^{-1} (m_1 - m_2) \dots\dots\dots(4.31) \end{aligned}$$

puesto que por definición  $C = E_1[(x - m_2)(x - m_1)] \therefore$

$$Var_1[p_{12}] = (m_1 - m_2)^T C^{-1} (m_1 - m_2) = r_{12} \dots\dots\dots(4.32)$$

Por consiguiente el valor esperado será:

$$E_1[p_{12}] = \frac{1}{2} r_{12} \dots\dots\dots(4.33)$$

Donde  $r_{12}$  representa la distancia entre  $p(x|w_1)$  y  $p(x|w_2)$  conocida también como la distancia de Mahalanobis. Así, para  $x \in w_1$ , el cociente  $p(x|w_1)/p(x|w_2)$  está distribuido con un radio  $(1/2)r_{12}$  y una varianza igual a  $r_{12}$ ; mientras que para  $x \in w_2$  es distribuido con un radio igual a  $(-1/2)r_{12}$  y la varianza igual a  $r_{12}$ . Como un precedente, cabe hacer notar que el valor de la varianza no cambia.

### 4.3.3 Funciones discriminante óptimas para modelos distribuidos normalmente

#### 4.3.3.1 Distribución normal

La función de densidad normal multivariada para  $M$  clases puede ser representada por:

$$p(x|w_k) = \frac{1}{\left[ (2\pi)^{n/2} |C_k|^{1/2} \right]} \exp \left[ -\frac{1}{2} (x - m_k)^T C_k^{-1} (x - m_k) \right] = N(m_k, C_k) \dots\dots\dots(4.34)$$

$k = 1, 2, \dots, M$

$n$  = dimensión del vector modelo.

Asumiendo que  $N$  es la función de densidad normal,  $m_k$  es el vector promedio que determina el centro del grupo, y  $C_k$ , es la matriz de covarianza que determina la forma

del grupo para clases  $k$ , definida respectivamente por los valores esperados sobre los patrones ó modelos pertenecientes para la clase  $k$ . Así

$$m_k = E_k[x] \text{ y } C_k = E_k[(x-m_k)(x-m_k)^T] \dots\dots\dots(4.35)$$

De todo lo anterior se deduce que una medida útil es la distancia de Mahalanobis ( $r^2$ ) [2 txt (cap 3)] desde el patrón o modelo  $x$  al medio  $m$ , la cual puede ser definida como:

$$r^2 = (x-m)^T C^{-1} (x-m) \dots\dots\dots(4.36)$$

La distancia de Mahalanobis ( $r^2$ ) entre dos clases puede ser expresada como:

$$r^2 = (m_i - m_j)^T C^{-1} (m_i - m_j) \dots\dots\dots(4.36b)$$

Cabe mencionar que si  $n = 1$  aproximadamente 95% de las muestras  $x$  caen en la región  $|x - m| < 2\sigma$ ,  $\sigma$  representa la desviación estándar y es igual a  $C^{1/2}$ .

Sabemos que la función discriminante para  $x \in w_k$  puede ser puesta en la siguiente forma:

$$d^i_k(x) = \log p(x|w_k) + \log(w_k) \dots\dots\dots(4.37)$$

Con la anterior función discriminante, aplicada a la densidad normal multivariada para una clase de  $M$ -patrones ó  $M$ -modelos, como sigue

$$p(x|w_k) = \frac{1}{\left[ (2\pi)^{n/2} |C_k|^{1/2} \right]} \exp \left[ -\frac{1}{2} (x-m_k)^T C_k^{-1} (x-m_k) \right] \quad k = 1, 2, \dots, M$$

la función del discriminante  $d^i_k(x)$  se hace

$$d^{ii}_k(x) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|C_k| - \frac{1}{2} (x-m_k)^T C_k^{-1} (x-m_k) + \log p(w_k) \dots\dots\dots(4.38)$$

como el primer término es el mismo para toda  $k$ , puede ser eliminado. Entonces la función discriminante se reduce a

$$d^{iii}_k(x) = -\frac{1}{2} (x-m_k)^T C_k^{-1} (x-m_k) + [\log p(w_k) - \frac{1}{2} \log|C_k|] \dots\dots\dots(4.39)$$

Esta es una función discriminante cuadrática, y puede ser puesta en forma mas compacta como:

$$d^{iv}(x) = -\frac{1}{2} r^2 + f(k) \text{ para } x \in w_k \dots\dots\dots(4.41)$$

$r^2 = (x-m_k)^T C_k^{-1} (x-m_k)$  es la distancia de Mahalanobis definida anteriormente y  $f(k) = \log p(w_k) - \frac{1}{2} \log|C_k|$ . Esta función esta determinada para dos casos, los cuales discutiremos a continuación.

**Caso 1.- Cuando las matrices de covarianza son iguales para diferentes clases (es decir  $C_i, C_j, C_k = C$ ). Cuyo significado físico es la separación de clases o grupos, los cuales son iguales en dimensiones y forma, pero los grupos son centrados sobre diferentes medios. Expandiendo la ecuación general para  $d_k(x)$  [2 txt (cap3)], obtenemos**

$$d_k(x) = -\frac{1}{2} \mathbf{x}^T \mathbf{C}^{-1} \mathbf{x} - \frac{1}{2} \mathbf{m}_k^T \mathbf{C}^{-1} \mathbf{m}_k + \mathbf{x}^T \mathbf{C}^{-1} \mathbf{m}_k + \log p(w_k) - \frac{1}{2} \log |C| \dots \dots \dots (4.42)$$

El primer y último términos en el lado derecho de la ecuación son los mismos para todas las clases (es decir para toda  $k$ ). Por consiguiente esta función discriminante puede ser escrita en forma mas compacta como sigue

$$d_k(x) = \mathbf{x}^T \mathbf{C}^{-1} \mathbf{m}_k + [\log p(w_k) - \frac{1}{2} \mathbf{m}_k^T \mathbf{C}^{-1} \mathbf{m}_k] \quad k = 1, 2, \dots, M \dots \dots \dots (4.43)$$

Obviamente ésta es una función discriminante lineal, si tomamos  $\mathbf{C}^{-1} \mathbf{m}_k$  como  $w_k$  y se tratan los 2 términos dentro de los corchetes como un término aumentado,  $w_{k,n+1}$ . Para un problema de 2 clases ( $M=2$ ):

$$d(x) = d_1(x) - d_2(x) = \mathbf{x}^T \mathbf{C}^{-1} (\mathbf{m}_1 - \mathbf{m}_2) + \log \frac{p(w_1)}{p(w_2)} - \left[ (\mathbf{m}_1 - \mathbf{m}_2)^T \mathbf{C}^{-1} (\mathbf{m}_1 + \mathbf{m}_2) \right] \dots (4.44)$$

**Caso 2.- Cuando la matriz de covarianza  $C_k$  es de forma diagonal  $\sigma_k^2 \mathbf{I}$ , donde  $\sigma_k^2 = |C_k|$ . El significado físico de esto es que el grupo tiene componentes iguales en los ejes principales, y la distribución es de forma esférica [2 txt (cap3)].**

Entonces la sustitución de  $\sigma_k^2 \mathbf{I}$  por  $C_k$  en la ecuación (4.53) se da:

$$d_k(x) = -\frac{1}{2} \frac{(\mathbf{x} - \mathbf{m}_k)^T (\mathbf{x} - \mathbf{m}_k)}{\sigma_k^2} + [\log p(w_k) - \frac{1}{2} \log |\sigma_k^2|] \dots \dots \dots (4.45)$$

debido a que  $\mathbf{C}^{-1} = \frac{1}{\sigma_k^2}$ . Cuando los eventos son estadísticamente independientes, y

cuando cada evento tiene la misma varianza,  $\sigma^2$ , entonces  $\sigma_k = \sigma_j = \sigma \quad \forall j, k$ , esto es,

$$C_k = C_j = \sigma^2 \mathbf{I} \dots \dots \dots (4.46)$$

y

$$d_k(x) = -\frac{1}{2} \frac{\mathbf{x}^T \mathbf{x} - 2 \mathbf{x}^T \mathbf{m}_k + \mathbf{m}_k^T \mathbf{m}_k}{\sigma^2} + \log p(w_k) - \frac{1}{2} \log \sigma^2 \dots \dots \dots (4.47)$$

De nuevo,  $\mathbf{x}^T \mathbf{x}$  y  $\frac{1}{2} \log \sigma^2$  son lo mismo para toda  $k$ . Podemos omitir dos términos en

$$d_k(x) = \frac{1}{\sigma^2} \mathbf{x}^T \mathbf{m}_k + [\log p(\mathbf{w}_k) - \frac{1}{2\sigma^2} \mathbf{m}_k^T \mathbf{m}_k] \dots \dots \dots (4.48)$$

la cuál también puede ser tratada como una función discriminante lineal.

Si además suponemos que  $C_k = C_j = \sigma^2 I$ , la suposición es hecha para permitir que  $p(w) = 1/k \forall k$ , donde  $k$  es una constante. El término “ $\log p(w_k)$ ” puede también ser suprimida desde la expresión  $d_k(x)$ . Entonces  $d_k(x)$  puede ser además simplificada como:

$$d_k(x) = \mathbf{x}^T \mathbf{m}_k - \frac{1}{2} |\mathbf{m}_k|^2 \dots \dots \dots (4.49)$$

la ecuación anterior es obviamente lineal.

Del análisis que se hizo anteriormente, para la función discriminante cuadrática, obtenida para la densidad normal multivariada para clases con  $M$ -patrones, se toma en cuenta que pueden ser simplificadas dentro a una forma que puede ser implementada por un mecanismo lineal, y de este modo hacer el problema mucho mas simple.

La ecuación (4.60) puede ser simplificada en otra forma, con la cual estamos familiarizados, puesto que se asume que  $C_k = C_j = \sigma^2 I$  y  $p(w_k) = p(w_j) = 1/k = una constante$ . Después suprimiendo los términos innecesarios, de modo que la ecuación (4.60), queda de la forma:

$$d_k(x) = \frac{1}{2} \left( \frac{(\mathbf{x}-\mathbf{m})^T (\mathbf{x}-\mathbf{m}_k)}{\sigma^2} \right) \dots \dots \dots (4.50)$$

o simplemente:

$$d_k(x) = (\mathbf{x}-\mathbf{m})^T (\mathbf{x}-\mathbf{m}_k) = -|\mathbf{x}-\mathbf{m}_k|^2 \dots \dots \dots (4.51)$$

con esta ecuación se obtiene también la distancia mínima para clasificar.

Para concluir, debemos mencionar que la función de densidad normal multivariada, es una función de probabilidad disponible para representar la distribución de variables aleatorias. Si  $k_n$ ,  $|W^{\frac{1}{2}}|$ , y  $f[(\mathbf{x}-\mathbf{m})^T W(\mathbf{x}-\mathbf{m})]$ , reemplaza a  $(2\pi)^{n/2}$ ,  $|C|^2$ , y

$$\exp\left[-\left(\frac{1}{2}(\mathbf{x}-\mathbf{m})^T C^{-1}(\mathbf{x}-\mathbf{m})\right)\right],$$

respectivamente, la función de densidad normal multivariada queda como:

$$p(x) = \frac{1}{(2\pi)^{n/2} |C|^{1/2}} \exp \left[ -\frac{1}{2} (\mathbf{x}-\mathbf{m})^T C^{-1} (\mathbf{x}-\mathbf{m}) \right] \dots\dots\dots(4.52)$$

puede ser generalizada como:

$$p(x) = k_n |W|^{1/2} f \left[ (\mathbf{x}-\mathbf{m})^T W (\mathbf{x}-\mathbf{m}) \right] \dots\dots\dots(4.53)$$

con  $k_n$  como la constante de normalización y  $W$  como la matriz de peso. Cuando los valores y funciones diferentes son dados para  $k_n, W, f$ , se pueden obtener diferentes tipos de función de densidad.

Un simple ejemplo ayudará como obtener la media, la covarianza y la función discriminante por el método de decisión estadística.

*Ejemplo:*

Dado un patrón de puntos  $(1,2)^T, (2,2)^T, (3,1)^T, (3,2)^T$ , y  $(2,3)^T$  que pertenecen a la clase  $w_1$ . Otro conjunto de puntos,  $(7,9)^T, (8,9)^T, (9,8)^T, (9,9)^T$ , y  $(8,10)^T$  que pertenecen a la clase  $w_2$ . Se requiere un límite para encontrar una decisión bayesiana y separarlos.

*Solución:*

$$m_1 = \frac{1}{N_1} \sum_{j=1}^{N_1} \mathbf{x}_{1j} = \frac{1}{5} \begin{bmatrix} 11 \\ 10 \end{bmatrix}$$

$$m_2 = \frac{1}{N_2} \sum_{j=1}^{N_2} \mathbf{x}_{2j} = \frac{1}{5} \begin{bmatrix} 41 \\ 45 \end{bmatrix}$$

Por definición:

$$C = E \left[ (\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})^{(T)} \right] = E \left[ \mathbf{x} \mathbf{x}^T \right] - \mathbf{m} \mathbf{m}^T$$

cuando es puesto en forma discreta

$$C_l = \frac{1}{N_1} \sum_{j=1}^{N_1} \mathbf{x}_{1j} \mathbf{x}_{1j}^T - \mathbf{m}_1 \mathbf{m}_1^T$$

por tanto

$$C_1 = \frac{1}{5} \left[ \binom{1}{2} (1 \ 2) + \binom{2}{2} (2 \ 2) + \binom{3}{1} (3 \ 1) + \binom{3}{2} (3 \ 2) + \binom{2}{3} (2 \ 3) \right] - \frac{1}{25} \binom{11}{10} (11 \ 10)$$

$$= \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix}$$

similarmente

$$C_1 = \frac{1}{N_2} \sum_{j=1}^{N_2} \mathbf{x}_{2j} \mathbf{x}_{2j}^T - m_2 m_2^T = \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix}$$

Tenemos entonces que:

$$C_1 = C_2 = C = \frac{1}{25} \begin{pmatrix} 14 & -5 \\ -5 & 10 \end{pmatrix}$$

El determinante y adjunto de  $C$  puede ser calculado como:

$$|C| = \frac{1}{25} \begin{vmatrix} 14 & -5 \\ -5 & 10 \end{vmatrix} = \frac{23}{5} \quad \text{adj } C = \begin{pmatrix} \frac{10}{25} & \frac{5}{25} \\ \frac{5}{25} & \frac{14}{25} \end{pmatrix}$$

la inversa de  $C$ ,  $C^{-1} m_1$  y  $m_1^T C^{-1} m_1$  son entonces, respectivamente como sigue:

$$C^{-1} = \frac{1}{|C|} \text{adj } C = \frac{1}{115} \begin{pmatrix} 10 & 5 \\ 5 & 14 \end{pmatrix}$$

$$C^{-1} m_1 = \frac{1}{115} \begin{pmatrix} 32 \\ 39 \end{pmatrix}$$

$$m_1^T C^{-1} m_1 = \frac{742}{5 \cdot 115}$$

La función discriminante para clase 1 es:

$$d_1(x) = x^T C^{-1} m_1 - \frac{1}{2} m_1^T C^{-1} m_1$$

$$= \frac{32}{115} x_1 + \frac{39}{115} x_2 - 0.65$$

similarmente obtenemos

$$C^{-1}m_1 = \frac{1}{115} \begin{pmatrix} 127 \\ 167 \end{pmatrix}$$

$$m_2^T C^{-1} m_2 = 22$$

El discriminante de la función para 2 clases es:

$$\begin{aligned} d_2(x) &= x^T C^{-1} m_2 - \frac{1}{2} m_2^T C^{-1} m_2 \\ &= \frac{127}{115} x_1 + \frac{167}{115} x_2 - 11 \end{aligned}$$

De todo lo anterior podemos determinar que la superficie de decisión estará dada por:

$$d(x) = d_1(x) - d_2(x) = 0$$

o

$$d(x) = -0.826x_1 - 1.11x_2 + 10.35 = 0.$$

#### 4.4 Clase de rechazo

En las aplicaciones prácticas del reconocimiento de patrones, frecuentemente sucede que usando una simple regla de decisión como las ecuaciones siguientes:

$$\text{Si } P(\mathcal{W}_1|x) > P(\mathcal{W}_2|x) \Rightarrow x \in \mathcal{W}_1$$

$$\text{Si } P(\mathcal{W}_1|x) < P(\mathcal{W}_2|x) \Rightarrow x \in \mathcal{W}_2 \dots\dots\dots(4.54)$$

$$\text{Si } P(\mathcal{W}_1|x) = P(\mathcal{W}_2|x) \Rightarrow \text{la decisión es arbitraria}$$

Entonces podemos establecer la ecuación anterior como:

$$\text{Si } P(\mathcal{W}_1|x) > P(\mathcal{W}_2|x) \Rightarrow x \in \mathcal{W}_1 \text{ sino } x \in \mathcal{W}_2$$

De acuerdo con la ecuación (4.18) La decisión dependerá en cómo se comparen las probabilidades de pertenencia, de acuerdo al umbral.

Es decir:

Predominio Equitativo.- La decisión del umbral de acuerdo a las distancias entre las medias, conforme a la distancia de Mahalanobis  $r^2 = (x - m_k)^T C_k^{-1} (x - m_k)$ .

Predominio en que  $w_1 > w_2$ . La decisión del umbral es desplazada a la clase más próxima con mayor predominio, por lo tanto el número de casos de las clases con menor predominio irá decreciendo dando así una clasificación un tanto errónea, de forma inconveniente.

En base a las anteriores decisiones, se pueden producir varias decisiones de frontera muy sensibles a cuestiones de ruido presentes en los datos de la imagen. En lo que se

refiere a los bordes de línea de la imagen, donde se encuentran muchos patrones situados cerca de estos, se puede hacer un cambio de asignación de clases por un ajuste no muy complicado.

Tomando en cuenta la unión de descripciones equivalentes para obtener una clasificación, un método usado en ciertas circunstancias es tomar en cuenta la existencia de ciertas clases especiales, denominadas “clases de rechazo” o “región de rechazo”.

Conociendo lo siguiente:

$w^*$  : como la clase de decisión;

$w_i$ : como las clases con probabilidad máximo a posteriori, es decir,  
 $P(w_i|x) = \max P(w_j|x)$  para todas las clases  $w_j \neq w_i$ .

La regla de Bayes puede entonces ser escrita simplemente como  $w^* = w_i$ .

Si se estipula ahora que las probabilidades a posteriori pueden ser mayor que cierto umbral de rechazo  $\lambda_r$ , implica que la clasificación en una cierta clase de rechazo  $w_r$ . La regla de Bayes es entonces reformulada como:

$$w^* = \begin{cases} w_i & \text{si } P(w_i|x) \geq \lambda_r \\ w_r & \text{si } P(w_i|x) < \lambda_r \end{cases} \dots\dots\dots(4.55)$$

# **5 Segmentación por campos aleatorios de Markov**

## **5.1 Introducción**

La teoría de los campos aleatorios de Markov (CAM) proporciona una manera conveniente para modelar la dependencia contextual de entidades, tales como el nivel de gris y otras características correlacionadas del espacio de una imagen. El uso práctico de los modelos de campos aleatorios de Markov se atribuye en gran parte a la equivalencia entre las distribuciones de probabilidad de los campos aleatorios de Markov y de la distribución de Gibbs (establecidas por Hammersley y Clifford 1971 y posteriormente desarrolladas por Besag 1974), permitiendo modelar matemáticamente problemas de visión y otras aplicaciones de análisis de imágenes, con un fundamento bayesiano [26 p.e].

Desde el punto de vista computacional, la característica local de estos campos conduce a algoritmos que se pueden implementar de manera local. Estos procesos son una extensión de las cadenas de Markov, en las cuales el índice de tiempo se substituye por un índice espacial.

Los modelos de visión estocásticos basados en los CAM se formulan dentro del enfoque bayesiano generalmente. Encuentran una solución óptima con el empleo de un estimador de probabilidad máxima a posteriori (MAP). La mayoría de los problemas de visión pueden ser considerados como de etiquetado usando restricciones, en base a conocimiento a priori y a diversas observaciones. En este caso, la solución óptima es definida por el etiquetado MAP y se calcula minimizando una función de energía a posteriori [27 p.e].

Los términos de probabilidad se definen empleando la formulación de Bayes. A partir de un modelo a priori y un modelo de probabilidad condicional, los términos de probabilidad deben reflejar información de observaciones (niveles de gris) y del modelo de segmentación (el modelo de Potts, en el caso de este trabajo). La teoría de los CAM proporciona herramientas para representar restricciones contextuales en la definición de la probabilidad a priori. Esta es la principal razón para el modelo de visión usando CAM [3 txt cap 6].

Un modelo de los campos aleatorios de Markov se puede resumir en 5 pasos:

1. Planteamiento del problema de visión como uno de etiquetado en el que la configuración de etiquetas representa una solución.
2. Planteamiento de un problema de etiquetado bayesiano en el que la solución óptima se define como la configuración de etiquetas MAP.
3. Empleo de la distribución de Gibbs para caracterizar la distribución a priori de las configuraciones del etiquetado.

4. Definición de la función de energía de la distribución de Gibbs.
5. Empleo de un algoritmo (por ejemplo, recocido simulado o simulated annealing) para minimizar la función de energía del modelo, obteniendo así la configuración óptima de etiquetas MAP.

### **5.2 Etiquetado en imágenes**

El etiquetado es una representación natural para el estudio de los CAM. Muchos problemas de visión pueden plantearse como problemas de etiquetado en los que la solución es un conjunto de etiquetas asignadas a píxeles o características de la imagen. Sea  $S$  un conjunto discreto con  $m$  estados.

$$\text{Sea } S = \{1, 2, \dots, m\} \dots \dots \dots (5.1)$$

Un estado (también llamado ‘site’) representa un punto o una región del espacio.

El conjunto de estados o ‘sites’ puede ser clasificado en términos de su homogeneidad o regularidad. Se puede denotar por:

$$S = \{(i, j) / 1 \leq i, j \leq n\} \dots \dots \dots (5.2)$$

al conjunto de localizaciones de los píxeles de una imagen 2D de tamaño  $n \times n$ .

Los estados que no representan regularidad espacial se consideran como irregulares. Para una imagen de  $n \times n$ , el píxel  $(i, j)$  puede ser reindexado convenientemente por un número único  $k$ , donde  $k$  toma valores en  $\{1, 2, \dots, m\}$  con  $m = n \times n$ . La interrelación entre estados o ‘sites’ se basa en sistemas de vecindad. Es por eso que una etiqueta es un evento que puede ocurrir a un estado.

Sea  $L$  el conjunto de etiquetas,  $L$  puede ser un conjunto discreto o continuo. En el caso continuo, el conjunto de etiquetas puede corresponder a  $\mathfrak{R}$  (conjunto de números reales) o a un intervalo compacto de  $\mathfrak{R}$ .

$$L_c = [x_l, x_h] \subset \mathfrak{R} \dots \dots \dots (5.3)$$

En el caso discreto, las etiquetas se toman dentro de un conjunto discreto de  $M$  valores,

$$L_d = [I_1, \dots, I_M] = \{1 \dots M\} \dots \dots \dots (5.4)$$

El conjunto de etiquetas puede ser caracterizado en términos de su continuidad. Otra característica esencial de un conjunto de etiquetas es poder establecer un orden entre ellas  $\mathfrak{R}$  pueden ordenarse por la relación “mas pequeño que”. Cuando un conjunto discreto como el de tonos de gris  $\{0, \dots, 255\}$  representa el valor de intensidades cuantizadas, entonces se trata de un conjunto ordenado, ya que para los valores de intensidad se cumple  $0 < 1 < 2 \dots < 255$ . Cuando un conjunto de etiquetas representa 256 símbolos (en relación a los tonos de gris) como tipos de texturas, entonces se considera como desordenado a menos que un orden por medio de etiquetas le sea asignado. En un conjunto de etiquetas ordenado, se puede definir una medida numérica (cuantitativa) de

## Capítulo 5. Segmentación por campos aleatorios de Markov

semejanza entre cualquier par de etiquetas. Para un conjunto de etiquetas sin relación de orden, la medida de semejanza es simbólica (cualitativa), normalmente tomando un valor *igual o no igual*.

El problema de etiquetado consiste en asignar una etiqueta del conjunto de etiquetas  $L$  a cada estado de  $S$ . Por ejemplo, para la detección de bordes en una imagen, el proceso consiste en asignar una etiqueta  $f_i$  dentro del conjunto  $L = \{\text{borde}, \text{no borde}\}$  para cada estado  $i \in L$ , donde  $S$  está formado por los píxeles de la imagen [26 p.e].

Llamando *etiquetado* a:  $f = \{f_1, \dots, f_m\}$

en donde a cada estado se le ha asignado una de las etiquetas de  $L$ .

Llamando *mapeo* a la función:  $f: S \rightarrow L$   
 en las que a cada estado de  $S$  se le asigna una única etiqueta  $f_i$  de  $L$ .

Cuando todos los estados tienen el mismo conjunto de etiquetas  $L$ , el conjunto de todos los etiquetados posibles es el producto cartesiano:

$$\tau = L \times L \times L \times \dots \times L = L^m \dots \dots \dots (5.5)$$

donde  $m$  es el tamaño de  $S$ .

En el problema de restauración de imágenes,  $L$  contiene los valores admisibles de cada píxel de  $S$  y  $\mathfrak{T}$  define todas las imágenes admisibles. Si  $L = \mathfrak{R}$ , entonces  $\tau = \mathfrak{R}^m$ . Si  $L$  es un conjunto discreto, entonces el tamaño de  $\mathfrak{T}$  es combinatorial. Para un problema con  $m$  estados y  $M$  etiquetas existen  $M^m$  etiquetados posibles en  $\mathfrak{T}$ .

En ciertas circunstancias, el conjunto de etiquetas puede no ser común a todos los estados. En un caso extremo, cada estado  $L$  puede tener su propio conjunto de etiquetas posibles  $L_i$  y esto conduce al siguiente espacio de etiquetados:

$$\tau = L_1 \times L_2 \times L_3 \times \dots \times L_m \dots \dots \dots (5.6)$$

En el proceso de una selección de etiquetas el concepto de contexto juega un papel importante en el análisis de imagen y en reconocimiento de patrones. En términos de probabilidad, las restricciones de contexto pueden expresarse localmente en términos 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  $P(f)$ .

Un aspecto importante de mencionar es que en los casos donde las etiquetas son independientes unas de otras, la probabilidad conjunta es el producto:

$$P(f) = \prod_{i \in S} P(f_i) \dots \dots \dots (5.7)$$

Lo cual implica también la independencia de la probabilidad condicional:

$$P(f_i / \{f_j\}) = P(f_i) \dots \dots \dots (5.8)$$

Por lo tanto, un etiquetado global  $f$  puede calcularse considerando cada etiqueta local  $f_i$ . En presencia de restricciones de contexto, las etiquetas son mutuamente dependientes. Cómo hacer entonces una inferencia global usando información local se convierte en una tarea no trivial, la teoría de los CAM proporciona una base matemática para solucionar este problema.

### 5.3 Sistema de vecinos y asociación de cliques

Los estados o ‘sites’ están relacionados unos con otros en base a un *sistema de vecindad*. El cual se define como:

$$N = \{N_i | \forall i \in S\} \dots \dots \dots (5.9)$$

en donde  $N_i$  es el conjunto de estados o ‘sites’ vecinos a de  $i$ . Con las siguientes propiedades de vecindad:

- 1).-  $i \notin (i,j)$  (un estado o ‘site’ no puede ser vecino de sí mismo); y
- 2).-  $i \notin N_j \leftrightarrow j \in N_i$  (la relación de vecindad es totalmente mutua).

Para una rejilla regular  $S$ , el conjunto de vecinos de  $i$  se define como el conjunto de estados próximos dentro de un radio  $r$ .

$$N_i = \{j \in S | \text{dist}(pixel_j, pixel_i)^2 \leq r, j \neq i\} \dots \dots \dots (5.10)$$

donde  $\text{dist}(pixel_j, pixel_i)$  es la distancia euclídea entre  $pixel_j$ , y  $pixel_i$  y  $r$  es un número entero. Cabe mencionar que los píxeles de la frontera tienen menos vecinos.

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 5.1(a), donde se 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 5.1(b). Los números  $n = 1, \dots, 5$  mostrados en la figura 5.1(c) indican los estados vecinos exteriores en el sistema de vecinos de orden  $Q$ -ésimo.

Cuando el orden de los elementos en  $S$  está especificado, el conjunto de vecinos puede determinarse de manera más explícita. Por ejemplo, sea  $S = \{1, \dots, m\}$  un conjunto ordenado de estados o ‘sites’, en los que sus elementos son los índices de los píxeles de una imagen 1D, un estado o ‘site’ interior  $i \in \{2, \dots, m-1\}$  tiene dos vecinos más cercanos  $N_i = \{i-1, i+1\}$ , y los estados de la frontera (los dos extremos) tiene uno,  $N_1 = \{2\}$  y  $N_m = \{m-1\}$ . El conjunto de estados en una rejilla regular y rectangular  $S = \{(i,j) | 1 \leq i, j \leq n\}$ , corresponde a los píxeles de una imagen en dos dimensiones  $n \times n$ , en donde un estado interno  $(i,j)$  tiene cuatro vecinos más cercanos  $N_{i,j} = \{(i-1,j), (i+1,j), (i,j-1), (i,j+1)\}$ , los estados en la frontera tienen tres y los de las esquinas tienen dos.

Generalmente, hablando de imágenes, estas son universalmente definidas en una superficie  $L$  rectangular y finita de  $M \times N$ .

$$L = \{(i,j): 1 \leq i \leq M, 1 \leq j \leq N\} \dots \dots \dots (5.11)$$

Con la definición de  $L$  descrita anteriormente en la ecuación 5.5, el llamado vecindario o vecindad de píxeles en una locación  $(i,j)$ , es denotado por  $N_{i,j}$ , la cuál es usualmente definido como:

- a).-  $(i,j) \notin N_{i,j}$ , y
- b).- si  $(m,n) \in N_{i,j}$ , entonces  $(i,j) \in \eta_{m,n}, \forall (m,n) \in L$ .

Estos incisos nos señalan la manera en como se asocia una determinada región de trabajo en un sistema vecinal o vecindario de píxeles.

La colección de todos los subconjuntos semejantes de  $L$  correspondientes a una definición específica de vecinos, es entonces llamada como un sistema de vecindad  $\eta$  para  $L$ , dicho de otra forma, como:

$$N = \{N_{i,j}: (i,j) \in L, N_{i,j} \subseteq L\} \dots \dots \dots (5.12)$$

En general, el conjunto de vecinos de  $N_i$  para un conjunto irregular de estados o 'sites'  $S$  tiene una dimensión y tamaño variables. El par  $\{S,N\}$  es un grafo donde  $S$  contiene los nodos y  $N$  determina las conexiones entre los nodos de acuerdo a la relación entre vecinos.

Un subconjunto  $C \subseteq L$  es llamado un clique si  $C$  consiste de un único píxel o un grupo de píxeles semejantes en un área vecinal.

Un 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  $\{i, j, k\}$  y así sucesivamente. El conjunto de cliques de un solo estado, de un par de estados, o de tres estados se denota como:

$$C_1 = \{i | i \in S\} \dots \dots \dots (5.13)$$

$$C_2 = \{\{i,j\} | j \in N_i, i \in S\} \dots \dots \dots (5.14)$$

$$C_3 = \{\{i,j,k\} | i,j,k, i \in S \text{ son vecinos unos de otros}\} \dots \dots \dots (5.15)$$

*El conjunto de cliques para  $\{S,N\}$  será entonces:  $C = C_1 \cup C_2 \cup C_3 \dots$*

Así mismo, es de especial interés para el procesamiento de imágenes el llamado sistema de vecindad homogénea  $N_c$ , con  $N_{i,j}$  siendo definido por el siguiente camino:

$$N_{i,j} = \{(m,n) \in L: 0 < (m - i)^2 + (n - j)^2 = d^2 \leq c\} \dots \dots \dots (5.16)$$

Hay que tomar en cuenta que puesto que  $i, j, m$  y  $n$  tienen que ser enteros,  $d^2$  puede también tomar solamente valores enteros así como  $c$ . Por lo tanto si  $c = 1$ , la vecindad de un píxel meramente consiste de sus cuatro vecinos más cercanos y es llamado sistema vecinal de primer orden. Para  $c = 2$  la vecindad contiene los ocho vecinos más cercanos del píxel central y es determinado como sistema vecinal de segundo-orden.

## Capítulo 5. Segmentación por campos aleatorios de Markov

Un punto importante por mencionar es que mientras  $c$  pueda ser un número natural, el valor de  $d^2$  puede formar una secuencia fija 1, 2, 3, 4, 5, 8, 9... . Si  $c$  es limitada a tomar solamente valores de esta secuencia, los números ordinarios de  $c$  en esta secuencia que es también definida como el orden del correspondiente sistema vecinal.

Los dos modelos en que  $c = 1$  y  $c = 2$ , son los modelos más ampliamente usados en sistemas vecinales en las distintas literaturas del procesamiento de imágenes [3 txt, 4 txt], asociados a los tipos de cliques; a continuación se muestran en la figura 5.1 ejemplos de estados vecinos de los tipos de cliques asociados, de acuerdo al orden vecinal.

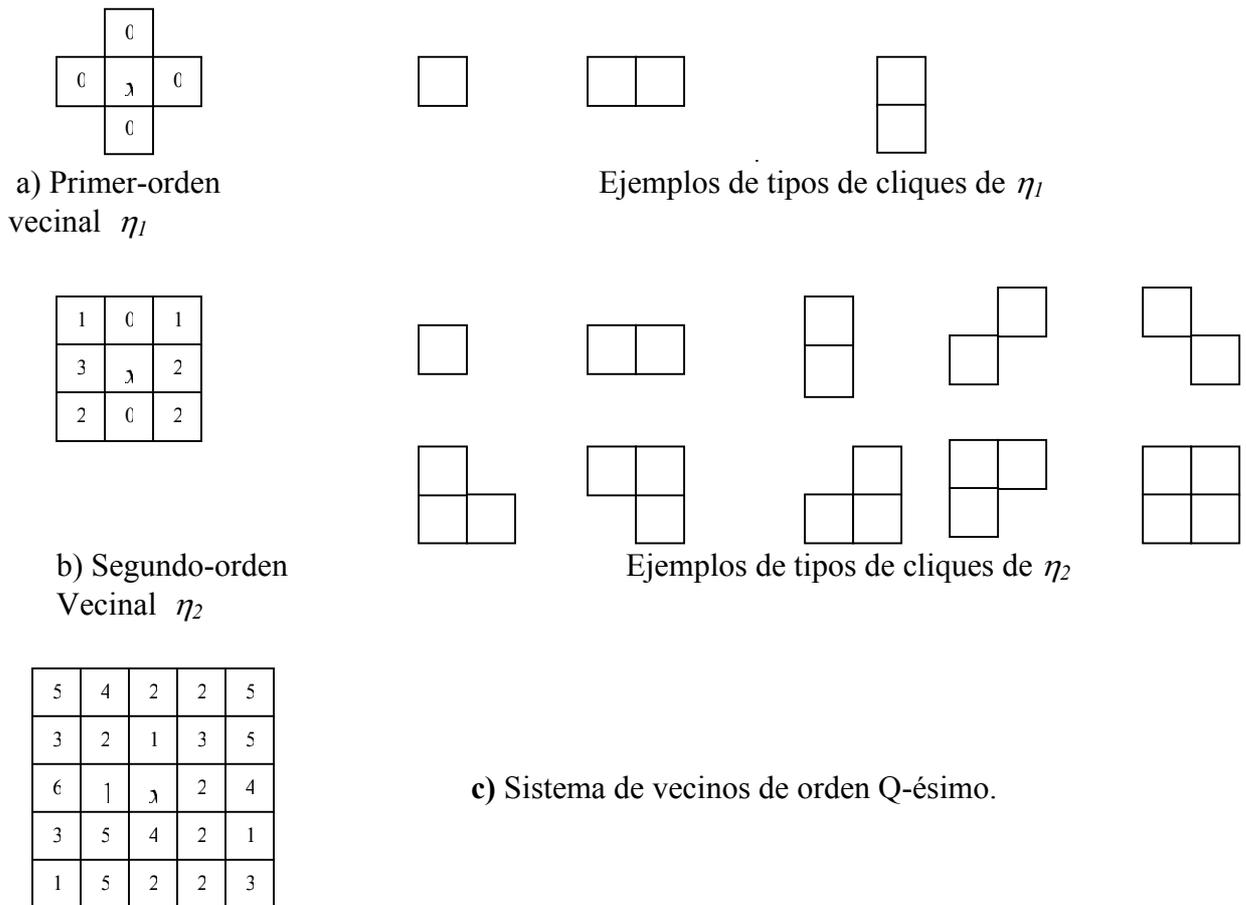


Figura 5.1.- Tipos de cliques asociados con primero, segundo y tercer orden, para un sistema vecinal [4 txt (cap 5)].

Es obvio que el número de tipos de clique crecerá rápidamente como el orden de un sistema vecinal se incrementa. En aplicaciones de tiempo real, sin embargo, solamente el primer y segundo orden del sistema vecinal tienen las principales aplicaciones.

Además, en el camino de determinar el tipo de región a la cual pertenece un píxel específico, estaremos interesados principalmente en la relación entre ese píxel con sus vecinos, pero no en la relación entre sus vecinos. Por lo tanto, los cliques los cuales pueden ser tomados en cuenta por el píxel central serán para un clique con un punto y para un clique con dos puntos, aquellos los cuales son compuestos del píxel siendo considerados y uno de sus ocho vecinos mas cercanos.

**5.4 Topologías (Asignación de etiquetas en los campos aleatorios de Markov)**

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  $L$ . Usaremos la notación  $F_i = f_i$  para denotar que el suceso  $F_i$  toma el valor  $f_i$  [26 p.e].

Un suceso conjunto  $\{F_1 = f_1, \dots, F_m = f_m\}$  abreviadamente  $f = \mathbf{f}$ , es una realización de  $F$ . Llamaremos configuración del suceso conjunto  $F$  a  $f = \{f_1, \dots, f_m\}$  [26 p.e].

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

Para un conjunto continuo de etiquetas  $L$ , denotamos a las funciones de densidad de probabilidad por  $p(F_i = f_i)$  y  $p(F = f)$  respectivamente.

**Una familia de variables (aleatorias) se dice que es un campo aleatorio de Markov sobre  $S$  con respecto a  $N_i$  si y sólo si:**

- a).-  $P(F = f) > 0 \quad \forall f \in \tau$
- b).-  $P(F_i = f_i / F_j = f_j, j \in d, j \neq i) = P(F_i = f_i / F_j = f_j, j \in N_i)$  Propiedad Markoviana.

La condición (a) es para exigir que  $F$  sea un campo aleatorio. La condición (b) se llama característica local. Se dice que la probabilidad de un suceso local a  $L$  condicionado a todos los sucesos restantes es equivalente a aquella condicionada a los sucesos vecinos de  $L$ . Puede demostrarse que la probabilidad conjunta  $P(F = f)$  de cualquier campo aleatorio se determina únicamente por estas probabilidades condicionales locales. 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  $P(f_i / N_i)$  se calcula sin importar la posición relativa del estado de  $i$  en  $S$ . La isotropía se considerará posteriormente cuando definamos las funciones “clique potenciales”.

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.

El concepto de campos aleatorios de Markov es una generalización de los procesos de Markov (PM) que han sido extensamente utilizados en análisis de secuencias. Los PM se definen normalmente en dominios de tiempo más que en dominios de espacio. Son una secuencia de variables aleatorias  $\dots, F_1, \dots, F_m, \dots$ , definidas sobre un conjunto de índices de tiempo  $\{\dots, l, \dots, m, \dots\}$  [4 txt (cap 5)]. Un PM unilaterial de orden n-ésimo satisface:

$$P(f_i / \dots, f_{i-2}, f_{i-1}) = P(f_i / f_{i-1}, f_{i-1}, f_{i-n}) \dots \dots \dots (5.17)$$

**Capítulo 5. Segmentación por campos aleatorios de Markov**

Un PM bilateral no causal depende no solamente del pasado sino también del futuro. Un PM bilateral de orden Q-ésimo satisface

$$P(f_i/\dots, f_{i-2}, f_{i-1}, f_{i+1}, f_{i+2}, \dots) = P(f_i/\dots, f_{i-n}, f_{i+1}, f_{i-1}, f_{i-n}) \dots \dots \dots (5.18)$$

Hay dos aproximaciones para especificar un campo aleatorio de Markov, en términos de probabilidades condicionales  $P(f_i/f_{Ni})$  y en términos de probabilidad conjunta  $P(f)$ . Discutiéndose así la aproximación mediante la probabilidad conjunta en vista de las desventajas de la aproximación mediante la probabilidad condicional: en primer lugar, no hay método obvio disponible para deducir la probabilidad conjunta de las probabilidades condicionales asociadas.

En segundo lugar, las probabilidades condicionales están sujetas a algunas condiciones de consistencia no obvias y altamente restrictivas. En tercer lugar, la especificación natural de equilibrio en un proceso estadístico es en términos de probabilidad conjunta en vez de en términos de distribución condicional de las variables. Afortunadamente, un resultado teórico sobre la equivalencia entre los campos aleatorios de Markov y la distribución de probabilidad de Gibbs proporciona medios de especificar la probabilidad conjunta de un campo aleatorio de Markov matemáticamente manejables.

Los CAM tienen el objetivo de fragmentar o dividir una imagen en áreas de textura similar. Con esta fragmentación, se proporciona una base para la detección de líneas, además de proporcionar la herramienta para la clasificación de la imagen y reconocimiento de patrones.

**5.4.1 Función de energía para la segmentación Markoviana**

1. Empleando la regla de Bayes:  $P(W_k|X) = \frac{P(X|W_k)P(W_k)}{P(X)} \dots \dots \dots (5.19)$

donde:

- $X$  es la imagen original.
- $W_k$  las clases con  $k = 0, 1, 2, \dots$
- $P(W_k|X)$  representa la probabilidad a posteriori.
- $P(W_k)$  es la probabilidad priori de la clase  $W_k$ .
- $P(X|W_k)$  es la probabilidad conjunta de  $X$  dado  $W_k$ .

$P(X)$  es la probabilidad total de  $X$ , debido a que es término común se puede eliminar, ya que en ningún momento depende de  $W_k$  el parámetro de interés.

Dado lo anterior se puede hacer una aproximación de las probabilidades condicionales del término  $P(X|W_k)P(W_k)$  mediante la función de energía que es explicará a continuación.

## Capítulo 5. Segmentación por campos aleatorios de Markov

La función de la energía hace referencia de cómo un píxel pertenece a una cierta clase determinada. De acuerdo a la figura 5.2, puede hacerse un proceso de clasificación en términos de  $P(X|W_k)$ . El nivel de gris de cada uno de los píxeles analizados se somete a la función de pertenencia de la figura 5.2; es decir a la clase a la que pertenece dicho píxel. Calculando así la función de la energía condicional de cada píxel respecto a cada determinada clase o etiqueta  $w_k$ . Cabe mencionar que el número de funciones depende del número de clases o etiquetas.

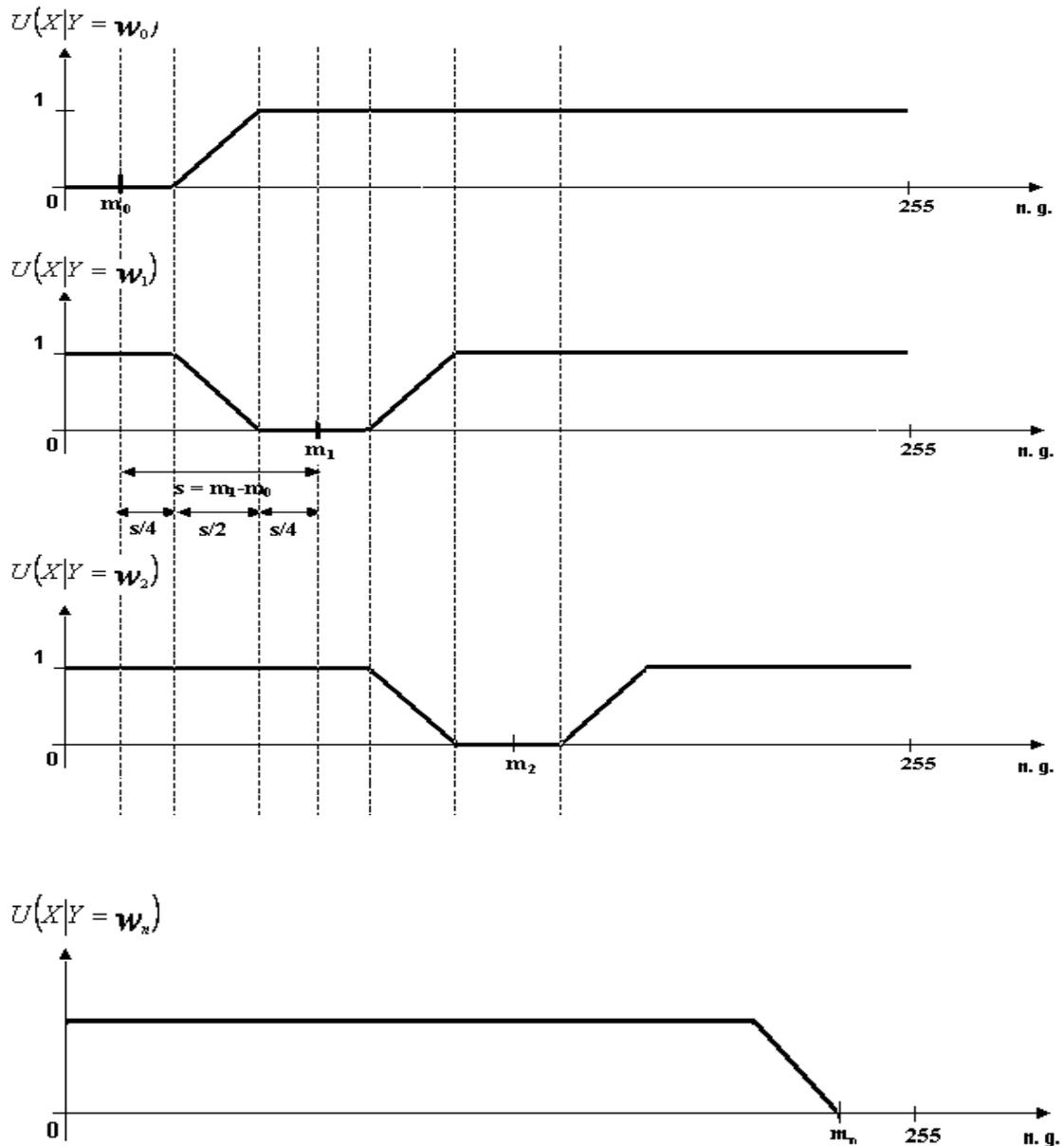


Figura 5.2.- Segmentación de las funciones de energía  $U(X|W_{k=n})$  [8 txt].

Si el nivel de gris del píxel analizado se encuentra en el intervalo  $0 \leq x < m_0$ , la energía condicional para la clase  $w_0$  (primera clase) será de 0.

## Capítulo 5. Segmentación por campos aleatorios de Markov

---

Si el nivel de gris del píxel analizado se encuentra en el intervalo  $x \geq m_n$ , la probabilidad condicional para la clase  $w_k$  (última clase) será de 0. Para las demás clases la energía condicional será de 1.

Si el nivel de gris del píxel analizado no se encuentra en ninguno de los 2 intervalos anteriores, se realiza la diferencia de las medias de las clases:  $S = m_{sig} - m_{act}$ , donde  $m_{sig}$  es la media de la clase siguiente y  $m_{act}$  es la media de la clase actual. El resultado de esta diferencia será dividido entre 4; quedando  $S/4$ .

Si el nivel de gris del píxel analizado se encuentra en el intervalo  $m_k \leq x \leq m_k + \frac{S}{4}$ , la probabilidad condicional para la clase  $w_{act}$  (clase actual) será de 0. Para las demás clases la probabilidad será de 1.

Si el nivel de gris del píxel analizado se encuentra dentro del intervalo  $m_k + S/4 \leq x \leq m_k + (3/4)S$ , y pertenece a la primera clase  $w_0$ , la energía condicional para la clase  $w_0$  se determina a través de la función:

$$i) E_C = \frac{2}{S}x - \frac{2}{S}m_k - \frac{1}{2} \text{ que será la pendiente positiva.}$$

y la diferencia de las medias  $S = m_{sig} - m_{act}$ .

Si el píxel pertenece a la última clase  $w_k$  se determina a través de la siguiente función:

$$ii) E_C = \frac{-2}{S}x + \frac{2}{S}m_k + \frac{3}{2} \text{ que será la pendiente negativa}$$

y la diferencia de las medias  $S = m_{act} - m_{ant}$ .

Para las demás clases en el caso de que el píxel analizado se encuentre en el intervalo de  $m_{ant} + S/4 \leq x \leq m_{ant} + (3/4)S$ , se utilizará la función (ii) con la correspondiente diferencia de las medias  $S = m_{act} - m_{ant}$ . En el caso de que el píxel analizado se encuentre en el intervalo de  $m_{act} + S/4 \leq x \leq m_{act} + (3/4)S$ , se utilizará la función (i), y la diferencia de medias  $S = m_{sig} - m_{act}$

donde:

- $E_C$  es la energía condicional de la clase  $w_k$
- $x$  es el píxel analizado.
- $m_k$  es la media de la clase  $k$  que se está analizando.

Tal y como se muestra en la figura 5.2,  $S$  se divide en 4 partes iguales, con segmentos que se superponen. De acuerdo a la figura en el segmento  $S/2$  la probabilidad para la clase 0 disminuye linealmente mientras que para la clase 1 la probabilidad se incrementa linealmente. Para la continuación de la gráfica será necesario seguir

comparando sucesivamente los valores de las medias de las clases restantes, es decir  $w_0, \dots, w_k$ .

En esta etapa la producción de la imagen, es conocida como una reconstrucción. Tal reconstrucción, usual y obviamente es un procedimiento matemático, el cual se define como un método de segmentación estocástica.

En 1953, Metrópolis propuso un algoritmo para simular un equilibrio térmico o termal de un sólido a cierta temperatura. El algoritmo de Metrópolis proporciona un simple, pero óptimo, método para la simulación de sistemas en donde se sujeta a pequeñas perturbaciones de forma aleatoria, resultando así un cambio en la energía  $\Delta E$  del sistema. Si  $\Delta E \leq 0$ , entonces ese desplazamiento es aceptado. Otro caso (en el caso de que  $\Delta E > 0$ ), el desplazamiento es aceptado con una probabilidad  $P(\Delta E)$ .

Lo cual se representa como:

$$P(\Delta E) = \exp\left(-\frac{\Delta E}{T}\right) \dots \dots \dots (5.20)$$

en donde  $T$  es la temperatura. Con el suficiente decrecimiento de  $T$ , el sistema puede alcanzar un equilibrio térmico.

Para el caso que compete a este trabajo, la segmentación de la imagen en cinco clases distintas, sostiene que el nivel de gris de un site analizado, depende únicamente de los valores de gris de sus sites vecinos, lo anterior en términos estadísticos. La energía local de la imagen en cada site se rige por el modelo de Potts. El algoritmo utilizado en la segmentación de la imagen se basa en un método de relajación óptimo, denominado *recocido simulado o simulated annealing*. El cual se describirá a continuación.

#### **5.4.2 Algoritmo de recocido simulado [8 txt]**

##### **1.- Inicialización**

- Determinar en forma aleatoria la solución  $Y^{(0)}$  (la clase de cada píxel debe determinarse mediante un generador de números aleatorios, con distribución uniforme, en el rango de numero de clases).
- Escoger una temperatura inicial  $T^{(0)}$  suficientemente elevada (en la mayoría de los casos se opta por usar,  $T^{(0)}=2.5$ ).

##### **2.-Hacer $N_{iter}=1$ , hasta $N_{iter}=n$**

###### **2.1.- Barrer secuencialmente la imagen, en cada etapa(n),**

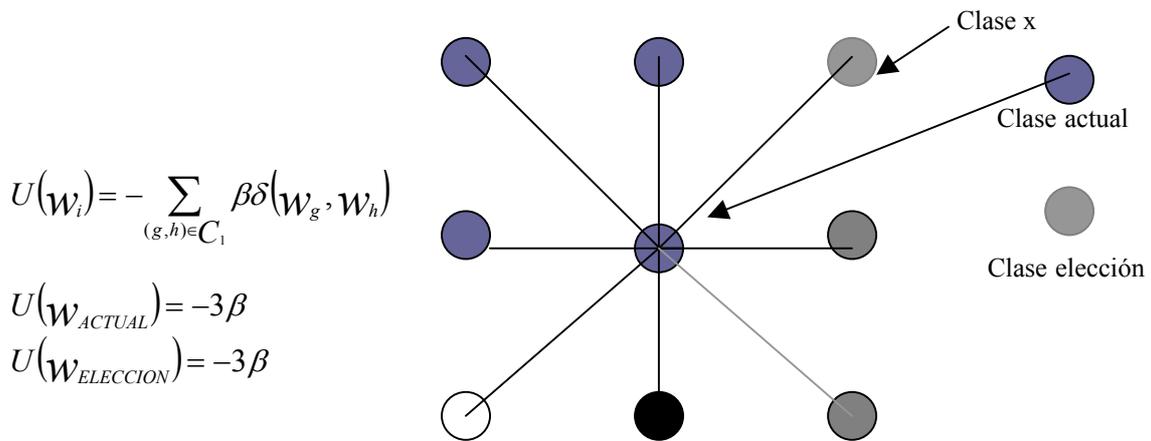
- Escoger un site:  $r = s^{(n)}$ , su valor es  $Y_r^{(n)}$  y su configuración de vecindad es  $V_r^{(n)}$ .
- Efectuar el sorteo de una variable aleatoria (“clase elección”) con distribución uniforme en el intervalo del número de clases:  $\xi \in \Omega$ .

**2.2.- Metrópolis y Potts**

**5.4.2.1 Criterio de Metrópolis y criterio de Potts**

En el criterio de Metrópolis se realizan barridos sobre la imagen y en cada site se tiene la posibilidad de efectuar un cambio en el estado del píxel.

En el modelo de Potts se tiene una configuración de vecindad, en este caso en 8-conexidad, lo que permite obtener la energía local del site (energía a priori). A continuación se muestra en la figura 5.3 un esquema del análisis en 8-conexidad y la forma en que se calcula la energía a priori:



**Figura 5.3. Esquema de 8-conexidad**

*Análisis en 8-conexidad:*

- Si "actual" = "x": prob\_actual = prob\_actual - β[actual]
- Si "elección" = "x": prob\_elección = prob\_elección - β[actual]

**2.3.-** Calcular la variación de la energía local asociada al site:

- $prob\_actual = prob\_actual + potencial[actual][i][j]$   
 $U(w_{actual}|X) \approx U(X|w_{actual})$
- $prob\_elección = prob\_elección + potencial[elección][i][j]$   
 $U(w_{elección}|X) \approx U(X|w_{elección})$

donde potencial  $[[i][j]]$  representan  $U(w_{actual}|X)$  y  $U(X|w_{elección})$  las cuales se calculan mediante las funciones de energía condicional; e  $[i][j]$  son las coordenadas del píxel analizado.

Esto es una vez que se tiene el site  $r = s^{(n)}$  y su valor  $Y_r^{(n)}$ , el valor del nivel de gris del site en X se proyecta en la función con número igual a  $Y_r^{(n)}$ , con lo que se obtiene el valor de energía  $U(X|w_{elección})$ ; posteriormente el valor del nivel de gris del site  $r = s^{(n)}$  en X se proyecta en la función con número igual al valor de la variable aleatoria calculada en el punto 2.1.

$\xi \in \Omega$ , con lo que se obtiene la energía  $U(X|W_{elección})$  para el actual site.

$U(W_{actual})$  y  $U(W_{elección})$  son calculados con el modelo de Potts como se muestra en la figura 5.1, donde “clase actual” es el valor  $Y_r^{(n)}$  del site  $r=s^{(n)}$  y “clase elección” =  $\xi$  del punto 2.1

Si  $U(W_{elección}|X) \leq U(W_{actual}|X)$ :  $Y[i][j] = \xi$ ,  $i$  y  $j$  son las coordenadas del site analizado, es decir Si  $prob\_elección \leq prob\_actual$  entonces  $Y[i][j] = "elección"$ .

En caso contrario:

$$p = \exp\left[-\frac{U(W_{elección}|X) - U(W_{actual}|X)}{T^{(n)}}\right] \text{ es decir}$$

$$p = \exp\left[-\frac{(prob\_elección - prob\_actual)}{T^{(n)}}\right]$$

- Se realiza un sorteo de una variable aleatoria  $\xi_2$ , uniformemente distribuida en el rango  $[0,1]$ .
- Si  $\xi_2 < p$ , entonces  $Y[i][j] = \xi$ ,

3.- Regresar al punto 2.1 hasta terminar un barrido, el cual consta de analizar todos los sites de la imagen X.

4.- Decrementar la temperatura:  $T^{(n)} = T^{(n-1)} * 0.95$

5.- Regresar al punto 2 hasta terminar el número fijado de iteraciones.

En la figura 5.4 se muestra el esquema del algoritmo de simulación, para  $n$  iteraciones.

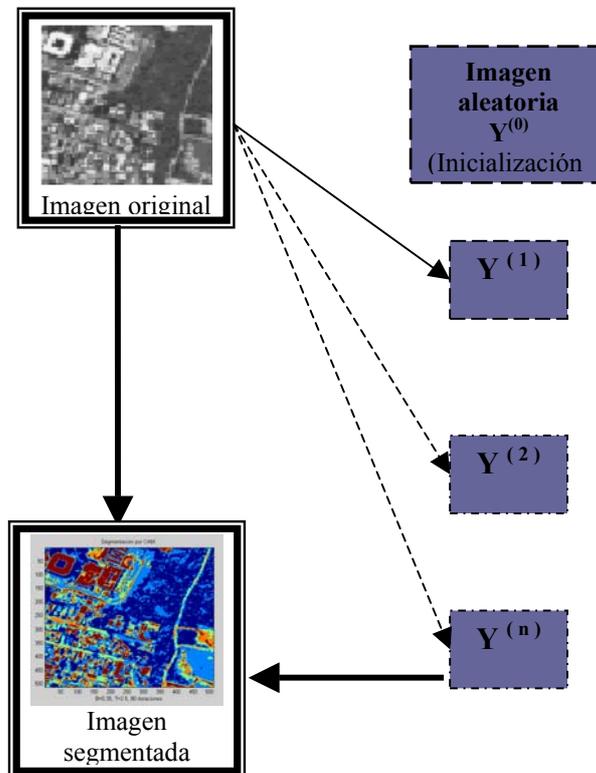


Figura 5.4. Esquema del algoritmo recocido simulado.

## Capítulo 5. Segmentación por campos aleatorios de Markov

El propósito de la segmentación es asignar una etiqueta a cada píxel, a partir de un conjunto finito. Esta segmentación consiste en minimizar la función de energía global asociada al criterio de Bayes:

$$U(Y|X) = U(X|Y) - U(Y) \dots \dots \dots (5.21)$$

Donde  $U(Y)$  es la función de energía de un modelo a priori seleccionado, y  $U(X|Y)$  representa un vínculo entre los niveles de gris y las clases. La distribución de probabilidad consiste en una aproximación condicional en términos de  $U(X|Y)$  por medio de la distribución gaussiana. La energía en términos asociados a  $U(X|Y = \mathcal{W}_k)$  se muestra en la figura 5.2. El eje de las abscisas denota los valores de los niveles de gris (0-255) y el eje de las ordenadas los valores de la energía. El valor  $S$  es la diferencia entre los valores de las medias de las ventanas de entrenamiento, para este ejemplo serían las clases 1 y 0:  $m_1 - m_0$ .

### 5.4.3 El campo aleatorio de Markov y la distribución de Gibbs

La limitación de continuidad espacial puede ser incorporada en el modelo de región asumiendo que el proceso de región  $X$  es un campo aleatorio de Markov con respecto a su sistema vecinal. Por definición, el CAM (campo aleatorio de Markov)  $X$  tiene la propiedad de que la distribución condicional de  $X$  en un píxel particular dados los valores de los otros píxeles en toda la imagen es igual a la distribución condicional dando solamente los valores de sus vecinos, es decir, con  $x_{i,j}, x_{mn} \in Q$  [4 txt (cap 5)],

$$P(X_{i,j} = x_{i,j} / X_{mn} = x_{mn} (m,n) \neq (i,j)) = P(X_{i,j} = x_{i,j} / X_{mn} = x_{mn}, (m,n) \in \eta_{i,j}) \dots \dots \dots (5.22)$$

Esta propiedad es muy importante y, cuando se trata de un píxel particular, nosotros podemos dejar afuera todos aquellos píxeles que se encuentran fuera de lugar de la vecindad.

Otra propiedad de los CAM de gran importancia es que, bajo una condición, el CAM  $X$  puede tener una distribución de Gibbs relativa para  $\{L, \eta\}$ . Una distribución de Gibbs es una probabilidad medida del atributo de un píxel particular. Cuando se tiene un conjunto de variables aleatorias  $F$  se dice que se cuenta con un campo aleatorio de Gibbs en  $S$  con respecto a  $N$ , sí y solo sí, sus configuraciones obedecen a distribuciones de Gibas [4 txt (cap 5)].

Derivada desde la propiedad de sus vecinos con la siguiente forma:

$$P(x_{i,j} = x) = \frac{1}{Z} \exp \left\{ -\frac{1}{T} \sum_{c \in C} V_c(x) \right\} \dots \dots \dots (5.23)$$

Donde  $x \in Q$ ,  $C$  es el conjunto de cliques relacionado con el píxel  $(i,j)$  como se discutió anteriormente:

- $V(x)$  es la llamada función potencial asociada con el clique  $C$ ,
- $T$  es la temperatura,
- $Z$  es la constante de normalización, definida por:

$$Z = \sum_x \exp \left\{ -\frac{1}{T} \sum_{c \in C} V_c(x) \right\} \dots\dots\dots(5.24)$$

Este valor discreto del CAM, es de vez en cuando llamado campo logístico multinivel (CML), requerido por muchos autores por ser apropiado para modelar la formación de una imagen en procesamiento.

La principal característica del CML, es que un parámetro positivo es asociado con cada tipo de clique, definido por:

$$V_C(x) = -\alpha_k \quad \text{si } x = q_k \dots\dots\dots(5.25)$$

Con una de sus componentes siendo según el píxel considerado, el potencial para un clique con dos puntos es definido de la siguiente manera:

$$V_C(x) = \begin{cases} -\beta & \text{si } \mathbf{X}_{mn} \text{ y } (m,n) \in C \\ \beta & \text{otro caso} \end{cases} \dots\dots\dots(5.26)$$

donde ambas  $\alpha_k$  y  $\beta$  son constantes positivas. De acuerdo con la definición de la distribución de Gibbs,  $x$  es el índice del tipo de región para ser determinada por el píxel siendo considerado y puede tomar un valor de  $Q$ . Para tomar el orden de la decisión se necesita evaluar todas las probabilidades para cada tipo de región que al píxel pertenece, y entonces toma el valor máximo.

En vista de la discusión anterior, es suficiente entonces para analizar los atributos del píxel central de él mismo y sus 8 vecinos. Por este motivo, la siguiente formula con  $k = 1, \dots, K$  puede ser mas preferible y clara.

$$V_C(x) = \begin{cases} -\beta & \text{si } \mathbf{X}_{mn} = \mathbf{Q}_k \text{ y } (m,n) \in C \\ \beta & \text{otro caso} \end{cases} \dots\dots\dots(5.27)$$

$$P(x_{i,j} = x) = \frac{1}{Z} \exp \left\{ -\frac{1}{T} \sum_{c \in C} V_c(x) \Big|_{x = q_k} \right\} \dots\dots\dots(5.28)$$

con

$$Z = \sum_{k=1}^K \exp \left\{ -\frac{1}{T} \sum_{c \in C} V_c(x) \Big|_{x = q_k} \right\} \dots\dots\dots(5.29)$$

**5.4.4 Campos aleatorios de Gibbs (GRF, Gibbs random fields)**

La ecuación (5.27) representa que  $Z$  es una constante de normalización llamada función de partición,  $T$  es una constante de temperatura generalmente con un valor igual a 1 y  $U(f)$  es la función de energía. La energía

$$U(f) = \sum_{c \in C} V_c(f) \dots\dots\dots(5.30)$$

es la suma de potenciales de cliques  $V_c(f)$  sobre todos los cliques posibles de  $C$ . El valor de  $V_c(f)$  depende de la configuración local en el clique  $c$ , la distribución gaussiana ocupa un papel fundamental en la distribución de Gibbs. Como se mencionó anteriormente una GRF es homogéneo si  $V_c(f)$  es independiente de la posición relativa del clique en  $S$ . Así mismo se dice que es isotrópico si  $V_c(f)$  es independiente de la orientación de  $C$ .

Para calcular la distribución de Gibbs es necesario evaluar la función de partición  $Z$ , siendo la suma de todas las configuraciones posibles en  $F$ . La función  $P(f)$  mide la probabilidad de ocurrencia de una configuración o patrón particular  $f$ .

Existe una característica que debe mencionarse, cuando se tienen configuraciones más probables, se tienen menores energías; en este caso la encargada de controlar la forma de distribución es la temperatura  $T$ . De manera que cuando la temperatura es alta, todas las configuraciones tienden a estar igualmente distribuidas, contrario así la temperatura es menor, es decir cuando este tiende a cero, la distribución se concentra alrededor de la energía mínima global. De modo que dadas  $T$  y  $U(f)$ , es posible generar una clase de patrones, muestreando el espacio de configuración  $F$  de acuerdo con  $P(f)$ . En el caso del etiquetado discreto, el potencial del clique  $V_c(f)$  puede especificarse por un número de parámetros.

Un ejemplo de esto sería si se tuviese a  $f_c = (f_i, f_{i'}, f_{i''})$  la configuración local de un clique triple  $c = \{i, i', i''\}$ ,  $f_c$  tomará un número finito de estados y por consiguiente  $V_c(f)$  toma un número finito de valores. En ocasiones resulta conveniente expresar la energía de una distribución de Gibbs como la suma de varios términos, cada uno atribuido a los cliques de cierto tamaño, esto es:

$$U(f) = \sum_{\{i\} \in C_1} V_1(f_i) + \sum_{\{i, i'\} \in C_2} V_2(f_i, f_{i'}) + \sum_{\{i, i', i''\} \in C_3} V_3(f_i, f_{i'}, f_{i''}) \dots \dots \dots (5.31)$$

Lo anterior implica una distribución de Gibbs homogénea debido a que  $V_1, V_2$  y  $V_3$  son independientes de la localización de  $i, i', i''$ . Un caso especial es cuando se consideran cliques de tamaño menores o iguales a dos. En donde la energía se escribe como:

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'}) \dots \dots \dots (5.32)$$

Considerando que en la segunda sumatoria de la ecuación anterior, en que las parejas forman a dos cliques distintos es decir  $\{i, i'\} \neq \{i', i\}$ , debido a que los estados del están ordenados. Por consiguiente esta probabilidad condicional se escribirá como:

$$P(f_i | f_{N_i}) = \frac{e^{-[V_1(f_i) + \sum_{j \in N_i} V_2(f_i, f_j)]}}{\sum_{f_i \in L} e^{-[V_1(f_i) + \sum_{j \in N_i} V_2(f_i, f_j)]}} \dots \dots \dots (5.33)$$

**5.4.5 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 (distribución de Gibbs) [4 txt (cap 5)].

La equivalencia de estas dos propiedades la establece el teorema de Hammersley-Clifford.

**Teorema 5.1.-**  $F$  es un campo aleatorio de Markov en  $S$  con respecto a  $N$  si y solo si  $F$  es un campo aleatorio de Gibbs en  $S$  con respecto a  $N$  [4 txt (cap 5)].

El teorema establece una probabilidad conjunta.

Sea  $P(f)$  una distribución de Gibbs en  $S$  con respecto al sistema de vecinos  $N$ . Si se considera la probabilidad condicional

$$P\left(f_i \mid f_{N_i}\right) = \frac{P\left(f_i, f_{N_i}\right)}{P\left(f_{N_i}\right)} = \frac{P(f)}{\sum_{f_i \in L} P(f')} \dots\dots\dots(5.34)$$

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 estado  $i$ .

Sustituyendo  $P(f) = \frac{1}{z} e^{-\left[\sum_{c \in C} V_c(f)\right]}$ , de modo que nos proporcione una formula para calcular la probabilidad condicional de las funciones potenciales, por lo tanto tenemos que:

$$P\left(f_i \mid f_{N_i}\right) = \frac{e^{-\sum_{c \in C} V_c(f)}}{\sum_{f_i \in L} \sum_{c \in C} V_c(f')} \dots\dots\dots(5.35)$$

Si se divide  $L$  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  $i$ . Por lo cual la probabilidad conjunta queda como:

$$P\left(f_i \mid f_{N_i}\right) = \frac{\left[ e^{-\sum_{c \in A} V_c(f)} \right] \left[ e^{-\sum_{c \in B} V_c(f)} \right]}{\sum_{f_i \in L} \left[ e^{-\sum_{c \in A} V_c(f')} \right] \left[ e^{-\sum_{c \in B} V_c(f')} \right]} \dots\dots\dots(5.36)$$

Como  $V_c(f) = V_c(f')$  para cualquier clique  $c$  que no contenga a  $i$  entonces  $e^{-\sum_{c \in B} V_c(f)}$  se simplifica tanto en el numerador como en el denominador. Por lo tanto, esta

**Capítulo 5. Segmentación por campos aleatorios de Markov**

probabilidad depende solamente de los potenciales de las cliques que contienen a  $i$ , lo cual implica la ecuación tanto en el numerador como en el denominador.

De acuerdo a la ecuación (5.35) la probabilidad depende solamente de la etiqueta en los vecinos de  $i$ . Con lo anterior se comprueba que un campo aleatorio de Gibbs en un campo aleatorio de Markov.

$$P\left(f_i \mid f_{N_i}\right) = \frac{\left| e^{-\sum_{c \in A} V_c(f)} \right|}{\sum_{f_i \in L} \left| e^{-\sum_{c \in A} V_c(f')} \right|} \dots\dots\dots(5.37)$$

Pero se debe tomar en cuenta de que la elección de funciones clique para un campo aleatorio de Markov no es única, porque pueden existir opciones equivalentes con una misma distribución de Gibbs. No obstante existe un potencial llamado canónico, para cada campo aleatorio de Markov.

**Definición 5.1.-** Sea  $L$  un conjunto de etiquetas numerables. Una función clique potencial  $V_c(f)$  esta normalizada si  $V_c(f) = 0 \ \forall i \in c$ , con  $f_i \in L$ .

Entonces se establece una relación matemática entre la distribución  $P(f)$  de un Campo Aleatorio de Markov y la representación canónica única de las funciones clique potenciales  $V_c(f)$  en la distribución de Gibbs.

**Teorema 5.2.-** Sea  $F$  un campo aleatorio en un conjunto finito  $S$  con características locales  $P\left(f_i \mid f_{N_i}\right)$ . Entonces  $F$  es un campo aleatorio de Gibbs con una función canónica definida por:

$$V_c(f) = \begin{cases} 0 & c = 0 \\ \sum_{b \subset c} (-1)^{|c-b|} \ln P\left(f^b\right) & c \neq 0 \end{cases} \dots\dots\dots(5.38)$$

donde  $|c-b|$  es el número de elemento del conjunto  $c-b$  y  $f^b$  esta dada por:

$$f_i^b = \begin{cases} f_i^b & i \in b \\ b & \text{otro caso} \end{cases}$$

la configuración  $f$  para los estados de  $b$  y valor de 0 para los estados fuera de  $b$ .

Para un conjunto no vacío  $c$ , la función clique potencial se puede obtener como:

$$V_c(f) = \sum_{b \subset c} (-1)^{|c-b|} \ln P\left(f_i^b \mid f_{N_i}^b\right) \dots\dots\dots(5.39)$$

donde  $i$  es cualquier elemento de  $b$ .

Usando el resultado anterior se puede calcular la función clique potencial  $V_c(f)$  si se conoce  $P(f)$ . Aunque hay que mencionar, que en los CAM que utilizan la distribución de Gibbs, la probabilidad  $P(f)$  se define después de  $V_c(f)$  con lo cual se dificulta calcular el potencial canónico de  $V_c(f)$  a partir de  $P(f)$ . Por lo consiguiente se calculará de una manera indirecta, es decir usando una forma no canónica para calcular  $P(f)$  y después usando el teorema 5.2 se puede obtener la representación canónica.

Definiendo 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, obteniendo así el conocimiento a priori sobre las interacciones de la etiquetas.

Las formas de las funciones potenciales determinarán la forma de la distribución de Gibbs. Para calcular la probabilidad conjunta de un CAM, que tiene una distribución de Gibbs, es necesario dar un valor a  $Z$ ,  $T$  y  $U(f)$ . Lo cual es posible calcular por medio del estimador MAP (máximo a posteriori) para CAM en donde  $V_c(f)$  no contiene parámetros desconocidos.

Para el propósito de la segmentación de imágenes, la característica de la imagen puede ser modelada de forma de  $K$  diferentes tipos de región, cada una siguiendo una distribución normal caracterizada por un valor diferente  $\mu_k$  y una varianza  $\sigma_k^2$ .

### 5.5 Aproximación máximo a posteriori

La teoría de los CAM proporciona una estructura conveniente para modelar entidades dependientes, en este caso los píxeles de una imagen dada, y otras características de espacio correlacionadas con estos. La teoría de los CAM muchas veces es usada en conjunto con la decisión estadística y teorías de estimación. La probabilidad de máximo a posteriori (MAP) es uno de los criterios estadísticos más usados para estos casos [4 txt (cap 5)].

Estableciendo la estimación probabilística, la ley condicional de observaciones puede ser condicionada. Esto es denotado por  $P(X|Y)$ , donde  $X$  es la imagen original y  $Y$  es el resultado de la segmentación,  $P(X|Y)$  describe el resultado.

Por Bayes:

$$P(Y|X) \approx P(X|Y)P(Y) \dots\dots\dots(5.40)$$

El estimador MAP consiste en encontrar la configuración  $y$  que maximiza la distribución a priori  $P(Y|X)$ . Un proceso equivalente es obtenido minimizando una función de energía:

$$y = \arg \max P(Y = y|X = x) = \arg \min U(y|x) \dots\dots\dots(5.41)$$

## Capítulo 5. Segmentación por campos aleatorios de Markov

Se asume un etiquetado con la relación máximo a posteriori (MAP) y campos aleatorios de Markov. La distribución a posteriori para un CAM se define por  $P(f|d)$ . De acuerdo con la regla de Bayes la probabilidad a posterior puede calcularse utilizando:

$$P(f|d) = \frac{p(d|f)P(f)}{p(d)} \dots\dots\dots(5.42)$$

donde  $P(f)$  es la probabilidad a priori de las etiquetas  $f$ ,  $p(d|f)$  es la función de densidad de probabilidad de las observaciones  $d$ , también conocida como función de probabilidad de  $f$  para una  $d$  fija, y  $p(d)$  es la densidad de  $d$  que es una constante cuando  $d$  es conocida.

Con una formulación de una restauración simple podemos ilustrar el etiquetado MAP. Asumiendo que la superficie es plana entonces la distribución conjunta a priori se define como:

$$P(f) = \frac{1}{z} e^{-U(f)} \dots\dots\dots(5.43)$$

donde  $U(f)$  es la energía a priori dada por

$$U(f) = \sum_i [f_i - f_{i-1}]^2 \dots\dots\dots(5.44)$$

Asumiendo que la observación es el valor verdadero más un ruido independiente gaussiano,  $d_i = f_i + e_i$ , donde  $e_i \sim N(\mu, \sigma^2)$ , entonces la densidad de probabilidad es

$$P(d|f) = \frac{1}{\prod_i \sqrt{2\pi \sigma_i^2}} e^{-U(d|f)} \dots\dots\dots(5.45)$$

donde

$$U(d|f) = \sum_{i=1}^m \frac{(f_i - d_i)^2}{2\sigma_i^2} \dots\dots\dots(5.46)$$

es la energía de probabilidad. La probabilidad a posteriori es:

$$P(f|d) \propto e^{-U(f|d)} \dots\dots\dots(5.47)$$

donde

$$U(f|d) = U(d|f) + U(f) = \sum_{i=1}^m \frac{(f_i - d_i)^2}{2\sigma_{i1}^2} + \sum_{i=1}^m (f_i - f_{i-1})^2 \dots\dots\dots(5.48)$$

es la energía a posteriori. En consecuencia el estimado de máximo a posteriori puede encontrarse minimizando esta energía a posteriori:

$$f^* = \arg \min_f U(f|d) \dots\dots\dots(5.49)$$

Donde los parámetros  $\sigma_i$  son conocidos y por consiguiente  $U(f|d)$  esta completamente definida, lo cual implica la solución del etiquetado de máximo a posteriori.

### 5.6 Modelos de Markov

La definición de algunos modelos útiles de campos aleatorios de Markov, es importante, porque contempla la forma de modelar algunas características importantes de una imagen como son diferentes tipos de región, de acuerdo a su textura de cada una, derivadas de las restricciones contextuales.

De manera tal que lo que implica en esto son las dos probabilidades más importantes en los CAM, las probabilidades condicional y conjunta, así como sus funciones de energía. Cabe mencionar que dependiendo de su función e energía, debemos agruparlos en Auto-modelos y Modelos logísticos multinivel (MLL) [26 p.e], los cuales se describirán a continuación.

#### 5.6.1.- Auto-modelos

Cuando existen dos etiquetas por consecuente existen restricciones contextuales, aunque estas se catalogan de menor orden, para expresar la información contextual. Estas restricciones son muy utilizadas debido a su forma simple y su no tan complejo costo computacional. Podemos decir que tales restricciones se encuentran incluidas dentro de la energía de Gibbs como funciones clique potenciales entre pares de estados o sites [26 p.e]. Con potenciales de cliques con 2 estados la función clique de energía potencial se representa como:

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{j \in N_i} V_2(f_i, f_j) \dots\dots\dots(5.50)$$

Esta ecuación es un caso especial de la ecuación (5.30), conocida como energía de *segundo orden*. Se puede especificar un CAM o un CAG mediante la selección apropiada de  $V_1$  y  $V_2$ .

Cuando  $V_1(f_i) = f_i G_i(f_i)$  y  $V_2(f_i, f_j) = \beta_{i,j}$ , donde  $G_i(f_i)$  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:

$$U(f) = \sum_{\{i\} \in C_1} f_i G_i(f_i) + \sum_{\{i,j\} \in C_2} \beta_{i,j} f_i f_j \dots\dots\dots(5.51)$$

**Definición 5.2.-**El modelo anterior recibe el nombre de auto-modelo. Los auto-modelos se pueden clasificar más a fondo según las suposiciones hechas sobre las  $f_i$  [26 p.e].

**Definición 5.3.-**Un auto-modelo se dice que es un modelo auto-logístico, si las  $f_i$  toman valores en el conjunto de etiquetas discreto  $L = \{0,1\}$  o  $L = \{+1,-1\}$  [26 p.e]. La correspondiente función de energía es de la forma:

$$U(f) = \sum_{\{i\} \in C_1} \alpha_i f_i + \sum_{\{i,j\} \in C_2} \beta_{i,j} f_i f_j \dots\dots\dots(5.52)$$

donde  $\beta_{i,j}$  es la interacción entre los coeficientes.

Cuando el sistema de vecinos  $N$  es una rejilla (2, 4, 8 vecinos en cada rejilla) entonces el modelo auto-logístico se reduce al modelo de Ising. La probabilidad condicional para el modelo auto-logístico con  $L = \{0,1\}$  es:

$$P(f_i | f_{N_i}) = \frac{e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}}{\sum_{f_i \in \{0,1\}} e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}} = \frac{e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}}{1 + e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}} \dots\dots\dots(5.53)$$

Si contamos con una distribución homogénea entonces:  $\alpha_i = \alpha$  y  $\beta_{i,j} = \beta$  sin importar  $i$  ni  $j$ .

**Definición 5.4.-** Un auto-modelo es un auto-modelo binomial si  $f_i$  toma valores en  $\{0,1,\dots,M-1\}$  y cada  $f_i$  tiene una distribución condicional binomial con  $M$  pruebas y probabilidad de éxito  $q$  [26 p.e].

$$P(f_i | N_i) = \binom{M-1}{f_i} q^{f_i} (1-q)^{M-1-f_i} \dots\dots\dots(5.54)$$

en donde

$$q = \frac{e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}}{1 + e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}} \dots\dots\dots(5.55)$$

Por consiguiente la correspondiente función de energía será:

$$U(f) = \sum_{\{i\} \in C_1} \ln \binom{M-1}{f_i} - \sum_{\{i\} \in C_1} \alpha_i f_i - \sum_{\{i,j\} \in C_2} \beta_{i,j} f_i f_j \dots\dots\dots(5.56)$$

Reduciéndose así al modelo autológico cuando  $M=1$ .

**Definición 5.5.-** Un auto-modelo es un auto-modelo-normal, llamado campo aleatorio de Markov gaussiano, si el conjunto de etiquetas  $L$  es  $\mathfrak{R}$  y la distribución de

probabilidad conjunta es normal multivariante [26 p.e]. Su función de probabilidad condicional es:

$$P\left(f_i \mid f_{N_i}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{1}{2\sigma^2} \left[ f_i - \mu_i - \sum_{j \in N_i} \beta_{I,J} (f_j - \mu_j) \right]^2} \dots\dots\dots(5.57)$$

La anterior ecuación refleja que se trata de una distribución normal con media condicional.

$$E\left(f_i \mid f_{N_i}\right) = \mu_i - \sum_{j \in N_i} \beta_{i,j} (f_j - \mu_j) \dots\dots\dots(5.58)$$

y varianza condicional

$$\text{var}\left(f_i \mid f_{N_i}\right) = \sigma^2 \dots\dots\dots(5.59)$$

La probabilidad conjunta es una distribución de Gibbs

$$p(f) = \frac{\sqrt{\det(B)}}{\sqrt{(2\pi\sigma^2)^m}} e^{\frac{1}{2\sigma^2} (f - \mu)^T B (f - \mu)} \dots\dots\dots(5.60)$$

Donde  $f$  se considera un vector,  $\mu$  es un vector  $m \times 1$  de las medias condicionales, y  $B = [b_{i,j}]$  es la matriz de interacción de  $m \times m$  cuyos elementos de la diagonal son 1 y el resto de los elementos  $(i,j)$  es  $-\beta_{i,j}$ , es decir,  $b_{i,j} = \delta_{i,j} - \beta_{i,j} = 0$ . Por lo tanto las funciones clique potenciales, de un solo estado y de dos estados, para el modelo auto-normal son:

$$V_1(f_i) = (f_i - \mu_i)^2 / 2\sigma^2 \text{ (un solo estado)} \dots\dots\dots(5.61)$$

$$V_1(f_i, f_j) = \beta_{i,j} (f_i - \mu_i) (f_j - \mu_j) / 2\sigma^2 \text{ (dos estados)}$$

**Definición 5.6.-** Otro modelo importante de mencionar el modelo simultáneo de auto-regresión (SAR). Este modelo, distinto del modelo auto-normal que está definido por  $m$  ecuaciones.

$$f_i = \mu + \sum_{j \in N_i} \beta_{i,j} (f_j - \mu_j) + \varepsilon_i \dots\dots\dots(5.62)$$

En donde  $\varepsilon_i$  son gaussianas independientes,  $\varepsilon_i \approx N(0, \sigma^2)$ . También genera la clase de todas las distribuciones normales multivariantes con función de probabilidad conjunta asignada en la ecuación (5.52).

**5.6.2 Modelo logístico multinivel (MLL)**

El modelo logístico multinivel (MLL) es una generalización del modelo auto-logístico. Al MLL también se le llama modelo generalizado de Ising. Hay M etiquetas discretas, en el conjunto de etiquetas  $L = \{1, \dots, M\}$ . Un clique potencial depende del tipo c (dirección, tamaño y orientación) de la clique y de la configuración local  $f_c = \{f_i | i \in c\}$ . Para cliques que contienen más de un estado ( $\#c > 1$ ) se definen las funciones clique potenciales como:

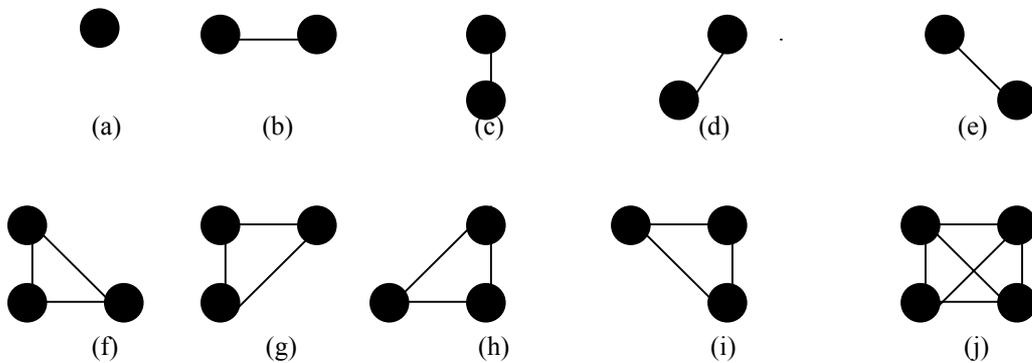
$$V_c(f) = \begin{cases} \zeta_c & \text{si todos los estados en c tienen la misma etiqueta} \\ -\zeta_c & \text{en otro caso} \end{cases} \dots\dots\dots(5.63)$$

donde  $\zeta_c$  es el potencial para los cliques de tipo c; para cliques de un solo estado depende de la etiqueta asignada al estado:

$$V_c(f) = V_c(f_i) = \alpha_I \text{ si } f_i = I \in L_d \dots\dots\dots(5.64)$$

donde  $\alpha_I$  es el potencial para la etiqueta con valor I.

A continuación en la figura 5.5 se muestra los tipos de cliques así como los parámetros asociados en sistemas de vecinos de segundo orden (es decir 8 vecinos).



**Figura5.5 Tipos de cliques y parámetros potenciales asociados en sistemas de vecinos de segundo orden 8 vecinos, los estados son los puntos y las líneas muestran las reacciones entre ellos [26 p.e].**

Asumiendo un modelo MLL de segundo orden, si y solo si  $\alpha$  (para cliques de un solo estado) y  $\beta$  (para cliques de dos estados) son diferentes de cero. Entonces la función potencial para cliques de dos estados se define como:

$$V_2(f_i, f_j) = \begin{cases} \beta_c & \text{si los estados de la clique } \{i, j\} = c \in C_2 \\ -\beta_c & \text{en otro caso} \end{cases} \dots\dots\dots(5.65)$$

Donde  $\beta_c$  es el parámetro  $\beta$  para las cliques de tipo c y  $C_2$  es el conjunto de cliques de dos estados. Para un sistema de cuatro vecinos hay cuatro tipos de cliques de dos

estados y cuatro tipos de  $\beta_c$ . Cuando el modelo es isotrópico los cuatro toman el mismo valor. Debido a su simplicidad el modelo en parejas MLL se ha utilizado extensamente para modelar regiones y texturas. Cuando el modelo MLL es isotrópico, la probabilidad condicional puede ser expresada como:

$$P\left(f_i = I \mid f_{N_i}\right) = \frac{e^{-\alpha_i - \beta_{n_i}(I)}}{\sum_{I=1}^M e^{-\alpha_i - \beta_{n_i}(I)}} \dots\dots\dots(5.66)$$

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

En el caso de que solo existan dos etiquetas, 0 y 1, se reduce a :

$$P\left(f_i \mid f_{N_i}\right) = \frac{e^{-\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}}{\sum_{f_i \in \{0,1\}} e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}} = \frac{e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}}{1 + e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}} \dots\dots\dots(5.67)$$

### 5.7 Análisis de clase de rechazo

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.

La etapa de segmentación puede ser implementada en base a que los píxeles centrales de los clusters sean obtenidos usando un algoritmo no contextual; entonces, basado en una clasificación bayesiana (como la propuesta en el capítulo 3), se puede llevar a cabo una primera etapa de segmentación, definiendo una “clase de rechazo”; finalmente este píxel rechazado será reclasificado vía un modelo de Markov.

La segmentación de la imagen puede ser propuesta por varios caminos: desde el concepto básico de *distancia mínima*, sustentada en la teoría de los CAM, propiamente en algoritmos en lo que a clasificación se refiere. Estos algoritmos muestran ventajas y desventajas pudiendo ser usados en combinación para obtener buenos resultados.

El esquema de segmentación requiere de los siguientes conceptos [3, 9 art]:

- a) Distancia-mínima.- Este concepto es obtenido de manera intuitiva partiendo del hecho de que entre elementos similares en un espacio dado pueden ser expresados en términos de su proximidad para cada conjunto de clusters.

b) Máxima-probabilidad.- Se consideran que los píxeles correspondientes a las diversas clases siguen un patrón o ley de probabilidad, es decir, cada clase puede ser descrita por medio de una función de probabilidad. Esta función se asemeja a la distribución real de cada clase, permitiendo así determinar la probabilidad de pertenecer a una clase dada para cada uno de los píxeles que componen la imagen.

c) Modelo markoviano.- Usando la teoría de CAM, y el concepto de estimación máximo a posteriori (MAP), minimizando la función de energía puede obtenerse por medio del algoritmo “*recocido simulado*” (su implementación y desarrollo se explica en el capítulo 6) el cual es controlado por medio de un parámetro de *temperatura* ( $T$ ), el cual se decrementa lentamente.

Otro desarrollo de modelos basados en contexto es el llamado “relaxation labeling” (RL) [Rosenfeld et al. 1976]. RL es una clase de procedimientos iterativos que utilizan restricciones de contexto para reducir ambigüedades en el análisis de imagen.

Definiendo un etiquetado homogéneo para ciertas regiones en la imagen el cual comprende estructuras bien definidas, haciendo referencia a pequeñas estructuras cuyas propiedades son realmente cercanas a un cluster o grupo de píxeles, lo que implica que su representación debe corresponder a determinada clase; y de esta forma ser preservadas estas estructuras.

Se define un criterio de rechazo dentro del concepto de máxima probabilidad, definiendo un parámetro llamado *umbral de rechazo ó limite de rechazo*, de tal forma que los píxeles cuya máxima probabilidad de pertenencia para una clase dada es menor este umbral, estos píxeles serán asignados a una nueva clase  $k+1$  (rechazo o clase de rechazo ó clase no definida), en otro caso estos píxeles son asignados a la función de máxima probabilidad que será maximizada [3 art].

Este umbral o límite de rechazo se toma de a cuerdo al valor de máxima probabilidad (0-1), tomando en cuenta que los valores que se encuentren por debajo de este umbral son asignados a la clase de rechazo; cuyas finas estructuras serán claramente diferenciadas sin ser alteradas. Un algoritmo que cumple a la perfección el criterio de máxima probabilidad es el llamado ICM (modo condicional iterativo) el cuál es implementado y desarrollado dentro del capítulo 6, sumamente útil para la relajación determinística.

Este criterio puede actuar como una reclasificación, pero que afecta solamente a los píxeles asignados para la clase de rechazo, tal y como trabaja el modelo de Potts cuya implementación y desarrollo también se tocan en el capítulo 6.

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

## Capítulo 5. Segmentación por campos aleatorios de Markov

---

En los casos donde las etiquetas son independientes unas de otras, la probabilidad conjunta es el producto  $P(f) = \prod_{i \in S} P(f_i)$

Lo cual implica también la independencia de la probabilidad condicional:

$$P(f_i / \{f_j\}) = P(f_i) \dots\dots\dots(5.68)$$

cuando  $i \neq j$

Por lo tanto, un etiquetado global  $f$  puede calcularse considerando cada etiqueta local  $f_i$ . Esto es ventajoso para solucionar el problema.

# 6 Esquemas de relajación

## 6.1 Introducción

Para propósitos de la presente tesis es de particular interés analizar esquemas de relajación aplicables a nuestros resultados de segmentación. En los métodos de segmentación puntuales y en los contextuales se busca tener buenos resultados proporcionando un modelo probabilístico (el más adecuado a la fuente de datos). Así, los resultados de segmentación pueden ser de gran calidad pero frecuentemente arrojando píxeles fuera del contexto local de segmentación (píxeles aislados). Obtener regiones más homogéneas y con ocurrencia de menos píxeles aislados es un propósito de esta tesis.

Los métodos de relajación tienen aplicación en problemas de visión por computadora, de la detección de bordes a la interpretación de imágenes en base al etiquetado de escenas [31 p.e].

El trabajo a desarrollar en estos métodos, sobre el etiquetado de la escena con relación a la incertidumbre de los valores de los píxeles de esta escena, es a través de **una aproximación (approach) discreta** de la relajación en el cual cada componente de la escena es asignado un sistema de interpretaciones posibles. Los etiquetados desfavorables se cambian examinando en primer lugar pares de la etiqueta mediante segmentos conectados, y asegurándose en segundo lugar de que las consistencias localmente definidas se pueden analizar juntamente para obtener un mejor resultado.

Los elementos básicos en métodos de relajación envuelven sistemas de características de los objetos respecto a un sistema de etiquetas. En el contexto de la visión, las características son generalmente puntos, bordes y superficies [31 p.e].

De una manera un tanto general se puede afirmar que el principal trabajo de los métodos de relajación, consiste en agrupar píxeles en conjuntos de regiones o etiquetas, de forma iterativa y conforme se relajan las condiciones de pertenencia. Partiendo de una serie de estimaciones de pertenencia inicial que se van refinando iterativamente, de acuerdo a datos del píxel a regularizar dentro de un contexto local, ya sea a través del nivel de gris, la media local, o la textura de la escena que se va analizar.

## 6.2 Marco teórico

Los CAM proporcionan las medidas de probabilidad sobre un dominio de definición que tenga relaciones de tipo espacial o temporal. El conjunto de posiciones donde se define el campo se denomina malla o rejilla, denotándose por un conjunto  $S$ .

En relación al contorno de este conjunto  $S$  se trazan diversos radios de acuerdo a las siguientes estructuras:  $\diagup$ ,  $\diagdown$ ,  $|$  y  $\_$ . Para cada posición  $s \in S$  se va a definir un espacio de estados  $A_s$  correspondiente a los niveles de gris. Para el contorno,  $s$  es un radio y el espacio de estados  $A_s$  para cada radio corresponde a cada uno de los puntos del radio por donde se establece el contorno [2 art].

Con  $x_s \in A_s$  se denota un valor de gris para del píxel o una posición para un radio correspondiente. Sea  $\Omega = (A_s)$  el espacio de todas las configuraciones, es decir será el conjunto de texturas o contornos posibles que se pueden definir en  $S$ . Definiendo  $x = (x_s)_{s \in S}$  se denota una configuración en particular, es decir una textura o contorno. Suponiendo que todos los espacios y configuraciones son finitos y todos los píxeles tienen el mismo intervalo de niveles de gris para el caso de la textura y el mismo intervalo de posiciones para todos los radios en el caso del contorno, se dice que el espacio  $\Omega$  es homogéneo.

Lo anterior establece la definición del campo. Ahora definiendo una medida de probabilidad o distribución de probabilidad en el espacio de configuraciones  $\Omega$ , si a cada configuración (textura o contorno)  $x \in \Omega$  se le asigna una probabilidad  $P(x) \geq 0$ , tal que

$$\sum_{x \in \Omega} P(x) = 1 \dots\dots\dots(6.1)$$

Para el caso de un suceso en el que  $E \subset \Omega$  corresponderá a un conjunto de texturas o un conjunto de contornos para los que se define su probabilidad simplemente como:

$$P(E) = \sum_{x \in E} P(x) \dots\dots\dots(6.2)$$

Los sistemas de vecinos para los métodos de relajación convienen que sean isótropos. Esto quiere decir que se comportarán de la misma manera en todas las direcciones de manera que sean invariantes a rotaciones en las direcciones.

En el caso del modelo de textura, sea  $s = (i, j)$  y  $t = (k, l)$  dos píxeles de la imagen a procesar tal que  $t \in \mathcal{N}(s)$ ; se define el orden  $c$  del vecindario (para el caso homogéneo e isótropo) como el menor entero que cumpla:

$$c \geq (k-i)^2 + (l-j)^2 \quad \text{para todos los vecinos } t = (k, l) \text{ de } s = (i, j) \dots\dots\dots(6.2)$$

En el caso del modelo de contorno, si  $s$  y  $t$  son radios tales que  $t \in \mathcal{N}(s)$ , se define el orden del vecindario (para el caso homogéneo e isótropo) como el menor entero que cumpla:

$$c \geq |s - t| \quad \text{para todos los vecinos } t \text{ de } s. \dots\dots\dots(6.3)$$

Dado un sistema de vecindario  $\mathcal{N}$  en  $S$ , se dice que un subconjunto  $C \subset S$  es un clique si dos o más elementos cualesquiera  $C$  (diferentes entre sí) son vecinos. Entonces el conjunto

de todos los cliques de  $S$  se denotarán como  $C$ . En el caso de vecindarios homogéneos los cliques se van poder clasificar en tipos, según la relación espacial entre los elementos que lo forman.

Para un campo aleatorio las dependencias de  $S$  van a ser en general locales. Esto quiere decir que un píxel va a depender de los píxeles cercanos o mejor dicho de sus píxeles vecinos, y que en un contorno el valor de éste píxel en una periferia determinada va a depender del valor del contorno en las periferias cercanas. Por esta razón se va a definir para cada posición  $s \in S$  un conjunto  $\mathcal{A}(s) \subset S$ , que corresponde a las posiciones de  $S$  de las que  $s$  depende. Los elementos  $\mathcal{A}(s)$  se denominan vecinos de  $s$ . De esta manera la colección de conjuntos  $\mathcal{A} = \{\mathcal{A}(s) : s \in S\}$  se denomina sistema de vecindario de la rejilla  $S$ . Un sistema de vecindario debe de cumplir dos propiedades principalmente:

- a) Que  $s$  no sea vecino de sí mismo:  
 $s \notin \mathcal{A}(s)$ .
- b) Que si  $s$  es vecino de  $t$ , éste último lo sea del primero:  
 $s \in \mathcal{A}(t) \Leftrightarrow t \in \mathcal{A}(s)$ .

Se establece que en general, los sistemas de vecindarios son *homogéneos*, es decir que sabiendo la posición de un píxel  $s \in S$  siendo  $S$  el conjunto de píxeles en una estructura matricial o rejilla; se puede determinar cuáles son los vecinos de otra posición  $t$  (conjunto de píxeles vecinos), desplazando a  $t$  de  $s$ ; afirmando de esta manera que son invariantes en el espacio.

### 6.3 Relajación probabilística

Una metodología combinada con el umbral de una imagen y con el proceso de relajación probabilística por medio de etiquetas es el RPL ((labeling probabilistic relaxation), propuesto para segmentaciones automáticas. La imagen primeramente se toma como un umbral, y se asigna un etiquetado individual de acuerdo a la distancia de cada píxel con el grupo o cluster, resultado de este umbral [1 art].

Un criterio *heurístico* es definido para determinar el número de clases que existe en cada sub-imagen. De este modo el costo de la programación se reduce si el proceso es implementado con múltiples niveles de resolución.

Para la relajación probabilística se presenta una técnica combinando el umbral para una segmentación automática y el proceso de relajación probabilística por medio de etiquetas (RPL). En particular, la técnica es ilustrada particionando la imagen por medio de ventanas o estructuras de celdas, dentro de diversas clases (pavimento, vegetación, construcción, agua, etc), de modo que la RPL es inherentemente sensible a la probabilidad del proceso de etiquetado, solucionando este problema con la técnica de la relajación probabilística. La ventaja de este método es tanto en costo de programación como la propiedad de suavizar la imagen.

**6.3.1 Umbralización**

A continuación se explican los fundamentos de técnicas de umbralización, las cuales tendrán empleo en el esquema de relajación probabilística. Las técnicas de umbral proveen una rápida y simple aproximación a la segmentación de la imagen. En la mayoría de las técnicas utilizadas, la información la representan en forma de histograma, en la selección del umbral, o se trata de obtener la mejor partición posible del histograma, con respecto al mismo criterio. De cualquier modo estos métodos son sensibles al ruido y por lo tanto propensos al error especialmente cuando la distribución de los objetos se superpone o coinciden unos con otros. Una muestra son las imágenes donde pueden confundirse los distintos tonos de gris, de la vegetación, el pavimento y la construcción.

Características de métodos de umbrales son:

- No se requiere una búsqueda exhaustiva y óptima.
- No se asume una distribución de probabilidad a priori.
- Puede ser obtenido un único umbral óptimo.

Para reducir errores locales, la imagen puede ser suavizada antes de hacer el umbral, así también podría disponerse de algún filtro adaptivo, con lo cual las diferentes regiones serían grandemente reducidas u homogeneizadas. No obstante algunos detalles finos son omitidos con respecto a ciertas regiones de la imagen, debido a la forma confusa de algunas regiones, es por eso que frecuentemente se opta por usar algún tipo de filtrado.

**6.3.2 Etiquetado de la relajación probabilística**

Frecuentemente los esquemas de etiquetado son **probabilísticos**. Las probabilidades se asignan a cada etiqueta con una estimación de la probabilidad de etiqueta de forma particular, determinando si es la correcta para esa característica. Las aproximaciones probabilísticas entonces son utilizados para maximizar (o reducir al mínimo) las probabilidades por el ajuste iterativo, considerando las probabilidades asociadas a las características vecinas [1 art]. Las estrategias de la relajación no garantizan necesariamente convergencia.

Ahora consideremos la estrategia de etiquetado más detalladamente. Asumiendo:

- es el sistema  $\{o_1, \dots, o_n\}$  de características del objeto de  $n$  que se etiquetarán.
- $L$  es el sistema  $\{I_1, \dots, I_m\}$  de las etiquetas posibles de  $m$  para las características.

Sea  $P_i(I_k)$  la probabilidad que la etiqueta  $I_k$  sea la etiqueta correcta para la característica del objeto  $o_i$ .

Los axiomas generalmente de la probabilidad se pueden aplicar a lo siguiente:

- Cada probabilidad satisface  $0 \leq P_i(I_k) \leq 1$  donde  $P_i(I_k) = 0$ , implica que la etiqueta  $i_k$  es imposible para la característica  $o_i$  e  $P_i(I_k) = 1$ , implica que la etiqueta es asignada correctamente.
- El sistema de etiquetas es *mutuamente exclusivo* y *exhaustivo*. Así podemos escribir para cada  $i$ :  $\sum_L P_i(I_k) = 1$

Así cada característica es descrita correctamente por *exactamente una* etiqueta del sistema de etiquetas. El proceso de etiquetado comienza con una inicial, y quizás con la arbitraria asignación de las probabilidades de las etiquetas. El algoritmo básico entonces transforma estas probabilidades en un nuevo sistema según la relajación.

Se repite este proceso hasta que el método de etiquetado converge o se estabiliza. Esto ocurre cuando poco o nada de cambio ocurre entre los sistemas sucesivos de valores de la probabilidad.

Los métodos populares toman a menudo *estrategias* estocásticas para actualizar las funciones de la probabilidad. Un operador puede considerar *la compatibilidad* de las probabilidades de las etiquetas como apremios en el algoritmo de etiquetado.

La compatibilidad  $C_{ij}(I_k, I_l)$  es una correlación entre las etiquetas definidas como la probabilidad condicional que la característica  $o_i$  tiene una etiqueta  $I_k$  dado que la característica  $o_j$  tiene una etiqueta  $I_l$ , es decir:  $C_{ij}(I_k, I_l) = P(I_k | I_l)$

Así, la actualización de las probabilidades de etiquetas es hecha considerando las probabilidades de las etiquetas para las características vecinas. Asumiendo que se han cambiado todas las probabilidades hasta un cierto paso,  $S$ , y ahora buscamos una probabilidad actualizada para el paso siguiente  $S+1$ .

Podemos estimar el cambio en  $P_i(I_k)$  con lo siguiente:

$$\delta P_i(I_k) = \sum_{j \in N} w_{ij} \left[ \sum_{l \in L} C_{ij}(I_k, I_l) P_j(I_l) \right]$$

donde el sistema de características vecinas,  $o_i$  y  $w_{ij}$  es un factor  $N$  de las etiquetas de los vecinos, definidos de una manera tal que:  $\sum_{j \in N} w_{ij} = 1$ .

La nueva probabilidad  $P_i(I_k)$  para la etiqueta en la generación  $S+1$  se puede programar de los valores *de la generación S*

$$P_i(I_k) = \frac{P_i(I_k)[1 + \delta P_i(I_k)]}{\sum_l P_i(I_l)[1 + \delta P_i(I_l)]} \dots\dots\dots(6.4)$$

El proceso de relajación probabilística por medio de etiquetas permite la introducción de la información contextual global, para reducir la ambigüedad local en el etiquetado de un píxel. El RPL, también provee una “limpieza general” para segmentaciones demasiado ruidosas. El RPL puede ser definido asumiendo lo siguiente:

Sea  $I_i$  un grupo o clase de píxeles y sea  $(L_1, L_2, \dots, L_m)$  siendo el posible grupo de clases o etiquetas. Con cada píxel  $I_i$ , a m-dimensional el vector de probabilidad  $(p_i(L_1), p_i(L_2), \dots, p_i(L_m))$ , es inicialmente asignado. La componente  $p_i(L_k)$  indica la probabilidad de que  $I_i \in L_k$ , donde  $L_k \in (L_1, L_2, \dots, L_m)$ , y se satisface

$$0 \leq p_i(L_k) \leq 1$$

$$\sum_{j=1}^m p_i(L_j) = 1 \dots\dots\dots(6.5)$$

Las probabilidades de etiquetado son entonces iterativamente actualizadas de forma independiente de acuerdo con una nueva  $p_i(L_j)$ , programada de la siguiente manera heurística [1 art]:

$$p_i^{n+1}(L_j) = \frac{p_i^{(n)}(L_j)(1 + q_i^{(n)}(L_j))}{\sum_{k=1}^m p_i^{(n)}(L_k)(1 + q_i^{(n)}(L_k))} \dots\dots\dots(6.6)$$

donde

$$q_i^{(n)}(L_k) = \frac{1}{c} \sum_{K \in N(i)} \sum_{\lambda=1}^m r_{ij}(L_k, L_\lambda) p_j^{(n)}(L_\lambda) \dots\dots\dots(6.7)$$

donde  $N(i)$  denota la 8-conexidad entre los vecinos, asociada al píxel  $I_i$ , y  $c$  es el número de píxeles de tales vecinos.

El coeficiente  $r_{ij}(L_k, L_\lambda)$  refleja el grado al cual la clasificación dentro de cada clase  $L_k$  es compatible con la clasificación de sus vecinos del píxel  $I_j$  dentro de la clase  $L_\lambda$ . El coeficiente  $r_{ij}(L_k, L_\lambda)$  puede ser elegido empíricamente o estimado usando la información mutua. En este experimento, podemos definir  $r_{ij}(L_k, L_\lambda)$  como:

$$r_{ij}(L_k, L_\lambda) = 1 \text{ si } k = \lambda$$

sino

$$r_{ij}(L_k, L_\lambda) = 0 \dots\dots\dots(6.8)$$

El significado de tal definición es que al encontrar un par de píxeles adyacentes que tengan la misma etiqueta podrán ser asumidos completamente compatibles uno con otro.

Tomando en cuenta la combinación del umbral con el RPL, la imagen es inicialmente recorrida en umbral y con una probabilidad de vector estimada de acuerdo a la distancia de Mahalanobis (que se definió propiamente dentro del capítulo 3) entre los píxeles de la clase  $I_i$  y las clases  $k$ , entonces esta estimación nos lleva a:

$$p_i^{(0)}(L_k) = \frac{1/d_{ik}}{\sum_{m=1}^m 1/d_{i\alpha}} \dots\dots\dots(6.9)$$

donde:

$$d_{ik} = \frac{(I_i - \mu_k)^2}{\delta_k^2} \dots\dots\dots(6.10)$$

con  $\mu_k$  y  $\delta_k^2$  denotan la media y la varianza de la clase  $k$ .

El proceso es iterativo y llega al final cuando un 99% de los píxeles son etiquetados con un componente de probabilidad mayor a 0.9. Dentro del capítulo final de resultados (capítulo 8) se podrá apreciar que los errores son reducidos en gran medida comparados con diferentes números de iteraciones, es decir, entre mayor sea el número de estas, aumenta la probabilidad de mejorar el resultado, obteniendo una imagen de mayor limpieza morfológica. Se puede asumir el siguiente criterio de mejora en la determinación de ventanas de análisis:

Cada ventana  $w_i$  esta definida sobre una sola clase. Posteriormente se tomará en cuenta el umbral más bajo y se le asignará a  $t_{Li}$  y el umbral más alto a  $t_{Hi}$ , de acuerdo a los siguientes pasos:

Paso1.- Asumir ventanas  $w_i$  con distintas clases; además determinar el umbral  $t_L$  y  $t_H$  para las ventanas. Entonces el criterio será:

$$\begin{aligned} p_1 p_2 (\mu_1 - \mu_2)^2 &> g_1 \\ p_2 p_3 (\mu_2 - \mu_3)^2 &> g_2 \\ p_{i-1} p_i (\mu_{i-1} - \mu_i)^2 &> g_i \dots\dots\dots(6.11) \end{aligned}$$

si lo anterior se satisface, entonces se asignará:

$$t_{Li} = t_L; \quad \text{y también} \quad t_{Hi} = t_H;$$

y se detiene, si no ir al paso 2 definido como:

Paso 2.- Asumir ventanas  $w_i$  consiste de  $i-1$  clases; entonces se determinará el umbral óptimo  $t$  de cada ventana como se hizo en el paso 1 de acuerdo a (6.11):

Si se satisface entonces se asignará:

$$\begin{array}{lll}
 t_{Li} = t; & t_{Hi} = t_{gH}; & \text{si } t \leq (t_{gL} + t_{gh}) / 2 \\
 t_{Li} = t_{gL}; & t_{Hi} = t; & \text{si } t > (t_{gL} + t_{gh}) / 2
 \end{array}$$

entonces  $w_i$  consiste de una clase solamente, y se asigna

$$t_{Li} = t_{gL}; \quad t_{Hi} = t_{gH}$$

El simple, pero robusto, criterio definido en el paso 1 provee hasta cierto punto la separación entre diferentes clases de objetos, con el valor medio de  $\mu_i$  de la probabilidad  $p_i$  respecto a la  $i$ -ésima clase del píxel, respecto a la ventana  $w_i$ . Declarando a  $t_{gL}$  y a  $t_{gH}$  como el umbral bajo y alto respectivamente, de la imagen o imágenes a procesar;  $g_1, g_2, \dots, g_i$  son constantes predeterminadas.

Una vez que los umbrales para cada ventana son determinados, el vector de probabilidad inicial asociado a cada píxel en una ventana particular puede ser determinado. La discontinuidad en el caso de cada una de las ventanas puede ser eliminada por medio de un filtro paso bajas, aunque en nuestro caso, no fue necesario, debido a que se logró programar el tamaño de las ventanas de forma que fueran todas de la misma dimensión, con respecto a la menor, de modo que no importa cual es la región específica a tomar de la imagen, siempre se tendrá una clase bien definida.

## **6.4 Relajación determinística**

### **6.4.1 Algoritmo modo condicional iterativo (ICM)**

Para el propósito de esta tesis se necesita trabajar en una estimación sobre el problema de segmentación, en base a un modelo aproximado de máximo a posteriori (MAP). Para esto nos basamos en los CAM, con el criterio de Gauss, para la definición de criterios de etiquetado de los píxeles, así como en el algoritmo ICM (Modo condicional iterativo).

La segmentación de la imagen es un problema importante, en lo que a programación se refiere, para percepción y análisis de la imagen. Usando los CAM y el criterio de Gauss, para el etiquetado de píxeles, expresando así su dependencia con los parámetros de relajación.

El algoritmo ICM, aproxima la proporción esperada de un etiquetado de píxeles correcta. El método ICM depende de una eficiente programación. La búsqueda de una correcta solución avanza iterativamente desde una configuración inicial de etiquetado hacia una configuración sucesora, solamente si la configuración sucesora tiene una distribución posterior alta [4 art]. Hasta aquí, el algoritmo garantiza la convergencia en un número finito de pasos. En otro caso, el algoritmo ICM es garantizado para alcanzar la búsqueda adecuada con respecto a una preclasificación, de modo que el algoritmo sirve para dar a una región preclasificada una forma regular, intentando mejorar esta preclasificación.

Cuando el conjunto de etiquetas es discreto, minimizar la energía a posteriori se convierte en un problema combinatorio. Como resulta difícil maximizar la probabilidad conjunta de un CAM, Besag, propuso un algoritmo determinístico llamado ICM que maximiza las probabilidades condicionales locales secuencialmente de la siguiente manera.

#### **6.4.2 Implementación del algoritmo ICM utilizado en la relajación determinística**

En el caso de este trabajo de tesis es de sumo interés encontrar regiones homogéneas en la imagen; para esto utilizamos la relajación determinística. En este caso, el contar con la ayuda de la relajación determinística, usando un sistema de vecindad espacial entre píxeles, asegura prácticamente una convergencia de un mínimo absoluto en el costo de la función, en otras palabras, será un costo menor en lo que a la programación y tiempo se refiere. En la relajación determinística, a diferencia de la relajación estocástica (la cual permite incrementos y decrementos en la transición de la energía por medio de la temperatura) solo se permiten las transiciones en decremento, por esto el algoritmo ICM parte del sentido de una segmentación correcta, a la cual se alcanzó mediante la segmentación bayesiana.

Existe un inconveniente en la utilización de este algoritmo; y es que la solución obtenida por este método será dependiente del orden en que se visiten los píxeles; esto conlleva a que se generen estructuras parásitas dentro de la imagen de segmentación a analizar. De lo anterior se propone un método de visita aleatoria para el total de píxeles que comprenden la imagen.

Se asume la convergencia del algoritmo ICM, cuando los píxeles locales son actualizados en una forma asíncrona, aunque se puede tener una convergencia "razonable" si los píxeles son visitados en un "*orden aleatorio*" (por llamarlo de alguna manera) tal que ellos no se traslapen unos a otros con sus vecinos [29 p.e].

Por lo tanto se define un camino  $(r, c) = w(i)$  tal que se visita todos los píxeles aleatoriamente. Primero se generan todas las posibles combinaciones renglón-columna (coordenadas) de acuerdo a las dimensiones de la imagen. Estas combinaciones se meten en un vector (arreglo), y se van sacando en desorden. De esta forma se asegura que se lean todos los puntos de la imagen, solo una vez cada uno, y sin un orden determinado. Cuando se termina una iteración se "regresan" todas las combinaciones al vector y se vuelven a leer, pero en un nuevo orden aleatorio por lo que también en cada iteración el barrido se hace de forma distinta, conservando hasta cierto punto la característica "aleatoria" (o pseudoaleatoria) del proceso.

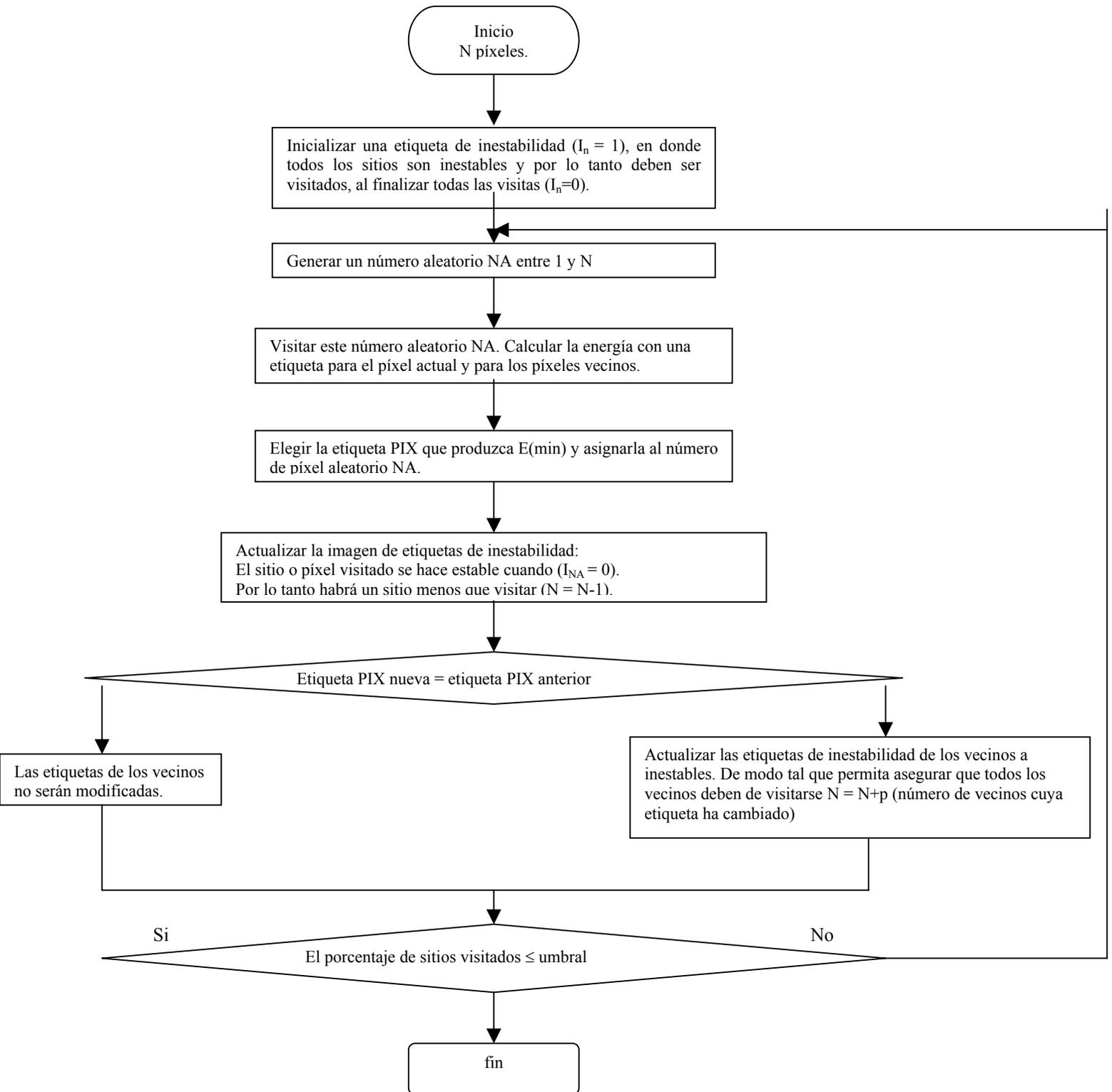
Para las ventanas de entrenamiento, las pruebas mostraron que el mejor camino es aquel que visita los píxeles en este orden aleatorio. Aunque cabe mencionar que procura visitar primero los píxeles en la frontera exterior, luego los píxeles en la frontera interior restante, etc., y así sucesivamente hasta hacer el barrido de toda la imagen.

En cada una de las visitas, los píxeles son visitados saltando de uno a otro. Puesto que los píxeles son visitados de esta manera, estos no se traslapan unos a otros con sus vecinos. En la práctica el barrido en forma aleatoria acelera la convergencia, y además permite que el barrido sea más eficiente. De este modelo de barrido surge lo siguiente [2 art]:

Mejora 1. Visitar los píxeles en forma aleatoria de afuera hacia adentro, desde la frontera, saltando de un nodo a otro.

Mejora 2. Mover los píxeles que no han sido movidos de ninguna manera, con una probabilidad  $p$  a una posición seleccionada aleatoriamente en una “8-vecindad”. Un nodo, que no es movido por el algoritmo durante una visita local, será movido con probabilidad  $p$  hacia un vecino arbitrario seleccionado de manera aleatoria.

En seguida se presenta un diagrama que pretende dar a conocer este razonamiento de visita aleatoria.



El porcentaje máximo al final es una medida de la exigencia que se desea imponer a la optimización del algoritmo, así como para detener el proceso. De esta manera se asegura que de acuerdo al mayor número de iteraciones deseadas, se alcance la convergencia con base al porcentaje de etiquetas de segmentación entre una iteración y otra.

Los pasos a seguir en el algoritmo ICM son:

- 1).- Comenzar en una buena configuración inicial  $\omega^0$ , con  $k = 0$ , siendo  $k$  el número de iteración.
- 2).- En cada iteración  $k$ 
  - a) escoger aleatoriamente las coordenadas de un píxel
  - b) a partir de la configuración actual  $\omega^k$ , determinar posibles configuraciones de energía,  $U(\eta)$ , donde:
 
$$\eta \in N_{\omega^k}$$
  - c) seleccionar la configuración de energía mínima
 
$$\omega^{k+1} = \arg \min_{\eta \in N_{\omega^k}} U(\eta)$$
  - d) regresar a 2a) hasta obtener el barrido de toda la imagen
- 3).- Repetir el punto 2, con  $k=k+1$ , hasta alcanzar el número fijado de iteraciones.

Para nuestro caso, las configuraciones de energía fueron calculadas empleando vecindades de 8-conexidad  $U_i(\eta) = -\sum \beta \delta_{ij}$ , donde:

- $i$  es la clase del píxel analizado
- $j$  es la clase de un píxel vecino y
- $\delta$  es la función impulso unitario.

Durante el cálculo de  $U(\eta)$ , debe de considerarse únicamente a las clases de los píxeles contenidos en la vecindad de análisis.

### **6.5 Relajación estocástica**

De acuerdo a la definición del campo y de la medida de probabilidad o distribución de probabilidad en el espacio de configuraciones  $\Omega$  y a la ecuación (6.1); se dice que un campo definido en la rejilla  $S$  con espacio de configuraciones  $\Omega$  y medida de probabilidad asociada  $P(x_s)$  es un *campo aleatorio o estocástico* si para todo  $x \in \Omega$  se cumple que  $P(x_s) > 0$ , es decir, si la distribución o medida de probabilidad cumple la condición de positividad. Por lo tanto, si el modelo de textura es un campo aleatorio, significa que todas las configuraciones serán posibles (ya sea con mayor o menor probabilidad) [29 p.e]. En el caso del contorno, de la misma forma, significa que todas las configuraciones de contornos serán posibles.

La esencia del procedimiento de la Relajación estocástica basado en el modelo de Gibbs consiste en observar una secuencia estimada de un elemento aleatorio escogido y a su vez la generación de un nuevo valor para el elemento, en concordancia con la distribución de

probabilidad propuesta. Entonces el nuevo valor es aceptado o rechazado de acuerdo con algún cálculo estadístico en base a una cantidad no grande de una secuencia de elementos vecinos.

El algoritmo para la reconstrucción por relajación estocástica RSR, puede ser visto como un algoritmo que continuamente se prueba para transformar la configuración actual de un píxel con sus píxeles vecinos.

### 6.5.1 Implementación

Como ya se definió previamente el propósito de la segmentación de la imagen, la imagen representada tiene que ser modelada de manera que de acuerdo a  $k$  diferentes tipos de regiones, cada una de ellas con una distribución normal caracterizada por un valor de su media  $\mu_k$  y su varianza  $\sigma_k^2$ . Ahora tomando en cuenta la distribución de Gibbs (la cual ha sido definida en el Capítulo 4) mientras  $T$  tiende a 1, una nueva medida de probabilidad generada de ese modo el píxel  $(i,j)$  debe de pertenecer a la región de tipo  $q_k$ , siendo estabilizado mediante la siguiente forma:

$$P(X_{ij} = q_k | Y_{ij}) \propto \exp \left\{ -\frac{1}{2\sigma_k^2} Y_{ij} - \mu_k - \sum_{c \in C} V_c(x) \right\}_{x=q_k} \dots \dots \dots (6.12)$$

Este modelo de la ecuación anterior contiene 2 elementos. El primero es un dato dependiente y restringe la región que será cercana a al dato. El segundo elemento depende únicamente de los atributos del píxel central con el vecindario de píxeles de forma que se imponga continuamente espacial. De acuerdo a la definición de la función potencial, es claro que si el número de píxeles correspondiente a la región o clase  $k$  en la vecindad es exactamente el mismo, el segundo término de la ecuación puede tomar el mismo valor para todas las regiones de la misma clase. Lo anterior no implica un cambio en el orden de los valores relativos de las probabilidades y los atributos del píxel analizado solamente serán determinados por la información. En contraste con lo anterior si estos números son diferentes uno de otro, el valor de probabilidad correspondiente para el máximo ellos, puede adquirir el mayor incremento y el píxel central puede ser más probable de que sea asignado a la región asociada. Si el valor del término dependiente llega a ser insignificante, en el caso de que  $\beta$  sea muy grande en primera instancia, el atributo de que el píxel pueda ser decidido principalmente por el atributo de sus vecinos.

En otras palabras, el espacio restringido puede tener efecto solamente cuando los números de los píxeles correspondientes al tipo de región  $k$  en la vecindad del píxel central difieren uno de otro, causando así que el atributo de cada píxel en la segmentación de la imagen sea más consistente con la mayoría de sus vecinos.

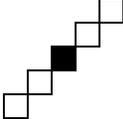
En esta etapa la producción de la imagen segmentada, es conocida como una reconstrucción. Tal reconstrucción usual y obviamente es un procedimiento matemático, el cual se define como relajación estocástica.

Se pueden asumir diferentes propuestas de conjuntos de configuración de vecindad. Tomando como base el conjunto integrado por 8-conexidad, y cuatro arreglos de vecindad en un número de  $\psi$  píxeles (adicionando 2 renglones y dos columnas mas) [3 art] para adoptar las siguientes formas de análisis:  $\diagup$ ,  $\diagdown$ ,  $|$ , y  $\_$ . Las funciones de energía correspondientes para cada configuración clique, serán las siguientes:

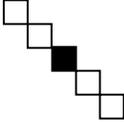
**Configuración 1:**

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i=j}$	
-------------------------------------------------	-------------------------------------------------------------------------------------

**Configuración 2:**

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{-i=j}$	
--------------------------------------------------	--------------------------------------------------------------------------------------

**Configuración 3:**

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i=-j}$	
--------------------------------------------------	---------------------------------------------------------------------------------------

**Configuración 4:**

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{0j}$	
------------------------------------------------	---------------------------------------------------------------------------------------

**Configuración 5:**

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i0}$	
------------------------------------------------	---------------------------------------------------------------------------------------

donde  $d_{i,j}$  representan las configuraciones de vecindades asociadas para el clique analizado, y el coeficiente  $\beta$  representa el factor de homogeneidad. En este caso todas las configuraciones tienen un valor de  $\psi = 5$ . Se observan las distintas configuraciones en las que el cuadrado del centro (a color) representa el píxel  $s$  y el cuadrado negro cada uno de

los píxeles  $t$  vecinos de  $s$ . Puesto que la imagen definida en la rejilla  $S$  tiene dimensiones finitas, los vecindarios de los píxeles cerca del borde de la imagen no pueden ser iguales que los de los vecindarios de los píxeles interiores. El concepto de homogeneidad no se puede cumplir para los píxeles de los bordes; sin embargo esto siempre ocurrirá, con la alternativa de hacer de mayor tamaño el contorno de la imagen a procesar. Pero si este no fuese el caso se asume un vecindario homogéneo suponiendo implícitamente el efecto de los bordes.

Posteriormente y después de analizar todas y cada una de las configuraciones de vecindades adaptivas, se procede a hacer una comparación entre ellas mediante el modelo de Potts, con el uso de un algoritmo que defina en base a la conectividad de las vecindades, cual de todas ellas tiene el menor decremento de la energía, para elegir la vecindad adaptiva que defina mejor el resultado de la regularización de la imagen.

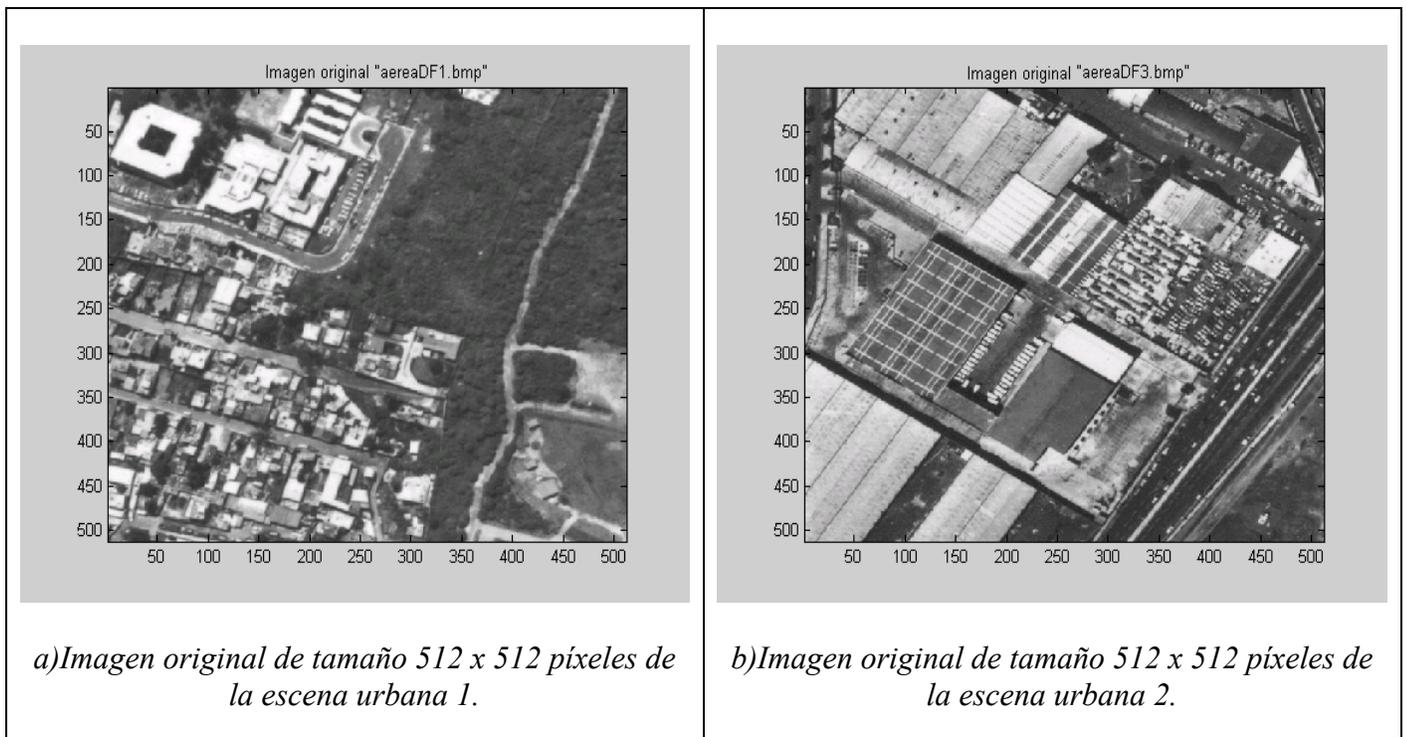
La corrección clasificación de píxeles aislados puede ser posible aplicando la teoría de la distribución de Gibbs, en el sentido de la vecindad existente del píxel central con sus vecinos, para ser clasificados de manera correcta los píxeles aislados. Esto ocurre cuando una zona o área determinada es visiblemente homogénea. Teóricamente los bordes obtenidos en la imagen destacan la proximidad en contexto de los límites contenidos en ella. Con los trazos asignados de acuerdo a las seis configuraciones, tan pronto como son llevadas a cabo, el número de píxeles de la imagen se comportan de manera tal que crean zonas circunvecinas o contornos perfectamente definidos, eliminando píxeles redundantes.

# 7 Resultados

## 7.1 Imágenes a procesar

A continuación se muestran las imágenes de prueba, sobre las cuales se aplicarán los algoritmos de segmentación y de relajación considerados en el presente trabajo de tesis. Ambas imágenes son de tamaño de 512 x 512 píxeles, cuantizadas a 8 bits por píxel y provienen de fotografías aéreas pancromáticas de escena urbanas de la ciudad de México.

Como parte previa del análisis que se lleva a cabo en este trabajo de tesis se hace la identificación en cada imagen de prueba de cinco diferentes regiones características de la escena, en base a las distintas clases obtenidas mediante las ventanas de entrenamiento, las cuales son: asfalto, construcción, áreas verdes o zonas boscosas, sombras u otros, de acuerdo a características de homogeneidad en niveles de gris, a ciertos tipos de texturas o estructuras finas.



Para el caso de la función discriminante óptima de Bayes se asignaron diferentes ventanas de entrenamiento de acuerdo a las distintas etiquetas o regiones características que se desean determinar y diferenciar dentro de la imagen, en base a que cada una de ellas tiene una textura o estructura fina, distintas unas de otras.

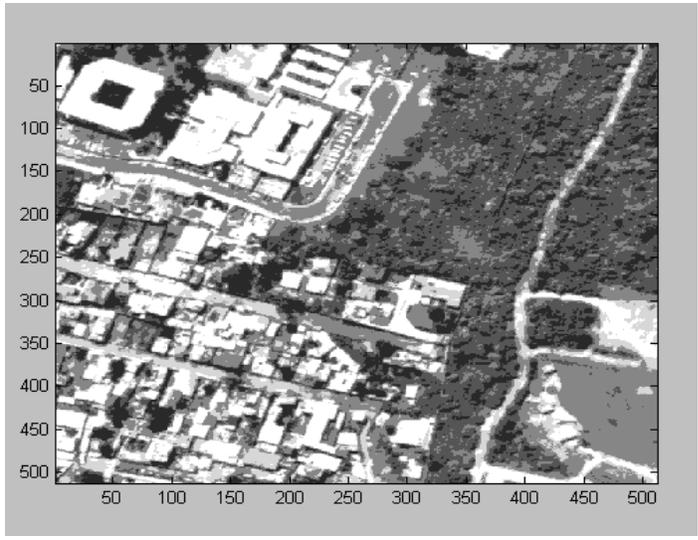
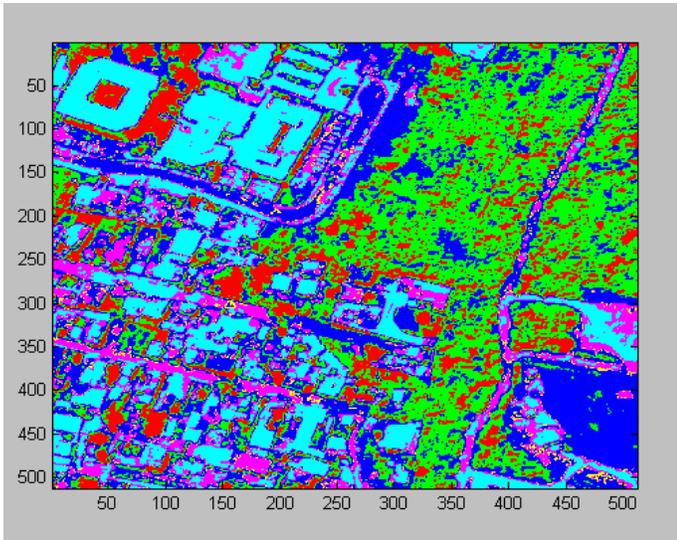


**7.2 Resultados de la segmentación aplicando la función discriminante de Bayes para la imagen 1**

*Imagen segmentada (en falso color) con la función discriminante, caso general:*

*Imagen segmentada (en niveles de gris) con la función discriminante, caso general:*

$$d_k^{iii}(x) = -\frac{1}{2}(x - m_k)^T C_k^{-1}(x - m_k) + [\log p(w_k) - \frac{1}{2} \log |C_k|] \quad k = 1, 2, \dots, M$$



*Imagen segmentada (en falso color) con la función discriminante, para el caso 1.-Cuando las matrices de covarianza son iguales para diferentes clases:*

*Imagen segmentada (en niveles de gris) para el caso 1 con la función discriminante:*

$$d_k(x) = x^T C^{-1} m_k + [\log p(w_k) - \frac{1}{2} m_k^T C^{-1} m_k] \quad k = 1, 2, \dots, M$$

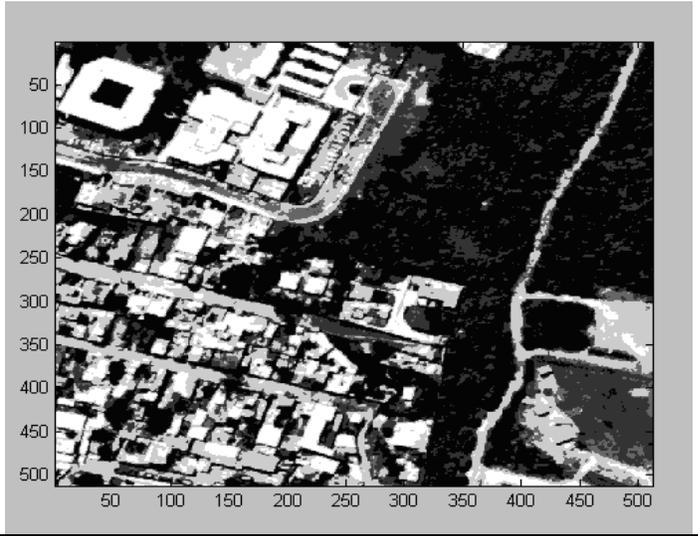
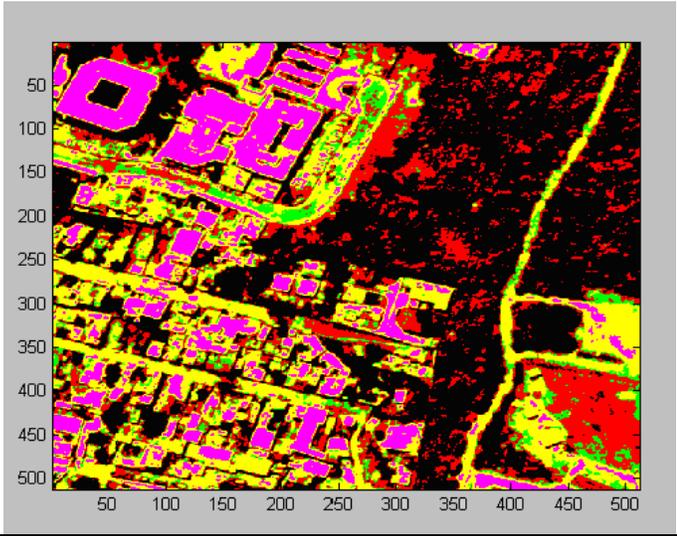


Imagen segmentada (en falso color) con la función discriminante para el caso 2.- Cuando la matriz de covarianza  $C_k$  es de forma diagonal  $\sigma_k^2 I$ , donde

$$\sigma_k^2 = |C_k|:$$

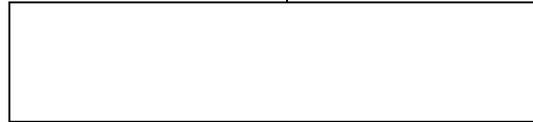
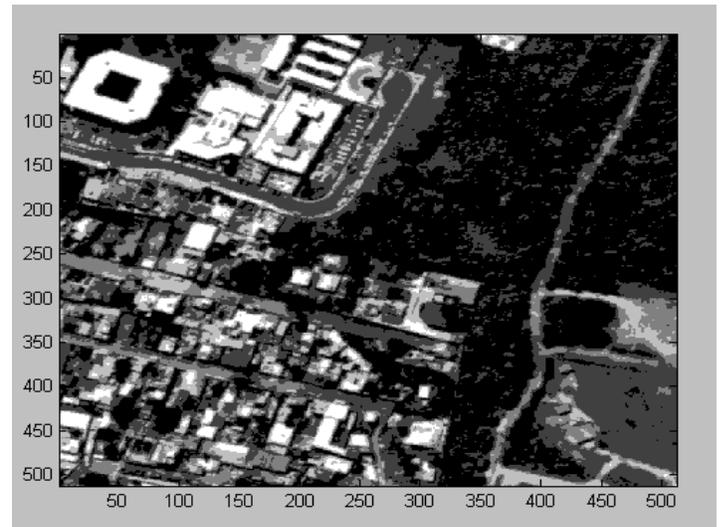
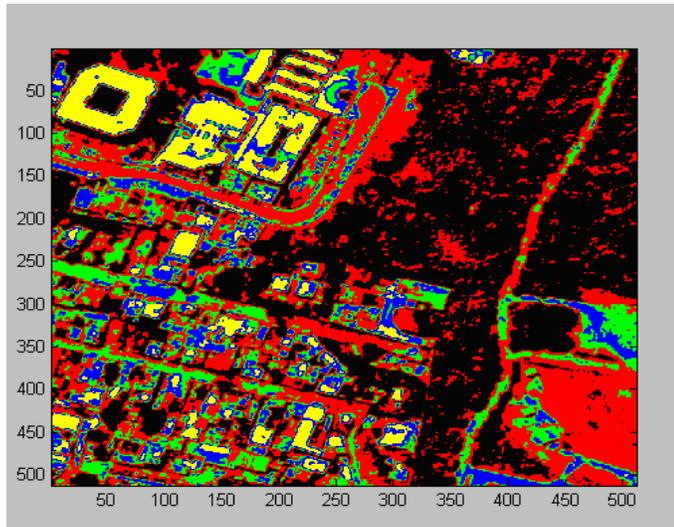


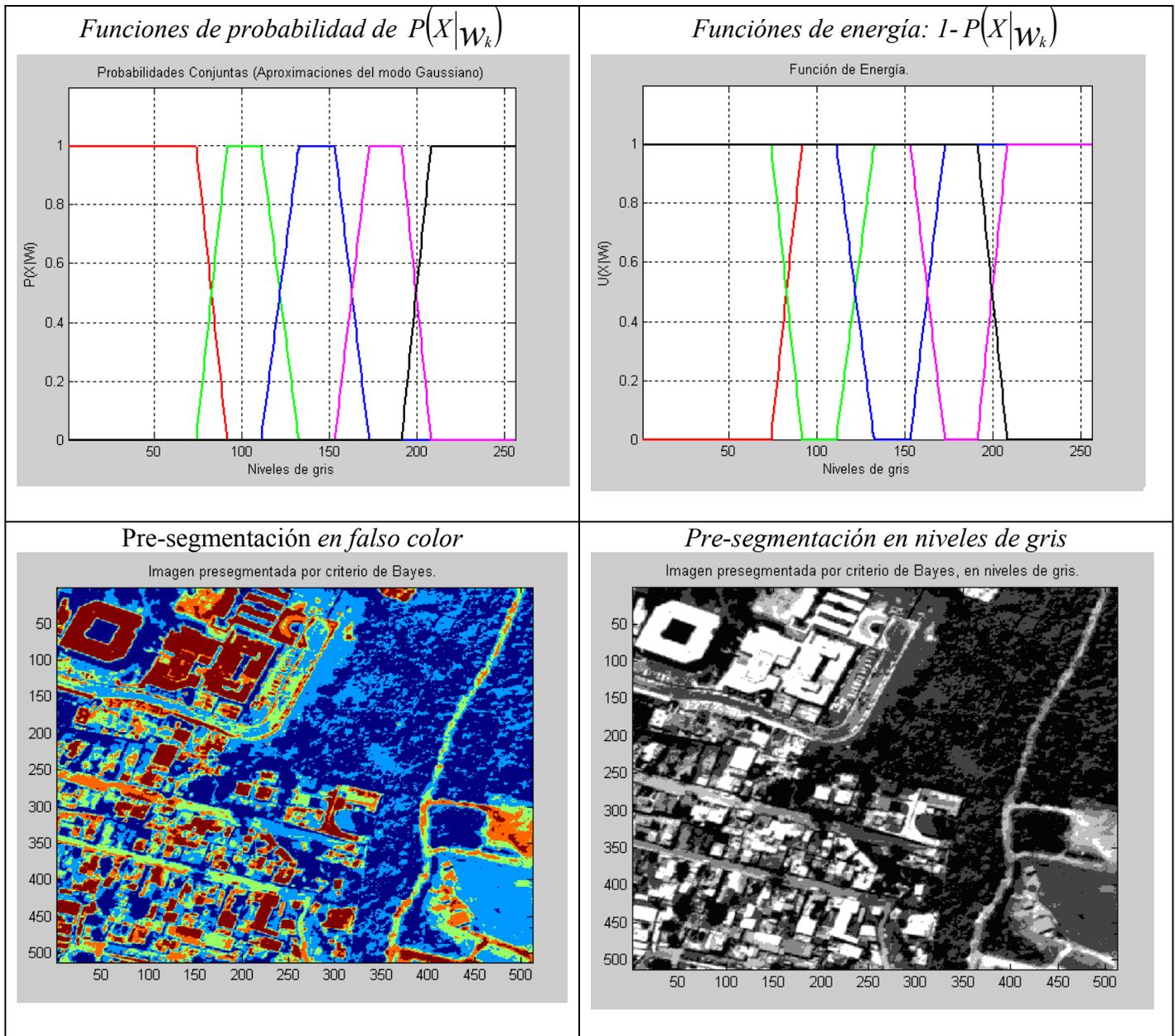
Imagen segmentada (en niveles de gris) para el caso 2 con la función discriminante:



Observando los resultados de la segmentación de la escena 1, por medio de la función discriminante, aunque los tres casos presentan una segmentación aceptable en la clase urbana y algunos errores de menor grado en las clases restantes, se toma la decisión de trabajar con el resultado arrojado por el caso 2, debido a que muestra un mejor agrupamiento de píxeles pertenecientes a la misma clase, dando como resultado una buena uniformidad, originando así una homogeneidad en la segmentación, definiendo de mejor manera la detección de contornos.

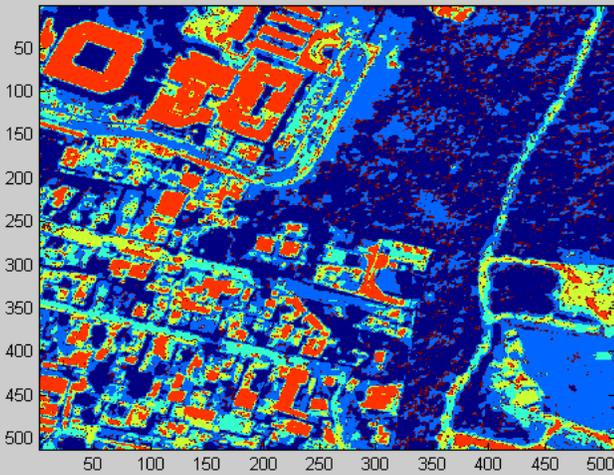
En cambio en el caso general y caso 1, se muestra de una forma mas directa la dispersión de píxeles, originando algunos inconvenientes en la identificación de contornos y regiones homogéneas.

**7.3 Resultados de la segmentación aplicando el criterio de Bayes y los CAM para la imagen 1**



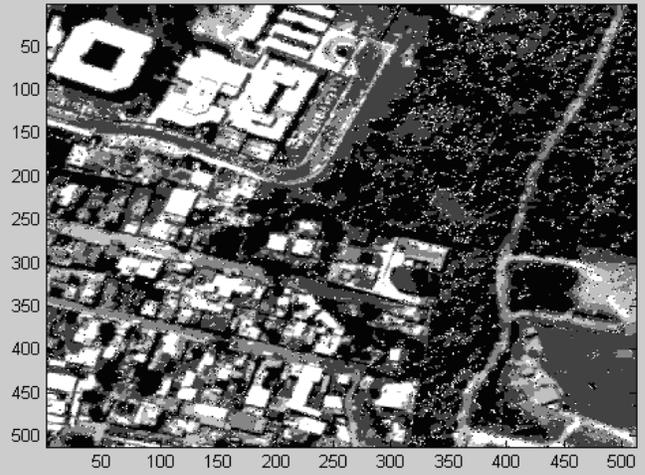
*Segmentación bayesiana  $P(W_k)$  actualizadas.*

Imagen segmentada por criterio de Bayes.



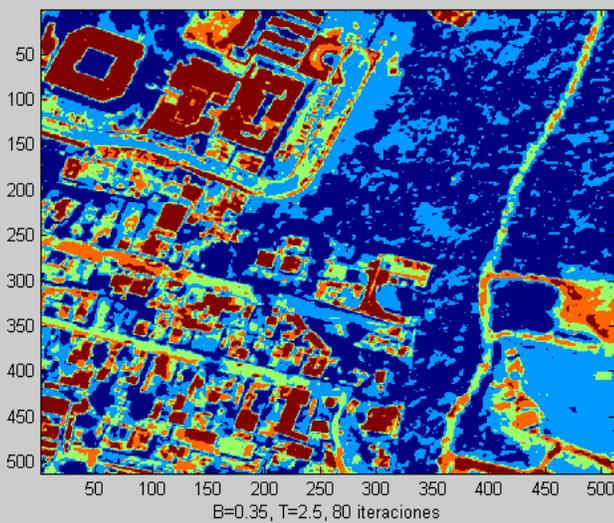
*Segmentación en niveles de gris*

Imagen segmentada por criterio de Bayes, en niveles de gris.



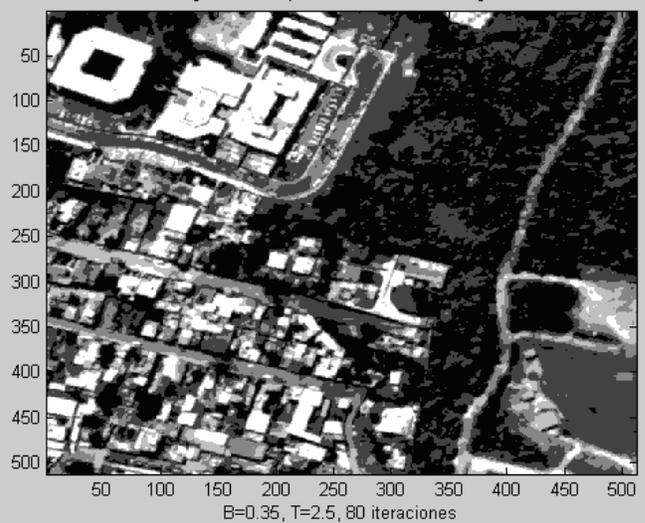
*Segmentación Markoviana*

Segmentación por CAM.



*Segmentación en niveles de gris*

Segmentación por CAM, en niveles de gris.



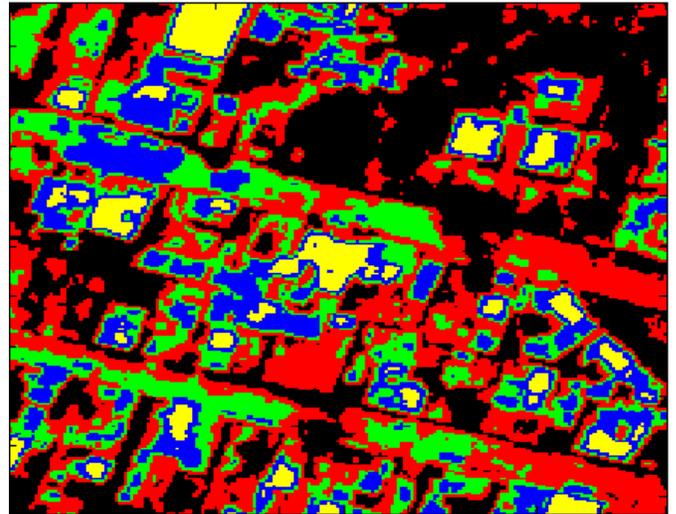
En el siguiente conjunto de esquemas se muestran las diferencias entre una segmentación y otra, comparando por un lado una subimagen de la original zona urbana 1, con cada subimagen segmentada por los distintos métodos.

7.3.1.- Esquema comparativo mediante un acercamiento (zoom) de los métodos de segmentación para la escena 1.

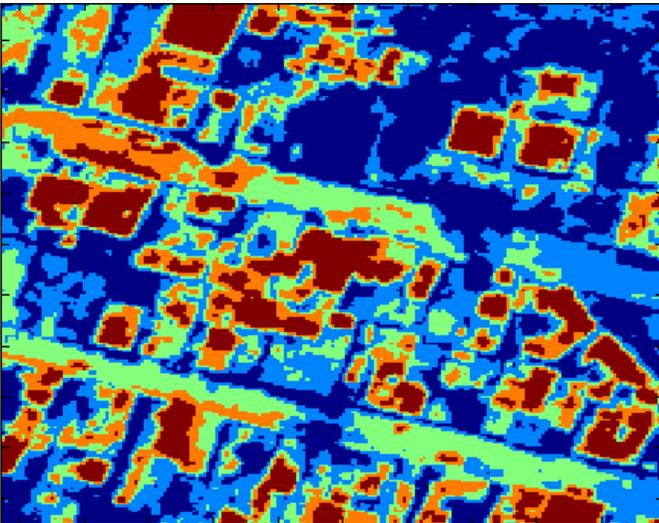
*Imagen original*



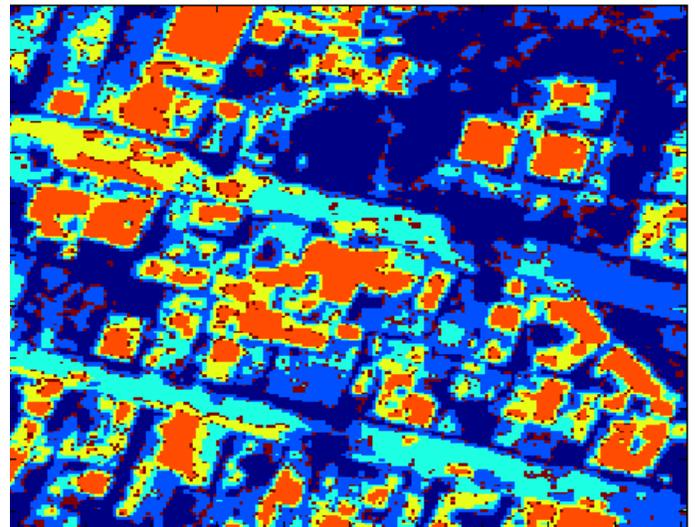
Función discriminante (caso 2).

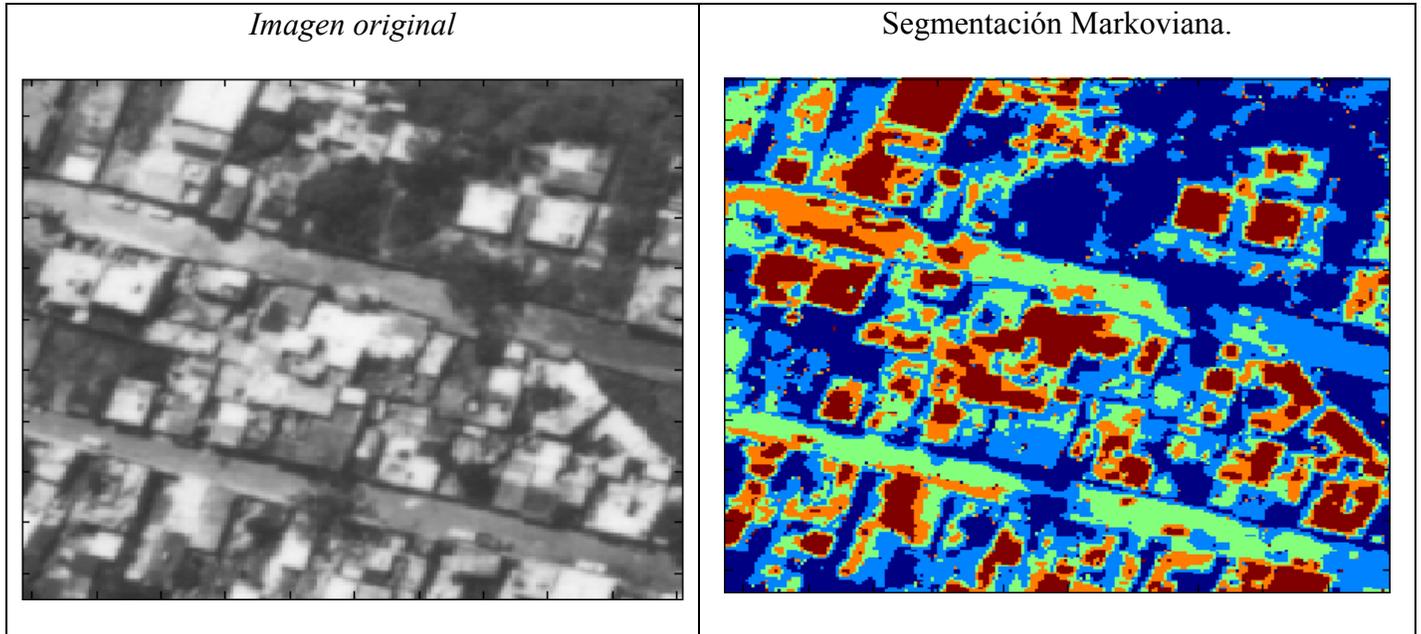


*Pre-segmentación Bayesiana.*



*Segmentación bayesiana  $P(W_k)$  actualizadas .*





*Tabla comparativa en relación al porcentaje de pertenencia de cada clase.*

	$W_1(\%)$	$W_2(\%)$	$W_3(\%)$	$W_4(\%)$	$W_5(\%)$
Función discriminante (caso 2).	33.7181	15.0379	7.4944	37.6884	6.0612
Segmentación Bayesiana.	30.1323	13.7760	11.6543	34.6275	9.8099
Segmentación Markoviana	30.8491	16.5039	10.4115	36.3602	5.8754

**Tabla 8.1. Porcentaje por método de segmentación.**

**7.4 Resultados de la segmentación aplicando la función discriminante de Bayes para la imagen 2**

Imagen segmentada (en falso color) con la función discriminante, caso general:

Imagen segmentada (en niveles de gris) con la función discriminante, caso general:

$$d_k^{iii}(x) = -\frac{1}{2}(x-m_k)^T C_k^{-1}(x-m_k) + [\log p(w_k) - \frac{1}{2} \log |C_k|] \quad k = 1,2$$

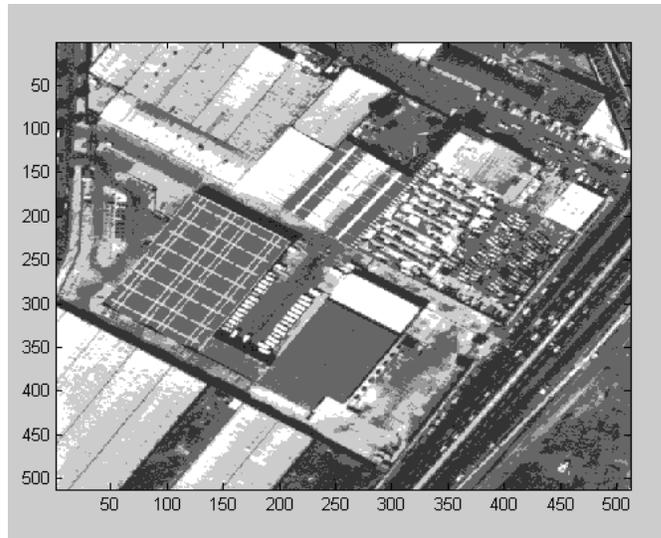
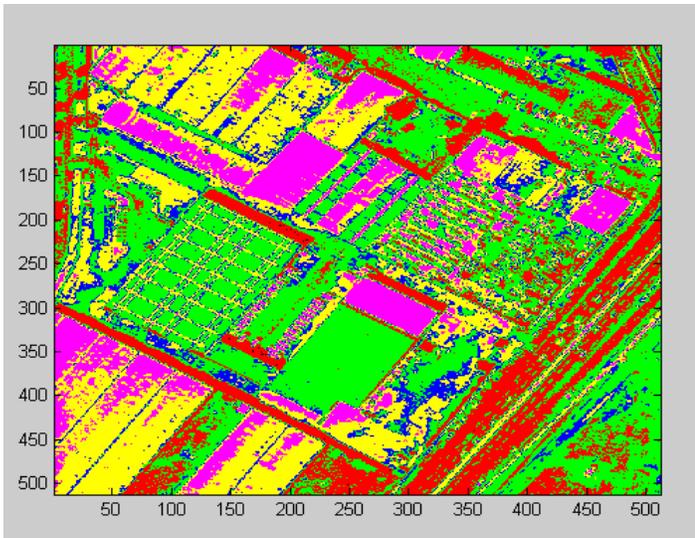


Imagen segmentada con la función discriminante en falso color, para el caso 1.-Cuando las matrices de covarianza son iguales para diferentes clases:

Imagen segmentada para el caso 1 con la función discriminante en niveles de gris:

$$d_k(x) = x^T C^{-1} m_k + [\log p(w_k) - \frac{1}{2} m_k^T C^{-1} m_k] \quad k = 1,2,...,M$$

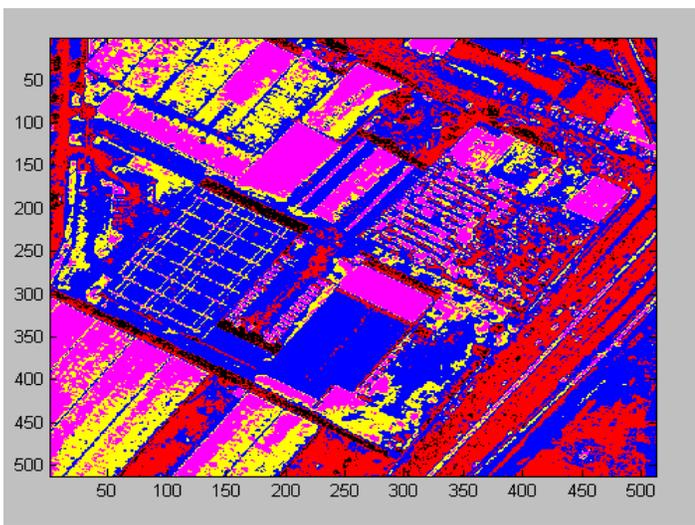
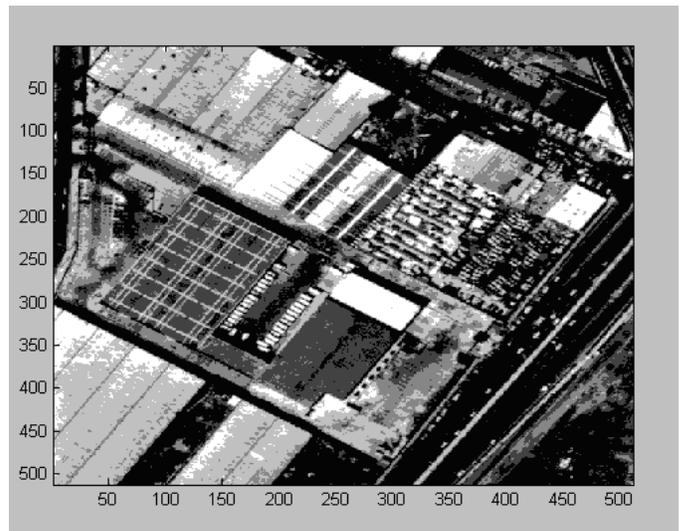
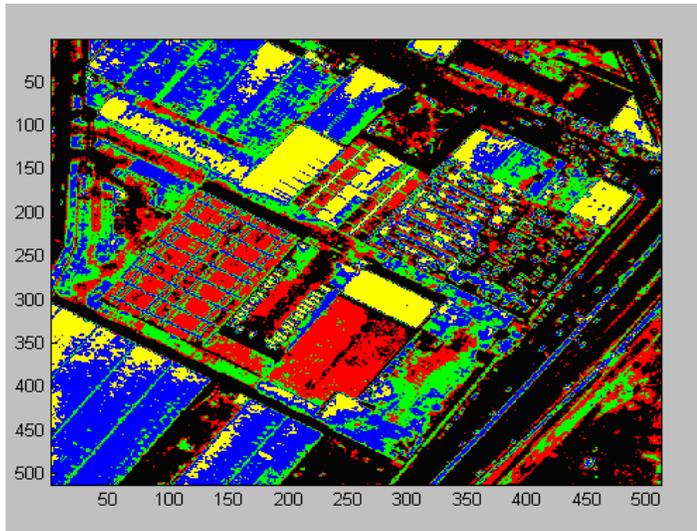
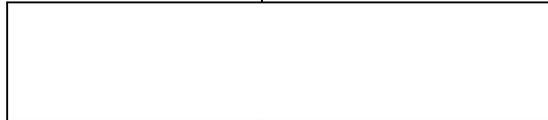


Imagen segmentada (en falso color) con la función discriminante, para el caso 2.- Cuando la matriz de covarianza  $C_k$  es de forma diagonal  $\sigma_k^2 I$ , donde

$$\sigma_k^2 = |C_k|:$$

Imagen segmentada (en niveles de gris) para el caso 2 con la función discriminante:

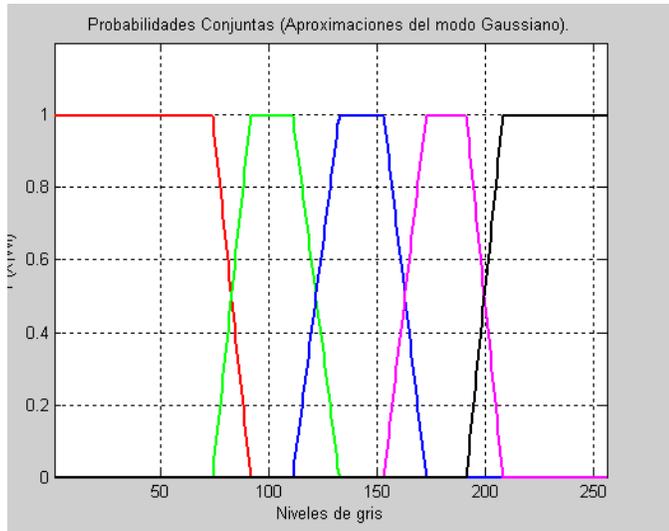


Observando los resultados de la segmentación de la escena 2, por medio de la función discriminante, se puede observar algunos errores de segmentación en determinadas zonas, tanto en el caso general como en el caso 1. Esto debido a la dispersión considerable de píxeles, originando cierta confusión de clases principalmente en áreas verdes, presentándose una definición de contornos errónea.

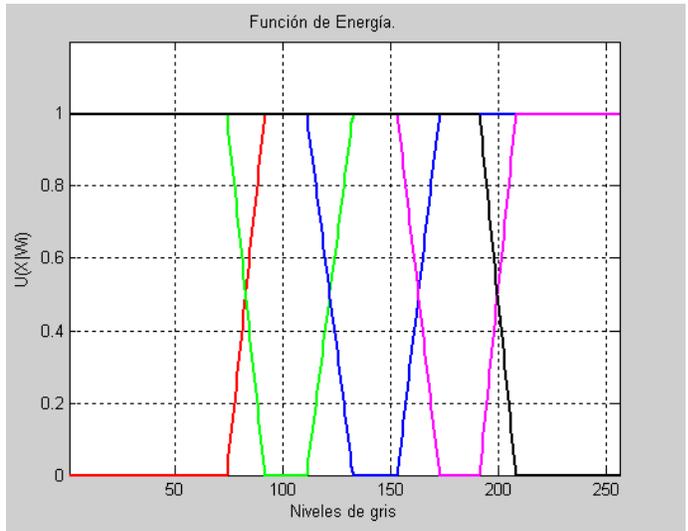
Se toma la decisión de trabajar de nuevo con el resultado arrojado por el caso 2, debido a que logra definir de mejor manera la frontera existente entre las clases y por consiguiente la detección de contornos es más acertada. de modo que origina un etiquetado más acertado que en los dos casos anteriores.

**7.5 Resultados de la segmentación aplicando el criterio de Bayes y los CAM para la imagen 2**

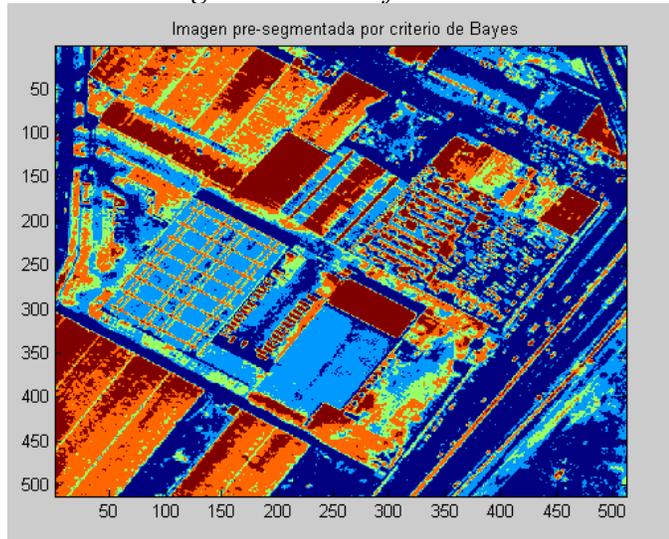
*Funciones de probabilidad de  $P(X|w_k)$ .*



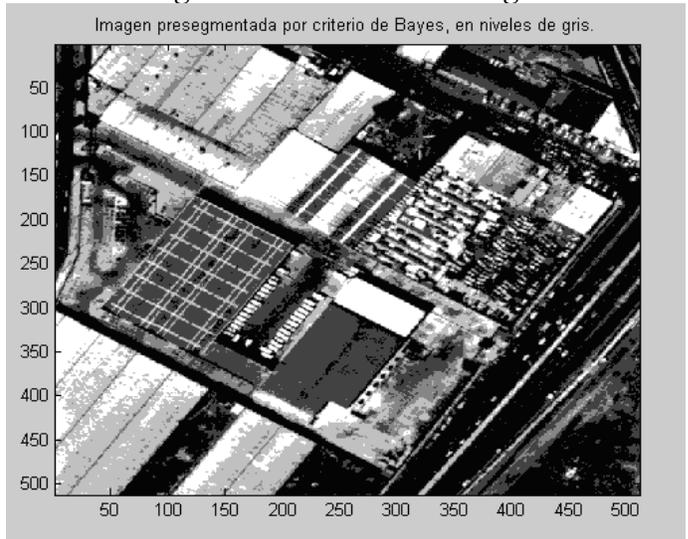
*Funciones de energía  $U(X|w_k)$ .*

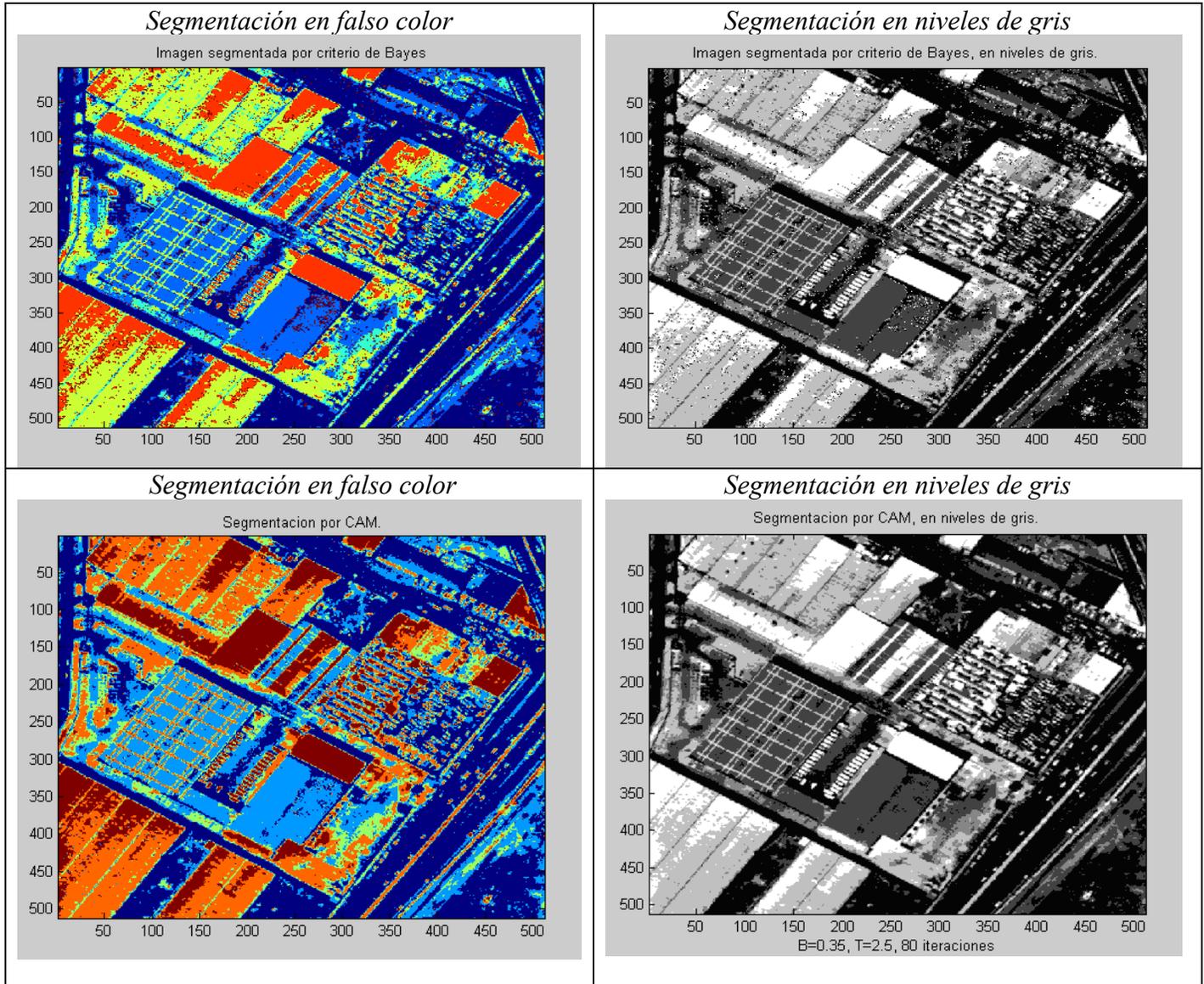


*Segmentación en falso color*



*Segmentación en niveles de gris*





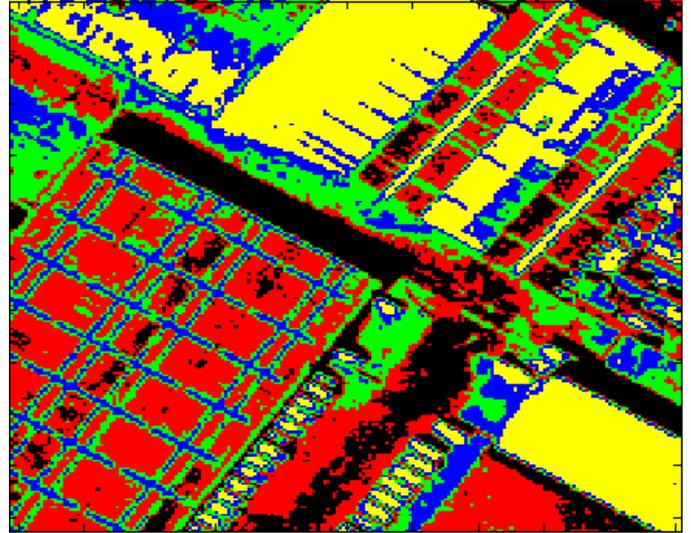
En el siguiente conjunto de esquemas se muestran las diferencias entre una segmentación y otra, comparando por un lado una subimagen de la original DF1 con cada subimagen segmentada por los distintos métodos.

7.5.1.- Esquema comparativo mediante un acercamiento (zoom) de los métodos de segmentación para la escena 2.

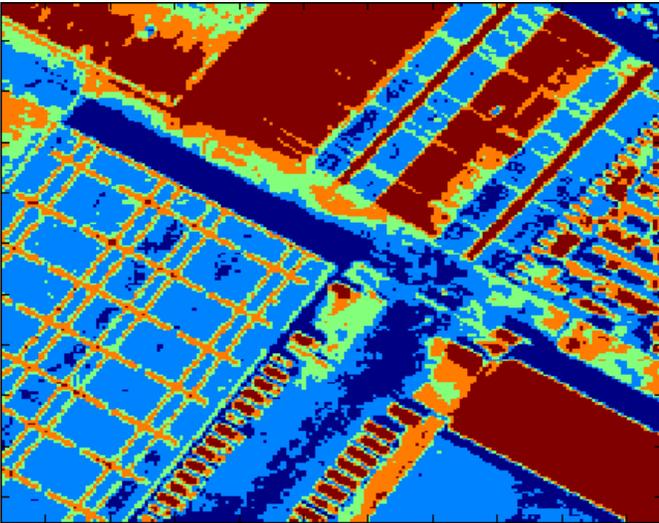
*Imagen original*



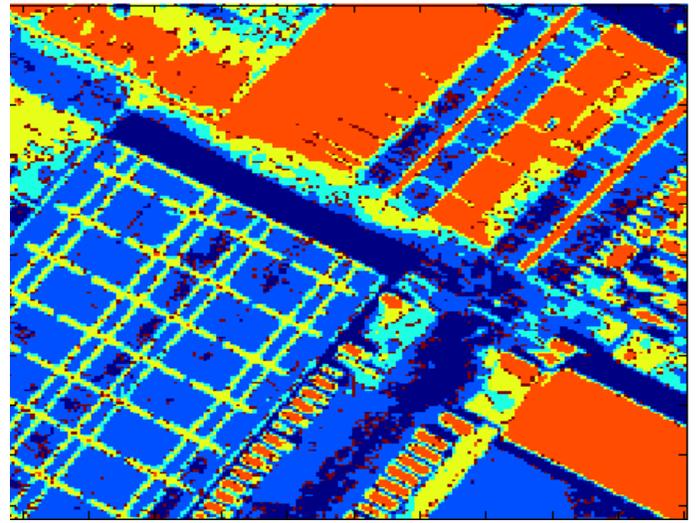
Función discriminante (caso 2).

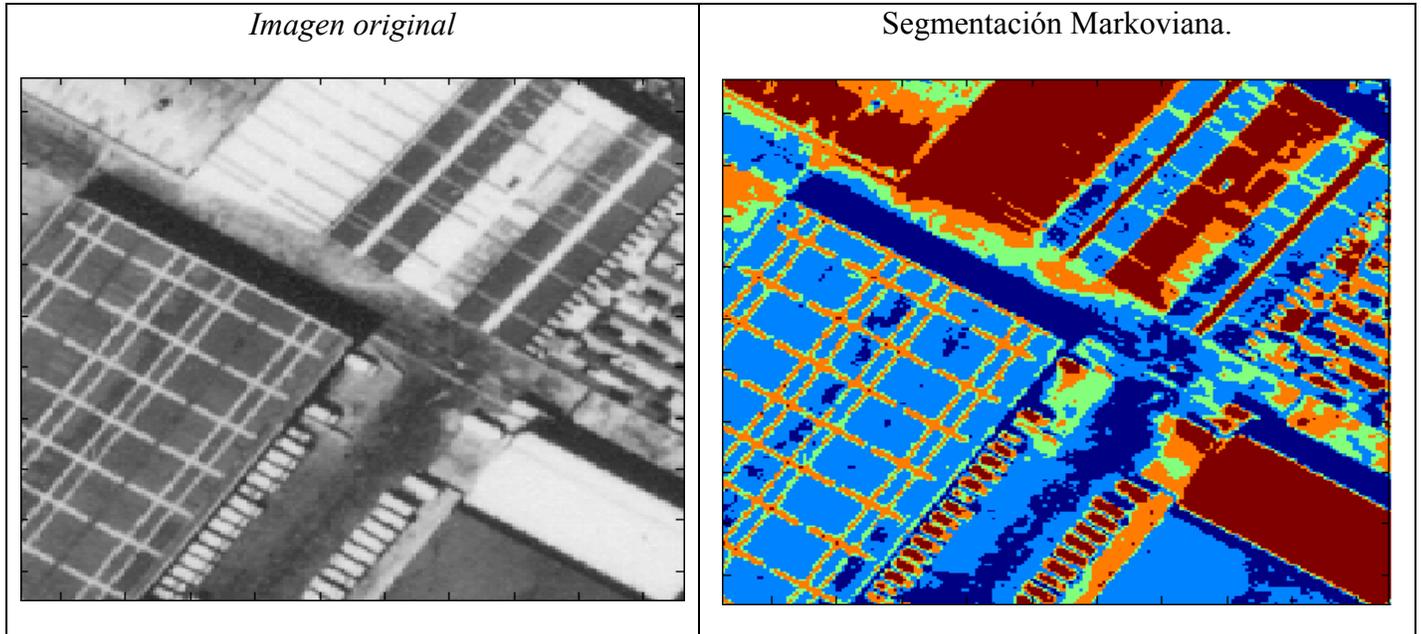


*Pre-segmentación Bayesiana.*



*Segmentación bayesiana  $P(W_k)$  actualizadas.*





*Tabla comparativa en relación al porcentaje de pertenencia de cada clase.*

	$W_1(\%)$	$W_2(\%)$	$W_3(\%)$	$W_4(\%)$	$W_5(\%)$
Función discriminante (caso 2).	9.0786	28.4714	12.3131	22.2904	27.8465
Segmentación Bayesiana	12.6625	30.7468	13.2748	19.8254	23.4905
Segmentación Markoviana	10.1147	27.3743	15.3515	22.8321	24.3275

**Tabla 7.2.- Porcentaje por método de segmentación.**

En el desarrollo de la implementación de esta teoría de decisión, se tomaron ventanas de entrenamiento que ayudaron a escoger las clases más significativas de la imagen original, en donde en cada una de estas ventanas se calcularon los parámetros estadísticos que fueron: la media, la varianza, la covarianza, la desviación estándar, la distancia de Mahalanobis, y obviamente las funciones discriminantes que representan a la probabilidad a posteriori  $p(x|\mathcal{W}_k)$ .

La segmentación por medio de la función discriminante óptima, es aceptable tanto en el caso general, caso 1, así como en el caso 2 en donde se manejan probabilidades a priori que son equiprobables, derivando un mejor resultado.

Como una muestra de comparación se usó el algoritmo de la regla de Bayes, en la cual se utilizó una pre-segmentación con las funciones bayesianas de segmentación mencionadas dentro del capítulo 4, para actualizar las probabilidades y posteriormente llevar a cabo una segmentación, con las mismas funciones. Este algoritmo tuvo un tiempo aproximado de 23.32 minutos.

Haciendo una comparación de segmentaciones, se puede constatar que la obtenida a través de la función discriminante óptima de Bayes arroja un resultado bastante alentador, con la mejora en relación al costo computacional, ya que el resultado se tiene de manera casi inmediata. Con lo anterior mencionado se puede decir que se ha logrado asentar las bases para hacer una buena clasificación de la imagen, con la previa segmentación obtenida en el método de la función discriminante óptima, ya sea en caso general, caso 1 ó caso 2, ejemplificándose esto en los resultados al implementar el algoritmo de la segmentación bayesiana

La segmentación por CAM es un resultado aceptable, sin quitar mérito al obtenido por medio de la función discriminante óptima de Bayes. Pero surge un gran inconveniente con este método, el tiempo de ejecución, , no obstante, cabe mencionar que se logró fuera de el menor tiempo posible a la hora de programarlo, haciendo una programación sencilla pero detallada, con esto se obtuvo un tiempo considerablemente aceptable, que fue de 34.25 minutos en un (procesador Pentium IV) . Además de considerar que el uso de los recursos de la computadora son limitados a la hora de esperar el resultado, debido a la manera en como se va realizando el número de iteraciones. Con esto se puede inferir que la mejor opción es la segmentación a través de la función discriminante óptima de Bayes, específicamente el caso 2, ya que se observa una mejor definición de los contornos de la imagen y así mismo un etiquetado más eficiente, que permite distinguir cada una de las regiones que componen la imagen original.

En este capítulo se presentaron los resultados de los diferentes métodos de segmentación propuestos, así como de los métodos de relajación con los cuales se busca obtener mejores resultados utilizando la segmentación, para la regularización de la imagen.

El mejor resultado dentro de los métodos de segmentación fue el de la aplicación de la función discriminante de Bayes, específicamente para el caso 2.

Tomando como base este resultado, se procede entonces a analizar los tres métodos de relajación descritos en el capítulo 6, en el orden siguiente: relajación probabilística, relajación determinística y por último la relajación estocástica.

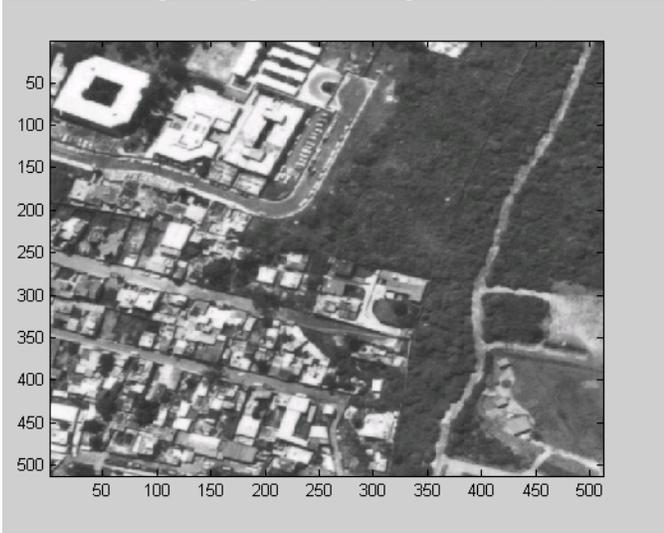
### ***7.6 Resultados del proceso de relajación***

Para el caso de los resultados obtenidos por los procesos de relajación, se asigna cierto número de iteraciones de acuerdo al funcionamiento de cada uno de los métodos. Posteriormente se hace un análisis de la limpieza morfológica, para así determinar el grado de píxeles aislados que se logró se fusionarán en base al contexto al cual pertenecen a lo largo del proceso. Las imágenes de limpieza morfológica mostradas, se obtienen mediante la diferencia entre la imagen segmentada y un resultado de relajación, conforme se va iterando.

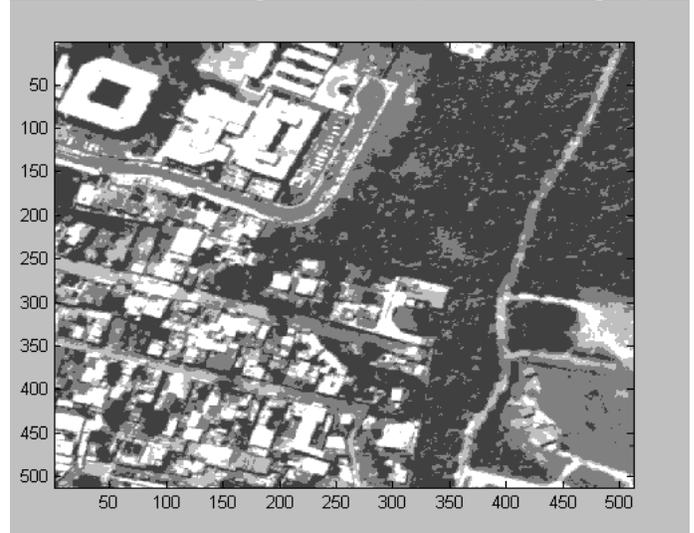
*7.6.1 Relajación probabilística para la escena urbana 1*

Se muestra a continuación el resultado de la aplicación del método de relajación probabilística aplicado en la imagen segmentada de la escena urbana 1, por medio de funciones discriminantes de Bayes, propiamente con el caso 2. Las imágenes de limpieza morfológica mostradas, se obtienen mediante la diferencia entre la imagen segmentada y un resultado de relajación, conforme se va iterando.

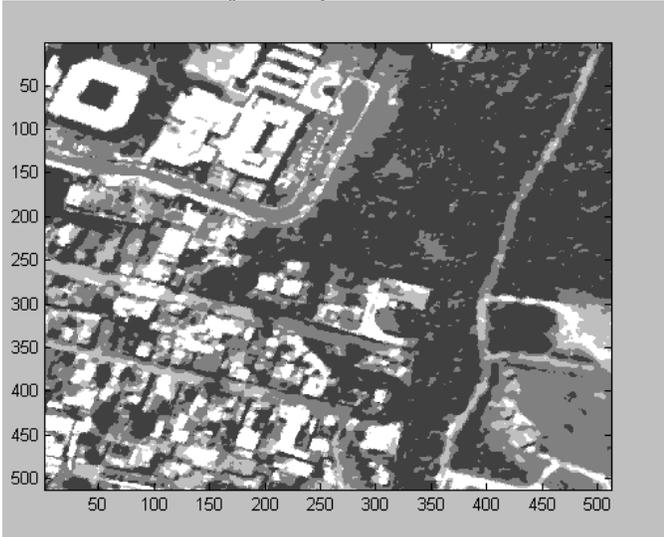
*Imagen original (sin segmentación)*



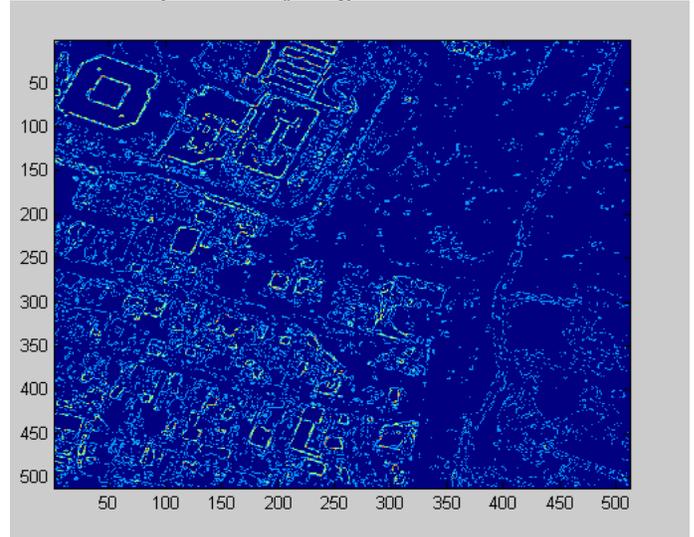
*Resultado de segmentación (en niveles de gris)*



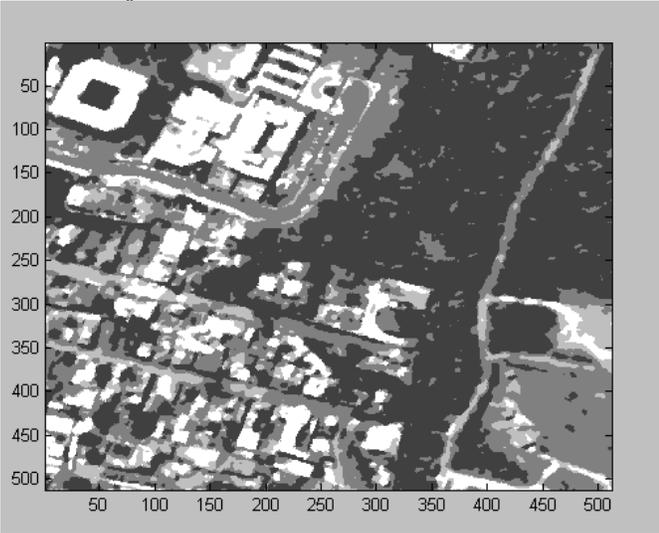
*Resultado de relajación probabilística: 1 iteración*



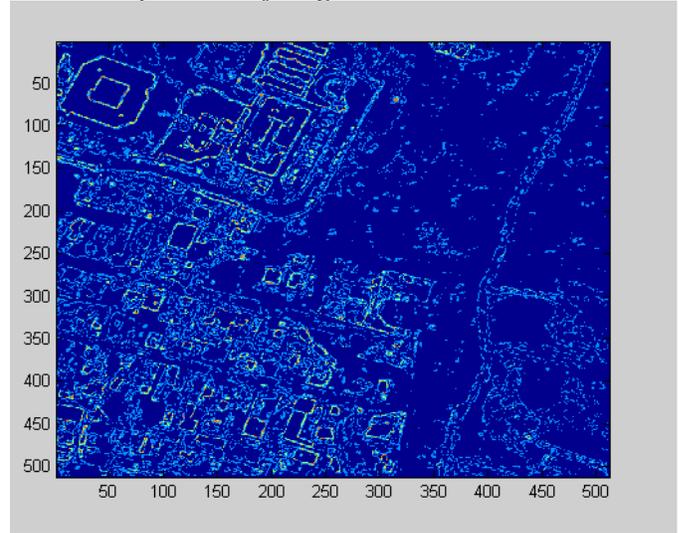
*Limpieza morfológica con 1 iteración*



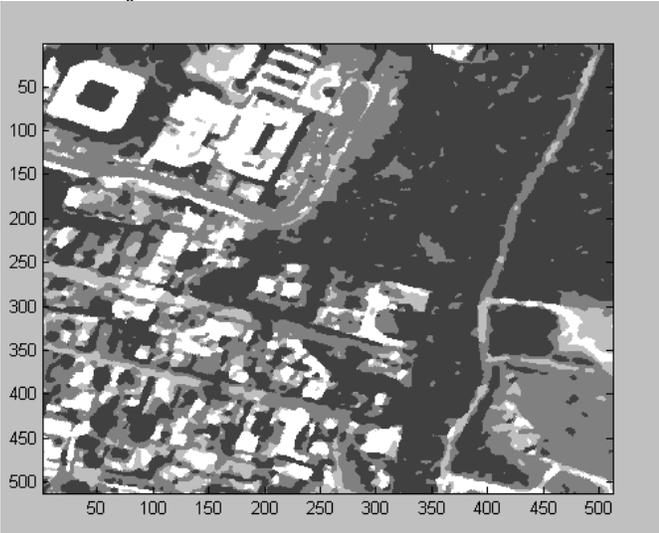
*Relajación Probabilística 2 iteraciones.*



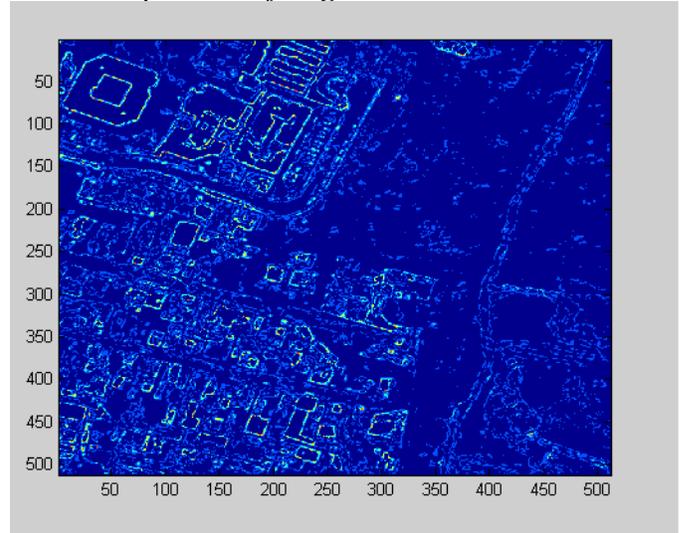
*Limpieza morfológica con 2 iteraciones..*



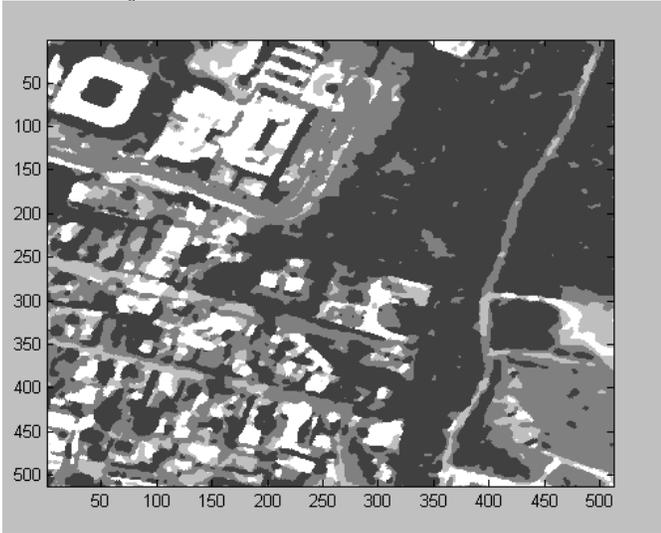
*Relajación Probabilística 3 iteraciones.*



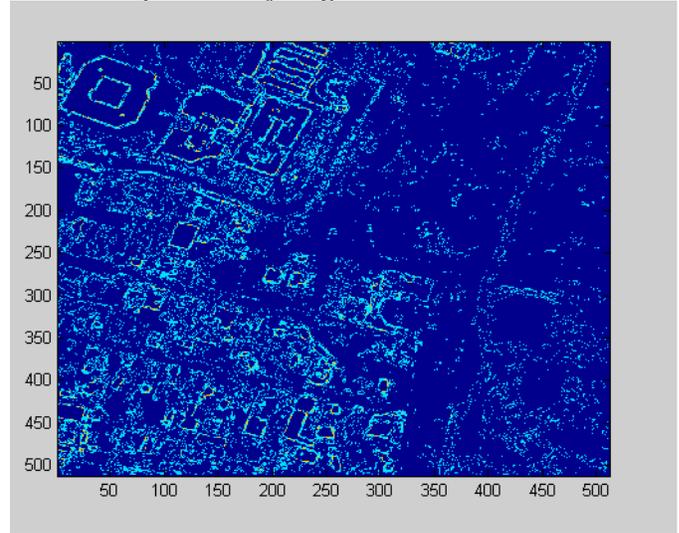
*Limpieza morfológica con 3 iteraciones.*



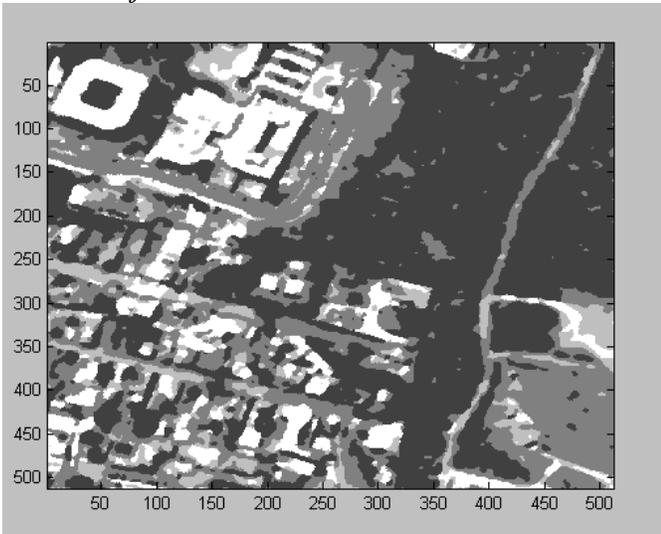
*Relajación Probabilística 4 iteraciones.*



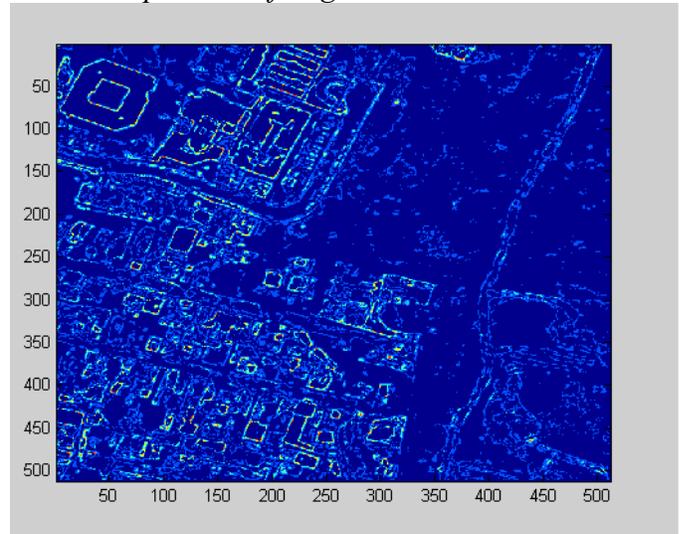
*Limpieza morfológica con 4 iteraciones.*



*Relajación Probabilística 5 iteraciones.*



*Limpieza morfológica con 5 iteraciones.*



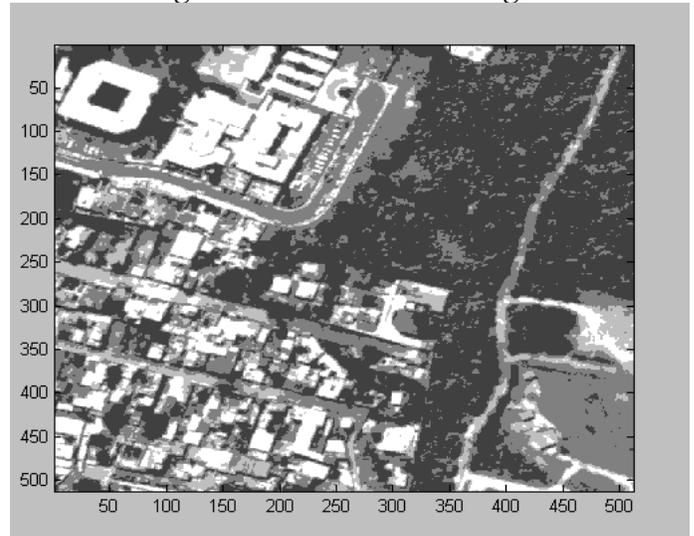
7.6.2 Relajación determinística para la escena urbana 1

Se muestra a continuación el resultado de la aplicación del método de relajación determinística aplicado en la imagen segmentada de la escena urbana 1, por medio de funciones discriminantes de Bayes, propiamente con el caso 2. Las imágenes de limpieza morfológica mostradas, se obtienen mediante la diferencia entre la imagen segmentada y un resultado de relajación, conforme se va iterando.

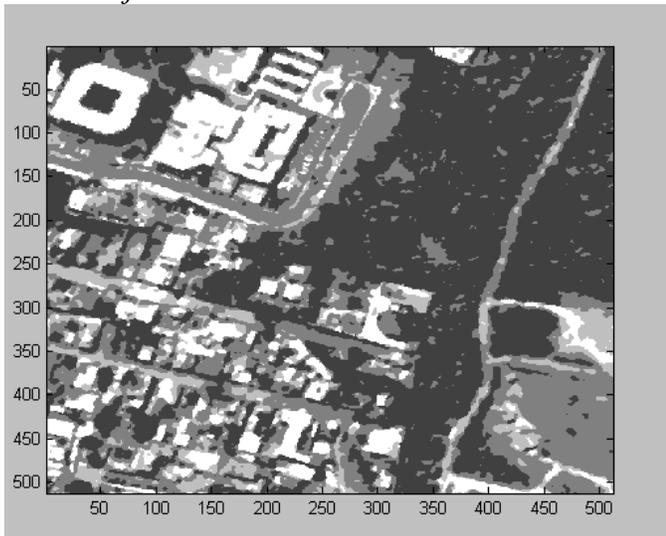
*Imagen original sin segmentación.*



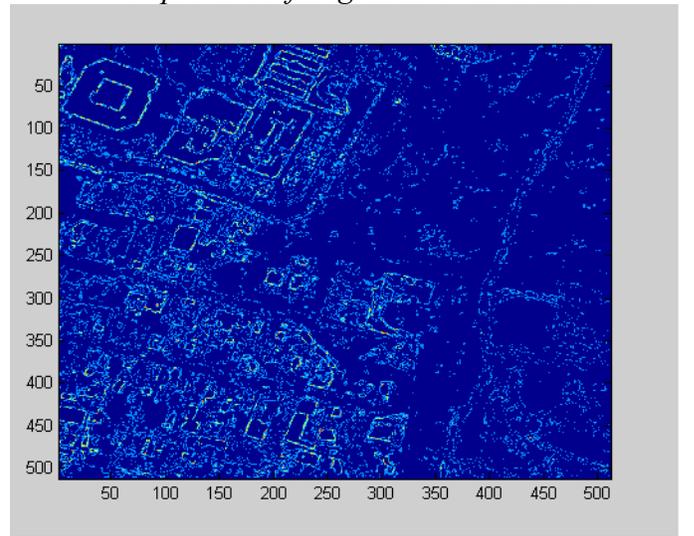
*Segmentación en niveles de gris.*



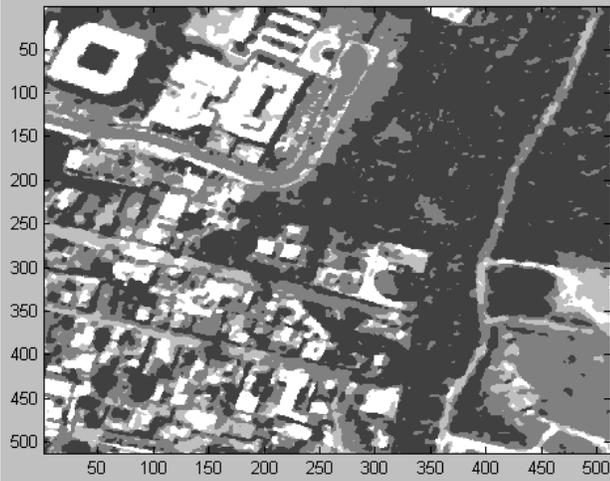
*Relajación determinística con 1 iteración.*



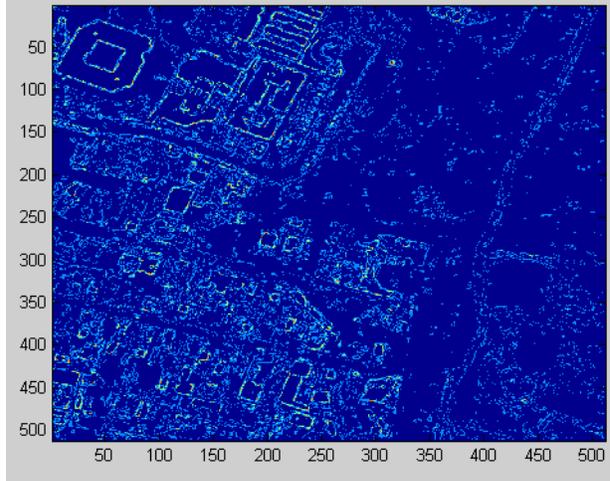
*Limpieza morfológica con iteración.*



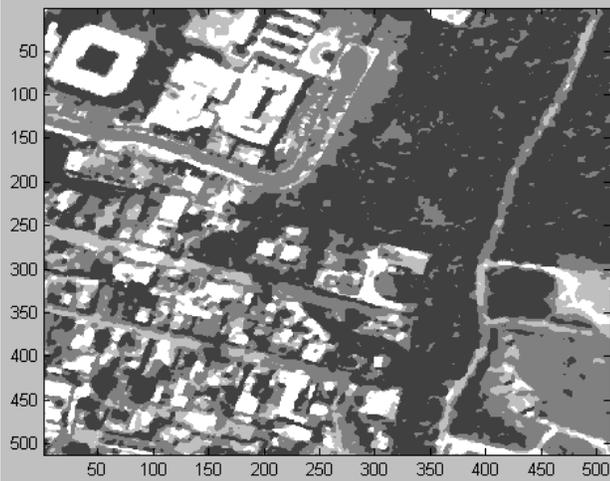
*Relajación determinística con 2 iteraciones.*



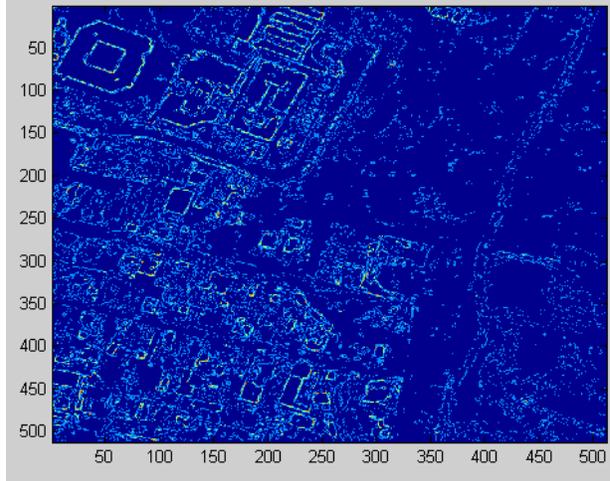
*Limpieza morfológica con 2 iteraciones.*



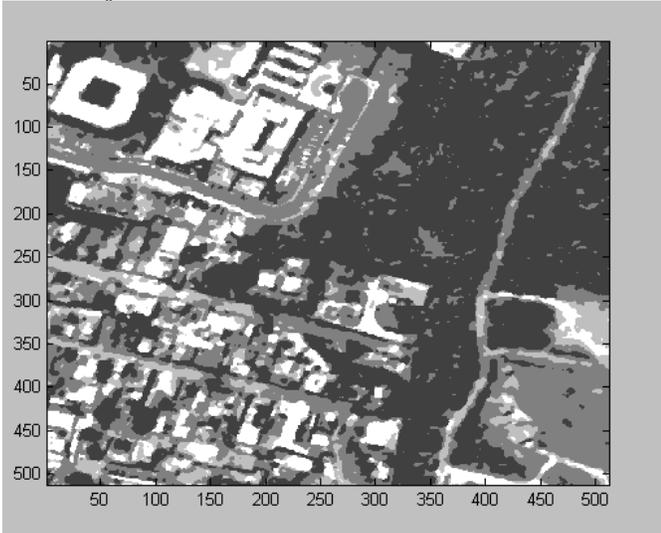
*Relajación determinística 3 iteraciones.*



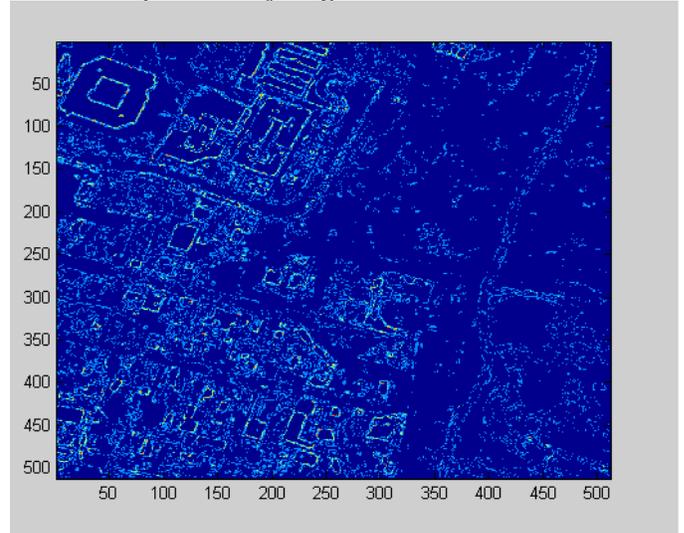
*Limpieza morfológica con 3 iteraciones.*



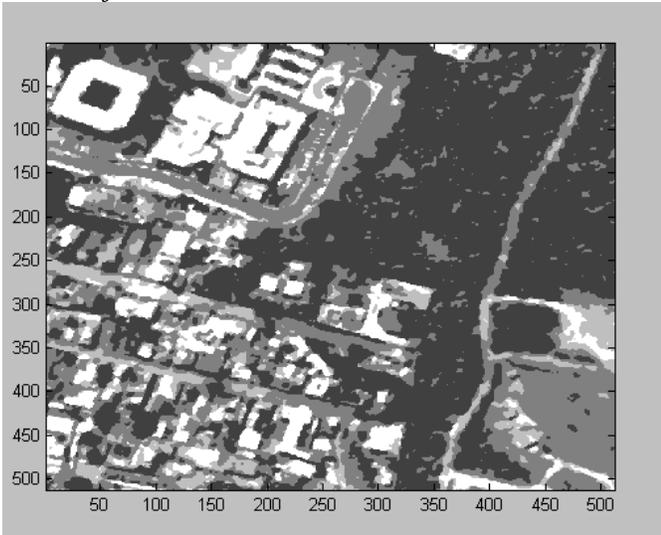
*Relajación determinística con 4 iteraciones.*



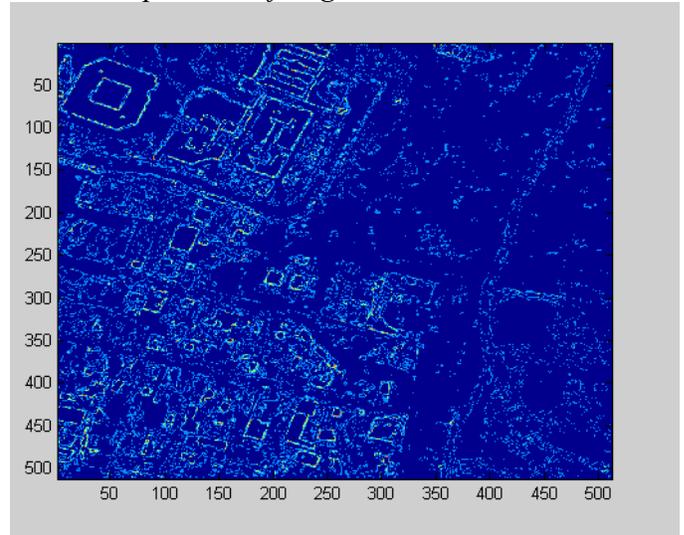
*Limpieza morfológica con 4 iteraciones.*



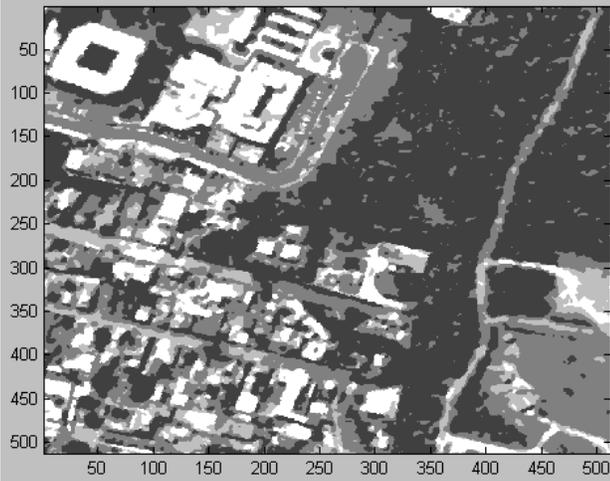
*Relajación determinística con 5 iteraciones.*



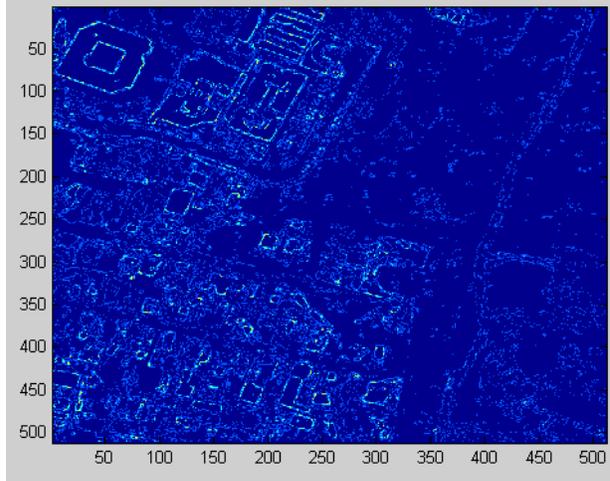
*Limpieza morfológica con 5 iteraciones.*



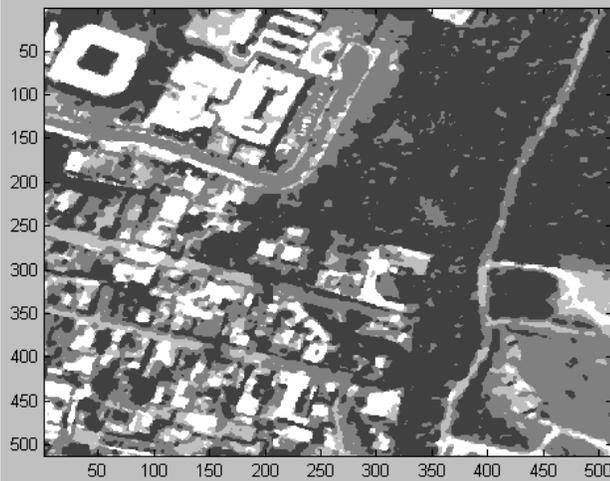
*Relajación determinística con 6 iteraciones.*



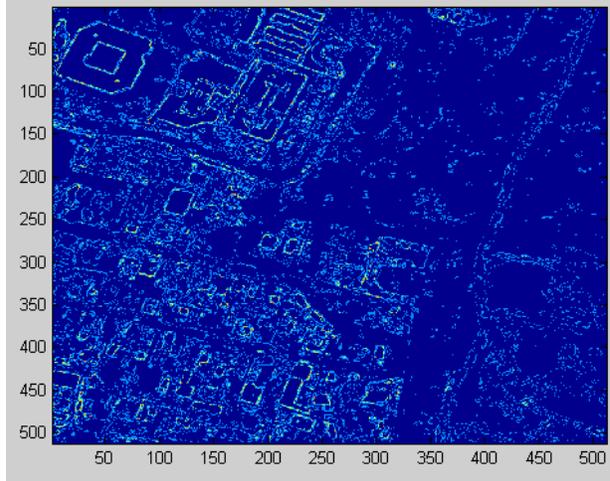
*Limpieza morfológica con 6 iteraciones.*



*Relajación determinística con 7 iteraciones.*

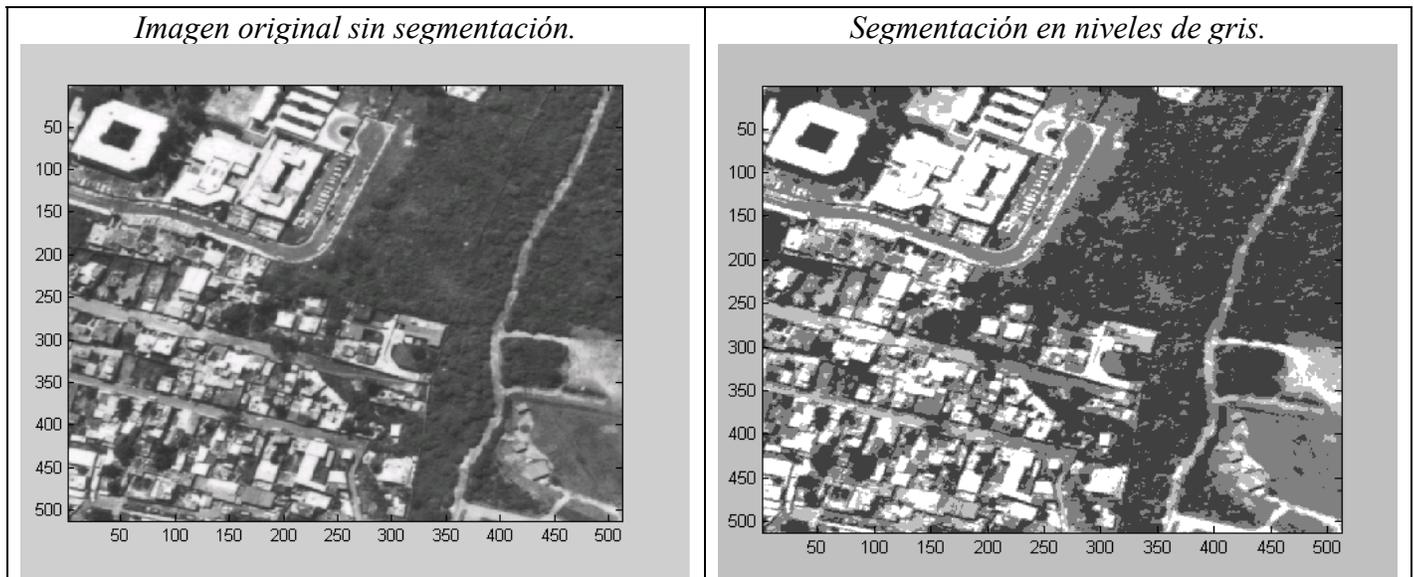


*Limpieza morfológica con 7 iteraciones.*



### 7.6.3 Relajación estocástica para la escena urbana 1, con diferentes configuraciones de vecindad y 40 iteraciones cada una

Se muestra a continuación el resultado de la aplicación del método de relajación estocástica aplicado en la imagen segmentada por medio de funciones discriminantes de Bayes, propiamente con el caso 2.



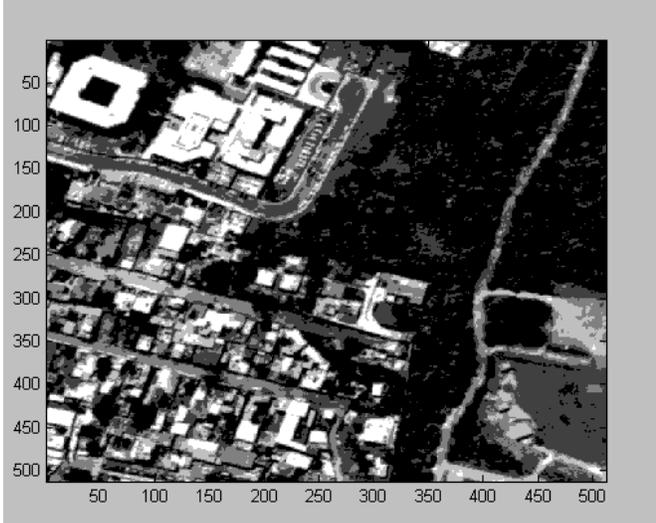
En este caso, se aplica un análisis con vecindades adaptivas a efectos de detectar en mejor forma estructuras finas. De acuerdo a lo mencionado en el capítulo 6, se procede al análisis de diferentes propuestas de conjuntos de configuración de vecindad, en base el conjunto integrado por 8-conexidad.

Las imágenes de limpieza morfológica mostradas para este método de relajación, se obtienen mediante la diferencia entre la imagen segmentada y un resultado de relajación en determinado número de iteraciones.

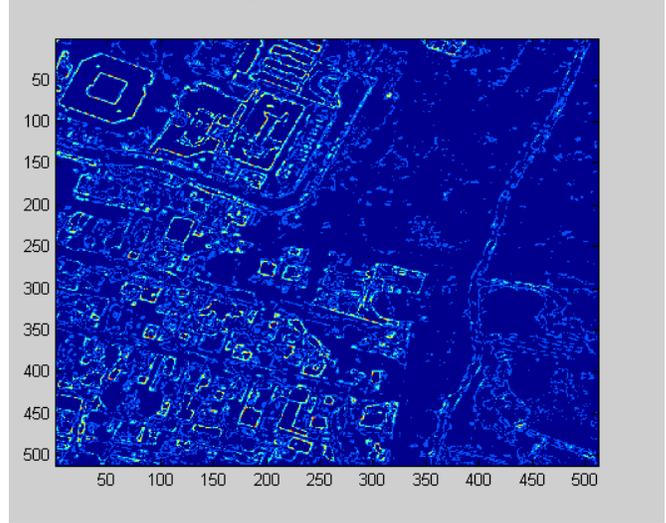
**Configuración 1:**

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i=j}$	
-------------------------------------------------	-------------------------------------------------------------------------------------

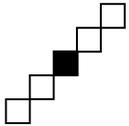
*Relajación estocástica 40 iteraciones.*



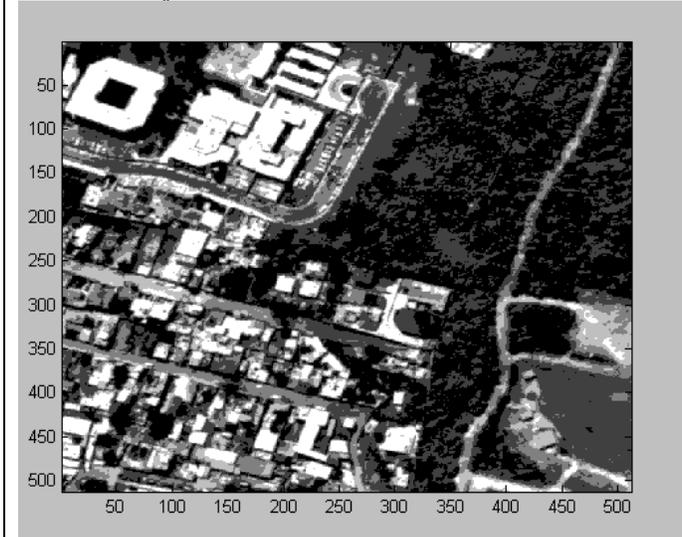
*Limpieza morfológica.*



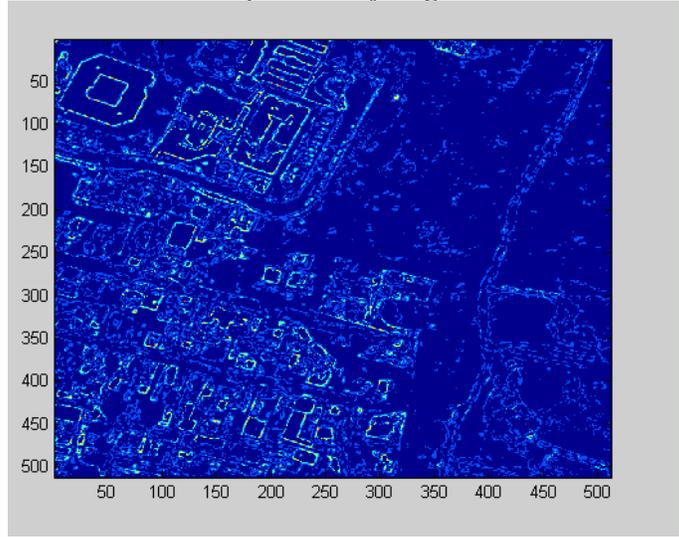
**Configuración 2:**

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{-i=j}$	
--------------------------------------------------	--------------------------------------------------------------------------------------

*Relajación estocástica 40 iteraciones.*

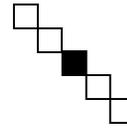


*Limpieza morfológica.*

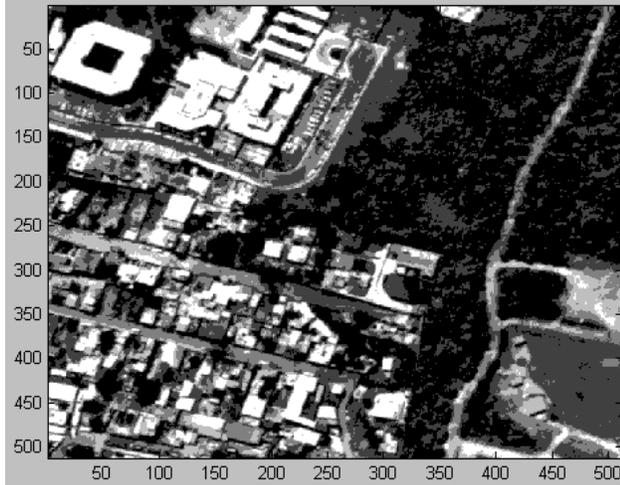


Configuración 3:

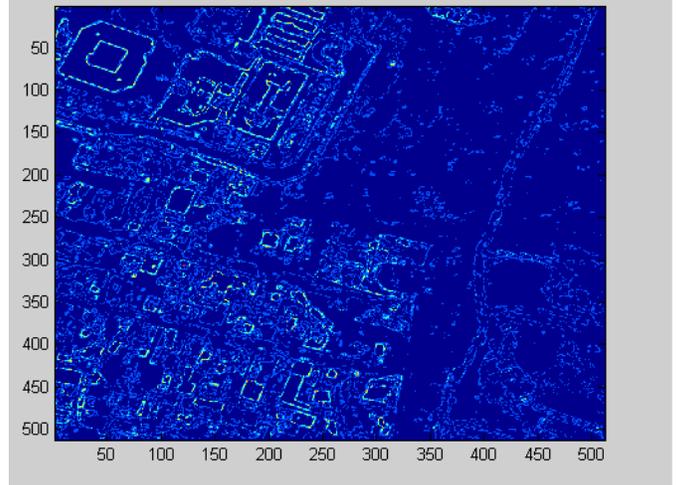
$$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i \rightarrow j}$$



Relajación estocástica 40 iteraciones.

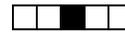


Limpieza morfológica.

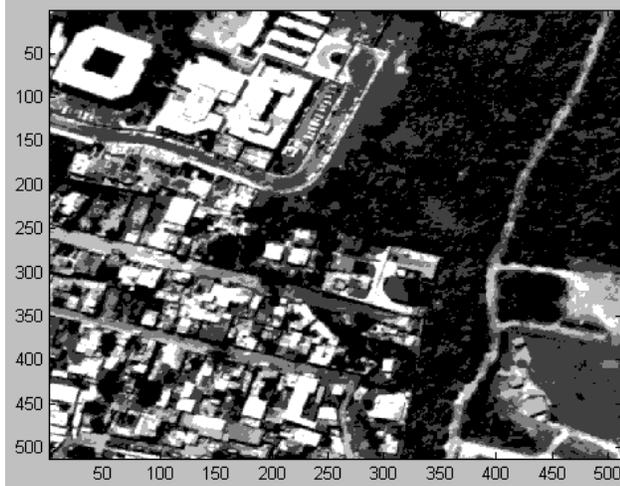


Configuración 4:

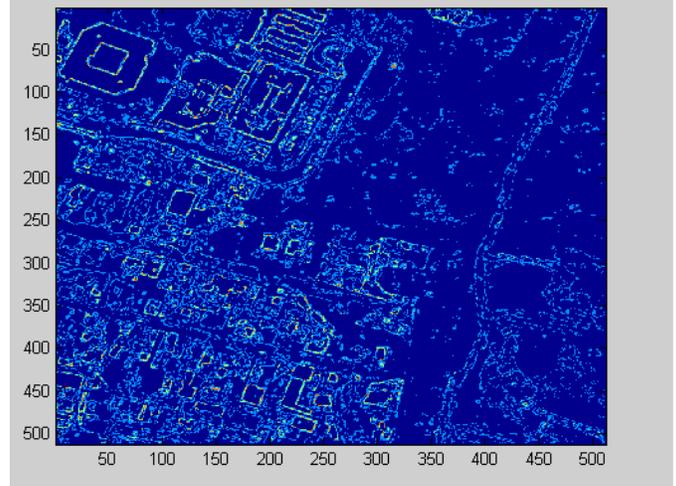
$$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{0j}$$



Relajación estocástica 40 iteraciones.



Limpieza morfológica.

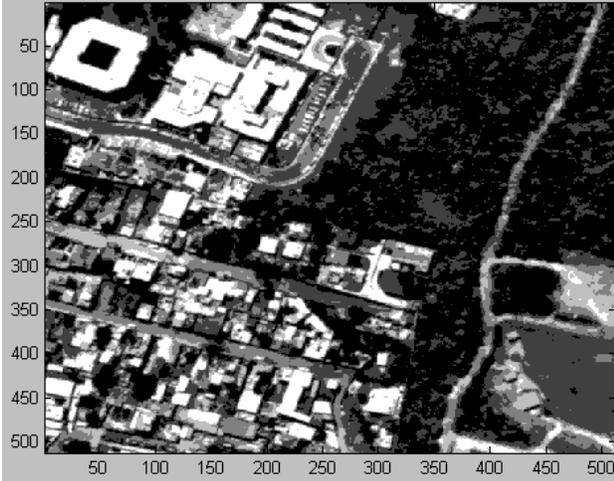


Configuración 5:

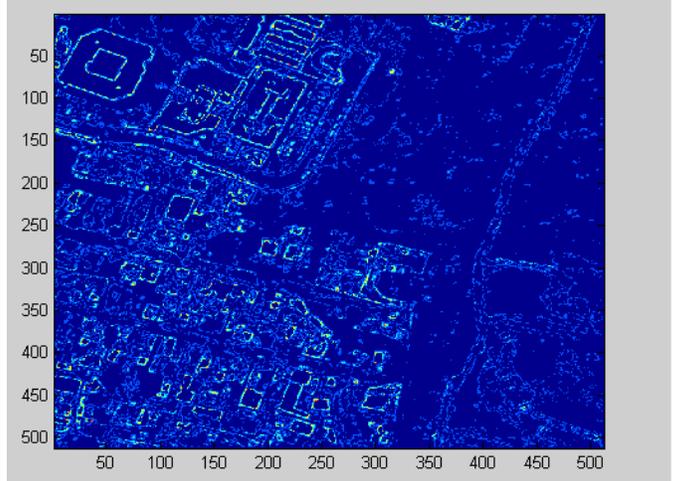
$$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i0}$$



*Relajación estocástica 40 iteraciones.*

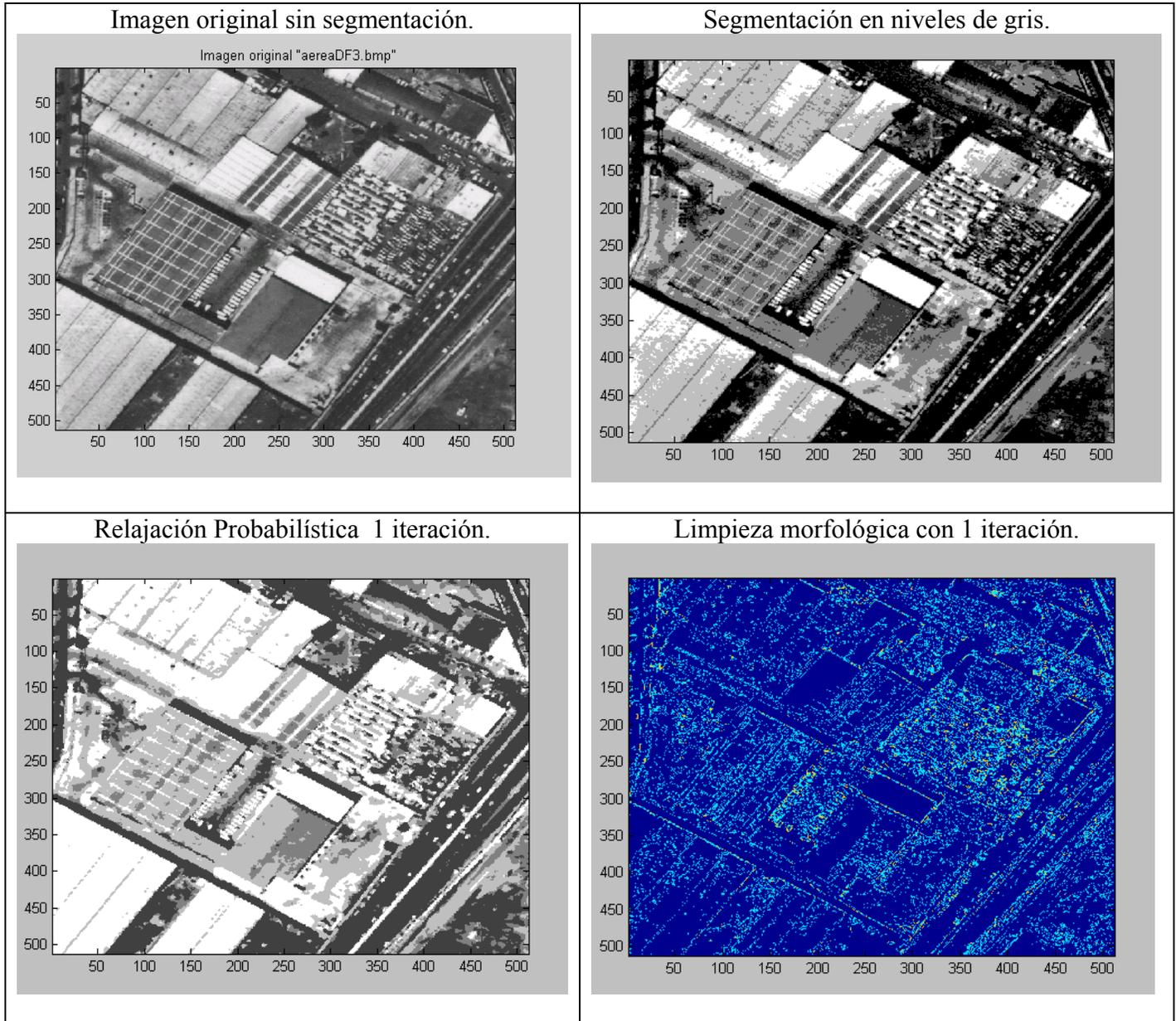


*Limpieza morfológica.*

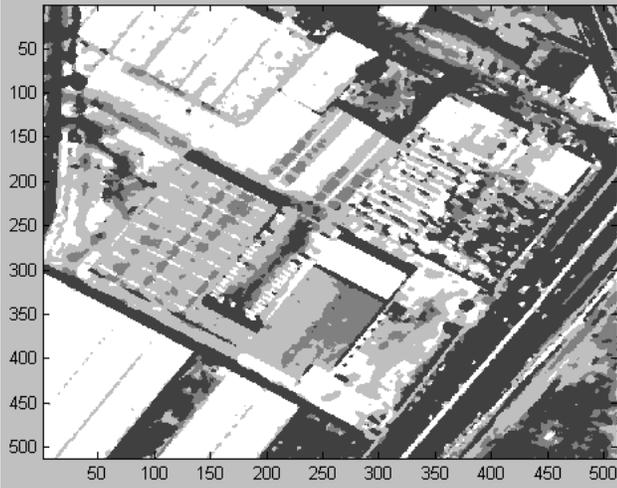


7.6.4 Relajación probabilística para la escena urbana 2

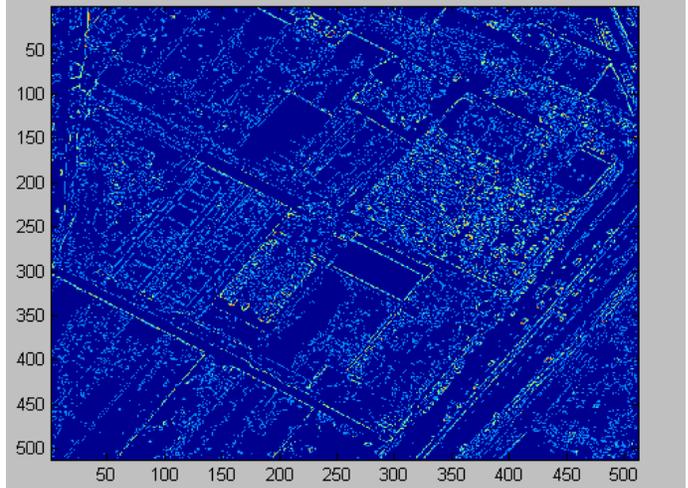
A manera de comparación, se muestra a continuación el resultado de la aplicación del método de relajación probabilística aplicado en la imagen segmentada de la escena urbana 2, por medio de funciones discriminantes de Bayes, propiamente con el caso general. Las imágenes de limpieza morfológica mostradas, se obtienen mediante la diferencia entre la imagen segmentada y un resultado de relajación, conforme se va iterando.



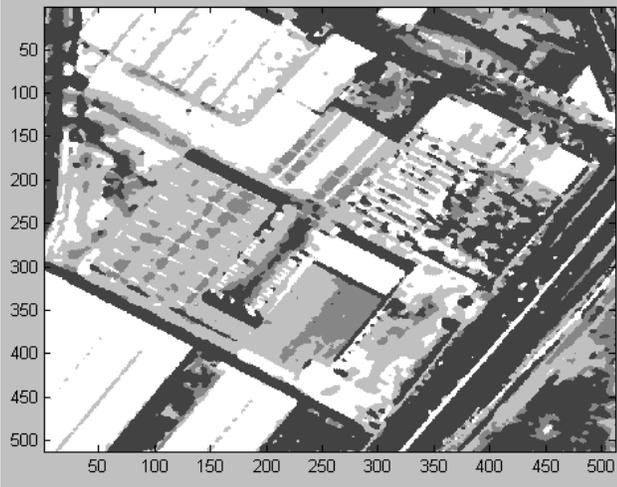
Relajación Probabilística 2 iteraciones.



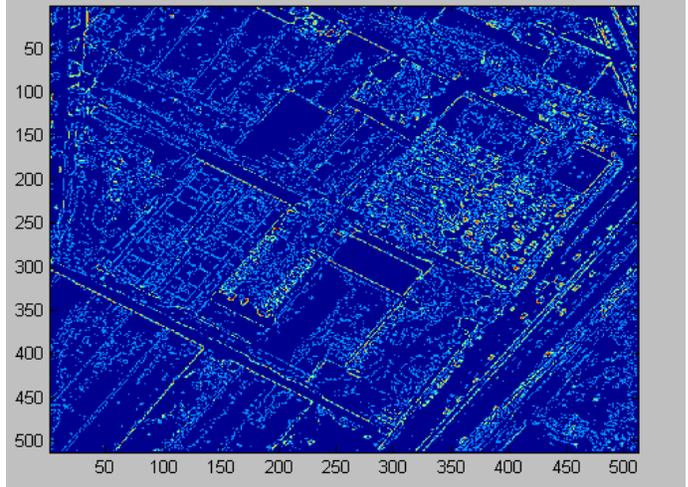
Limpieza morfológica con 2 iteraciones..



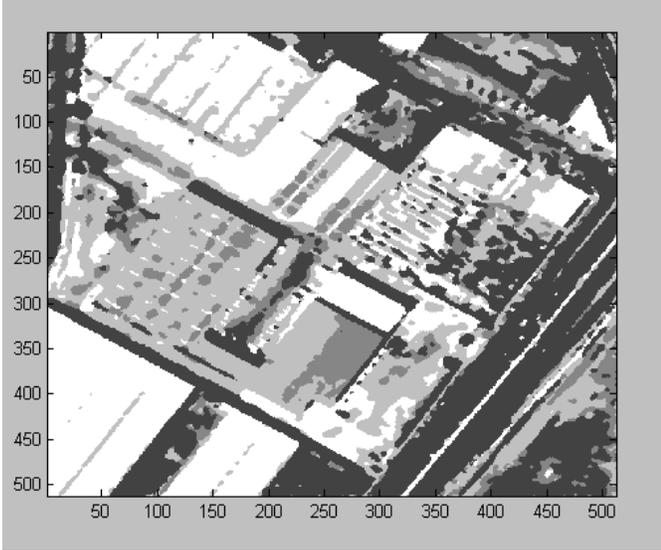
Relajación Probabilística 3 iteraciones.



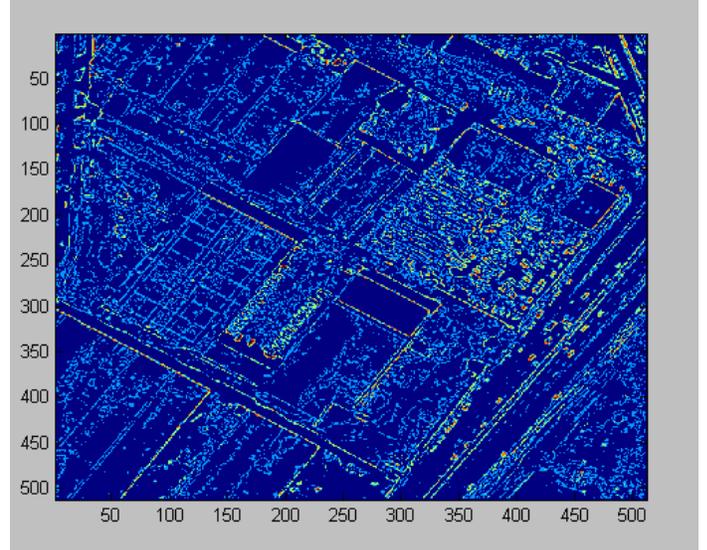
Limpieza morfológica con 3 iteraciones.



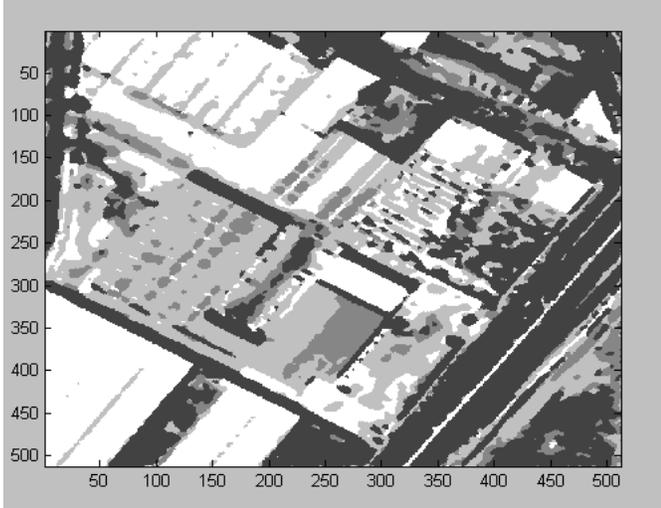
Relajación Probabilística 4 iteraciones.



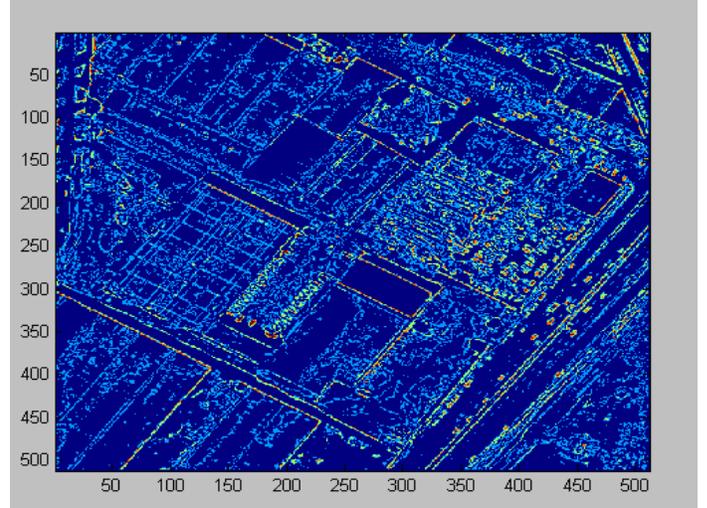
Limpieza morfológica con 4 iteraciones..



Relajación Probabilística 5 iteraciones.



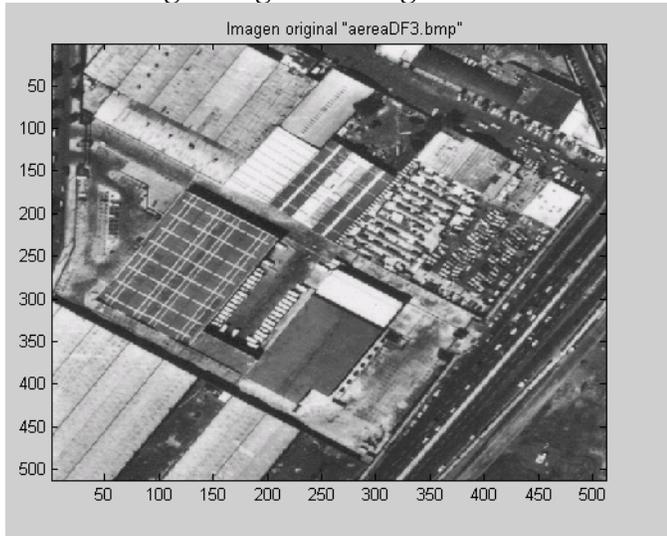
Limpieza morfológica con 5 iteraciones.



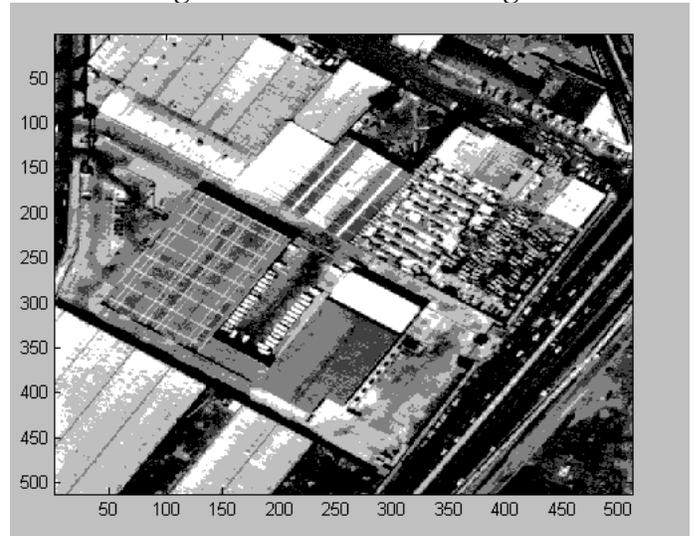
7.6.5 Relajación determinística para la escena urbana 2

A manera de comparación, se muestra a continuación el resultado de la aplicación del método de relajación determinística aplicado en la imagen segmentada de la escena urbana 2, por medio de funciones discriminantes de Bayes, propiamente con el caso general. Las imágenes de limpieza morfológica mostradas, se obtienen mediante la diferencia entre la imagen segmentada y un resultado de relajación, conforme se va iterando.

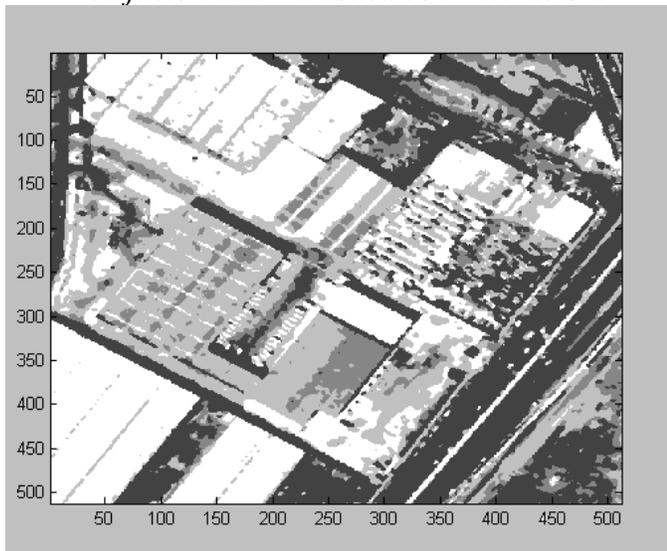
Imagen original sin segmentación.



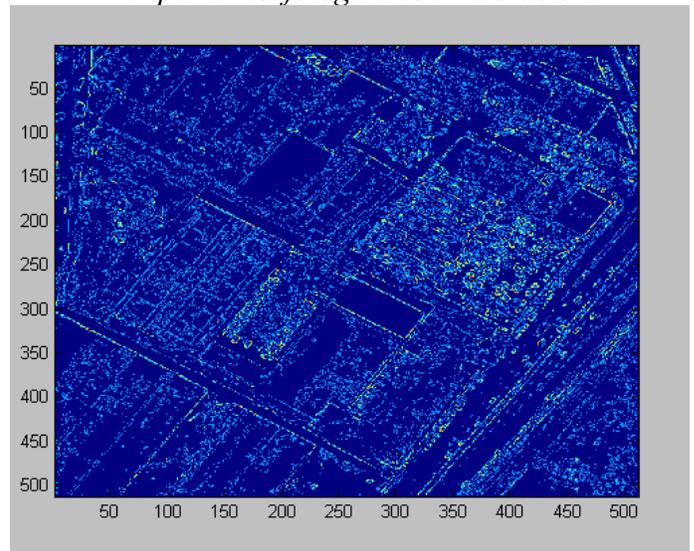
Segmentación en niveles de gris.



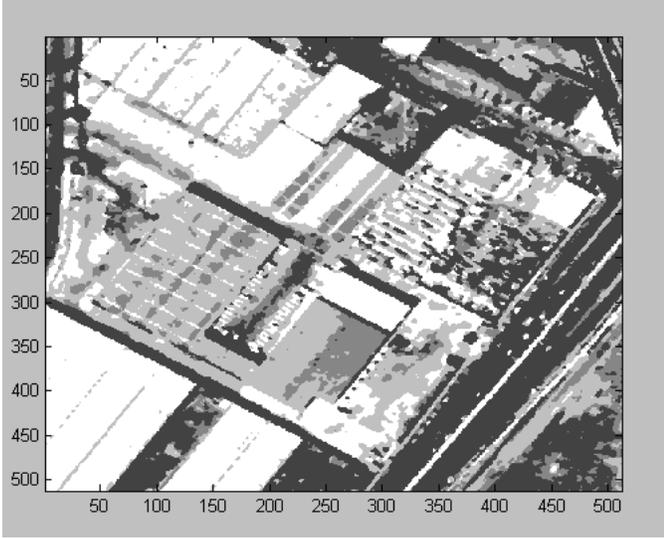
Relajación determinística con 1 iteración.



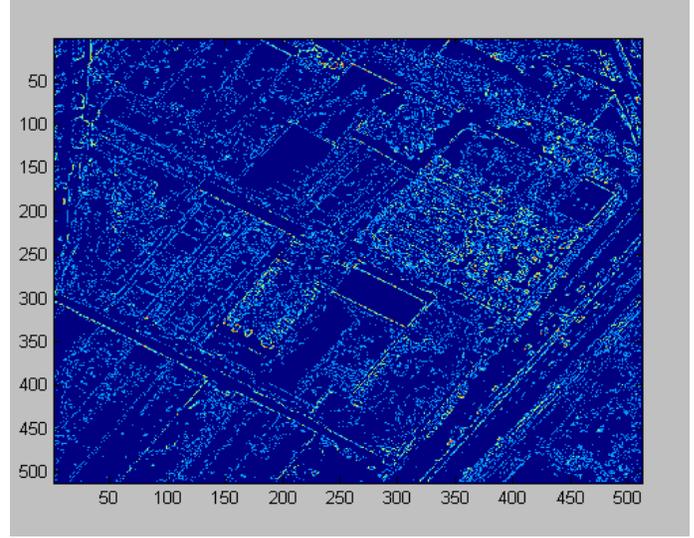
Limpieza morfológica con 1 iteración.



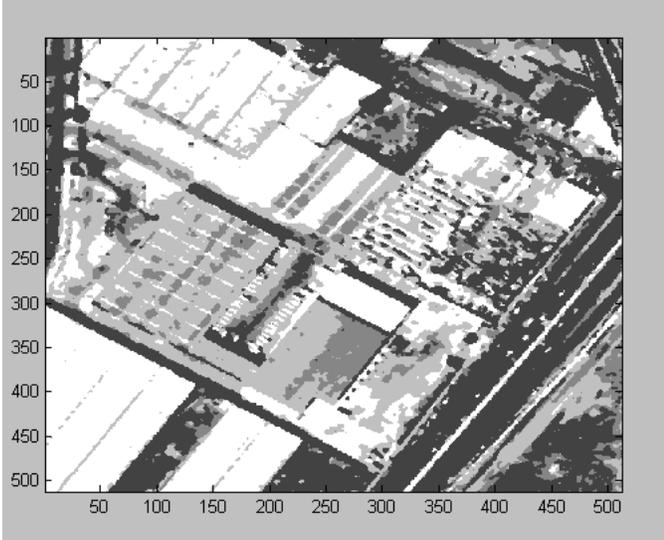
*Relajación determinística con 2 iteraciones.*



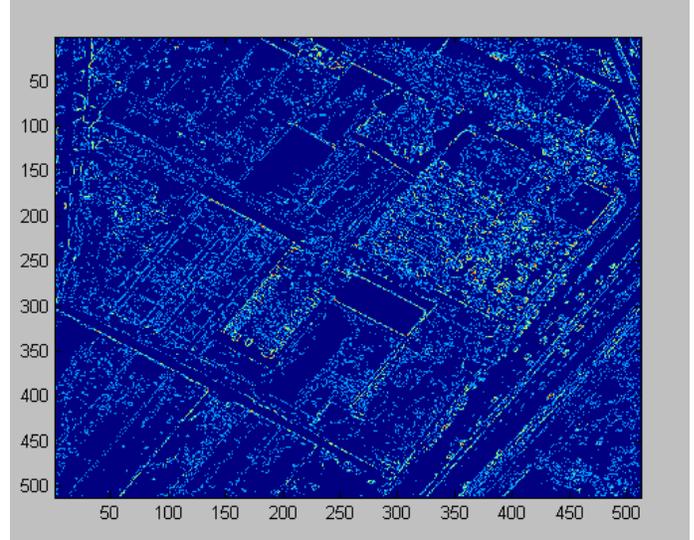
*Limpieza morfológica con 2 iteraciones.*



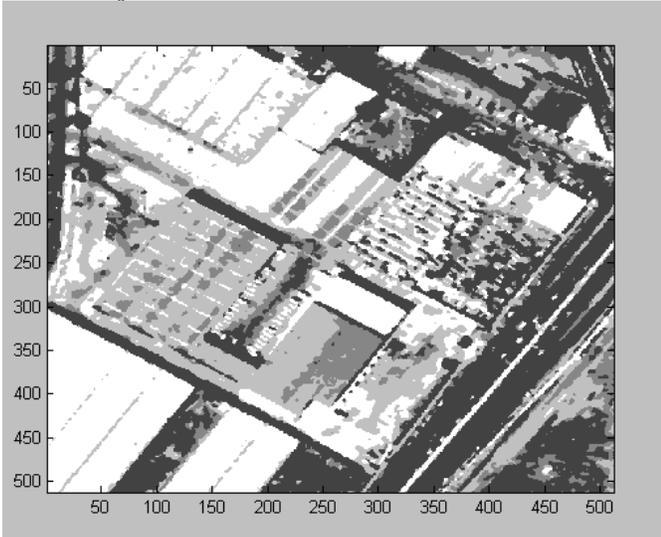
*Relajación determinística 3 iteraciones.*



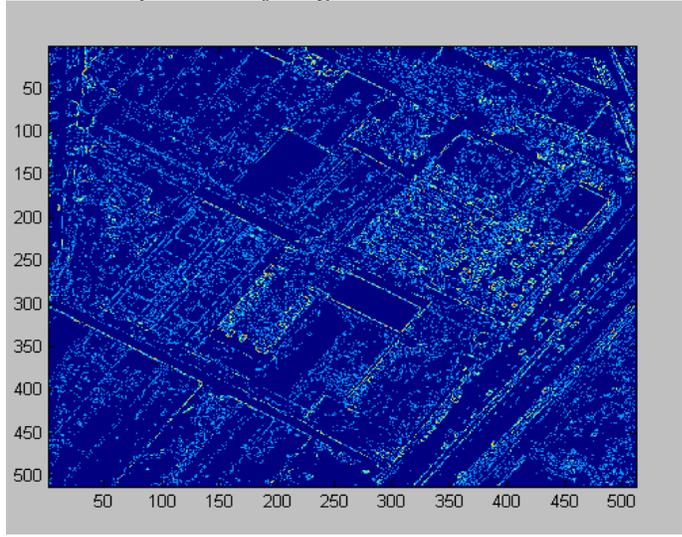
*Limpieza morfológica con 3 iteraciones.*



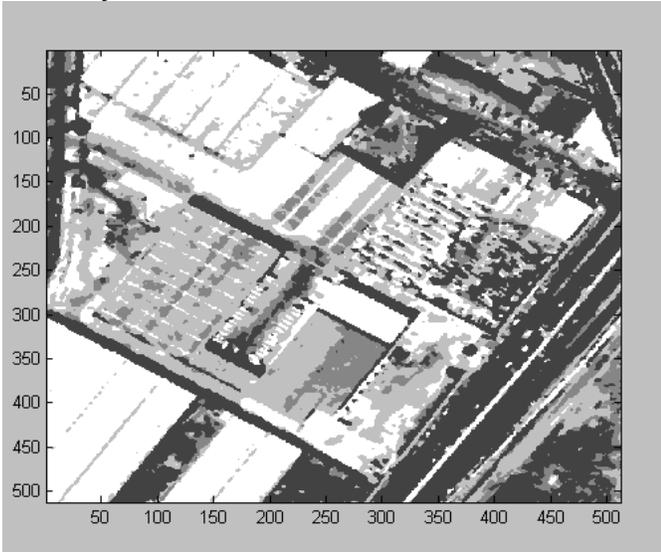
*Relajación determinística con 4 iteraciones.*



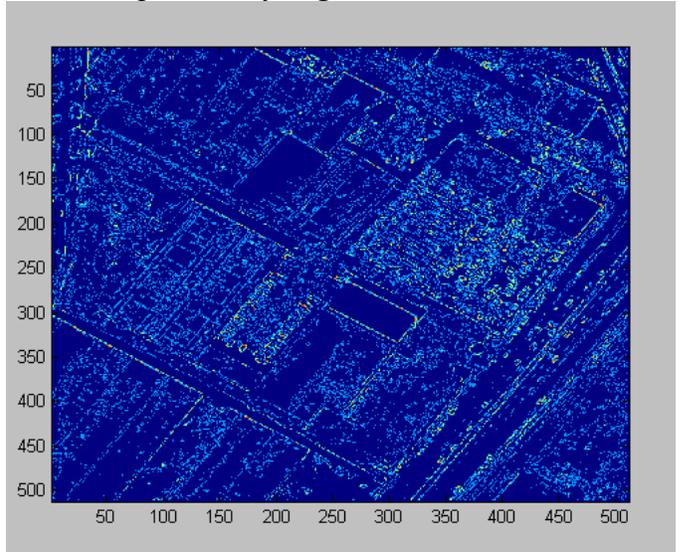
*Limpieza morfológica con 4 iteraciones.*



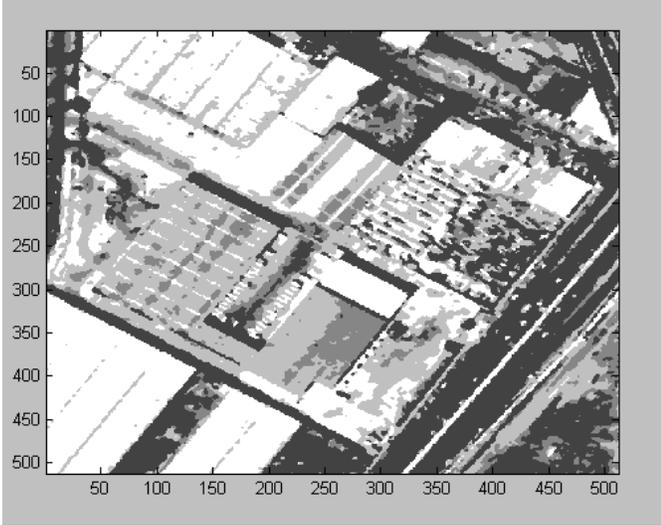
*Relajación determinística con 5 iteraciones.*



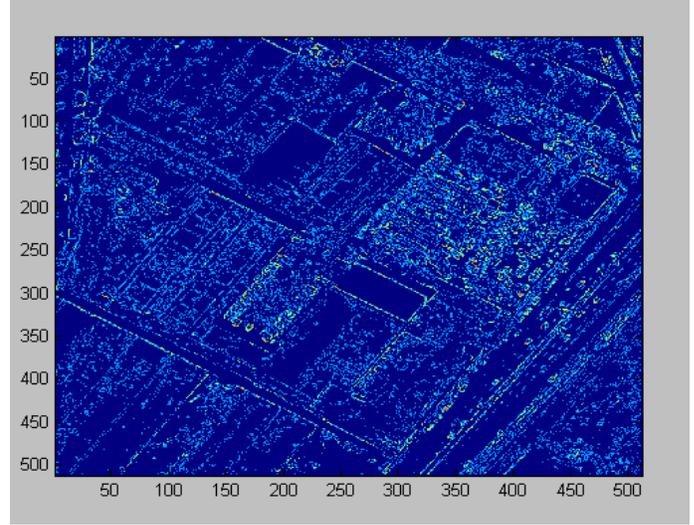
*Limpieza morfológica con 5 iteraciones.*



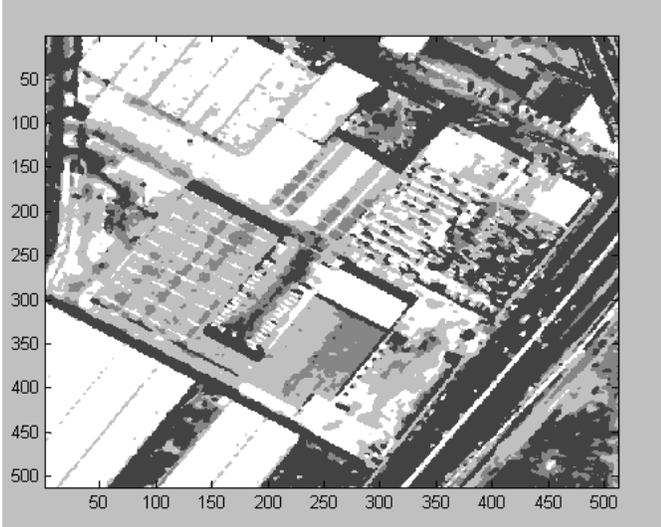
*Relajación determinística con 6 iteraciones.*



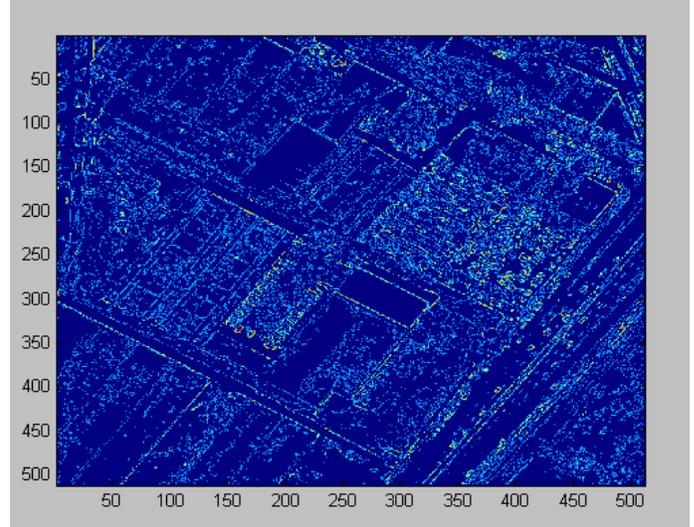
*Limpieza morfológica con 6 iteraciones.*



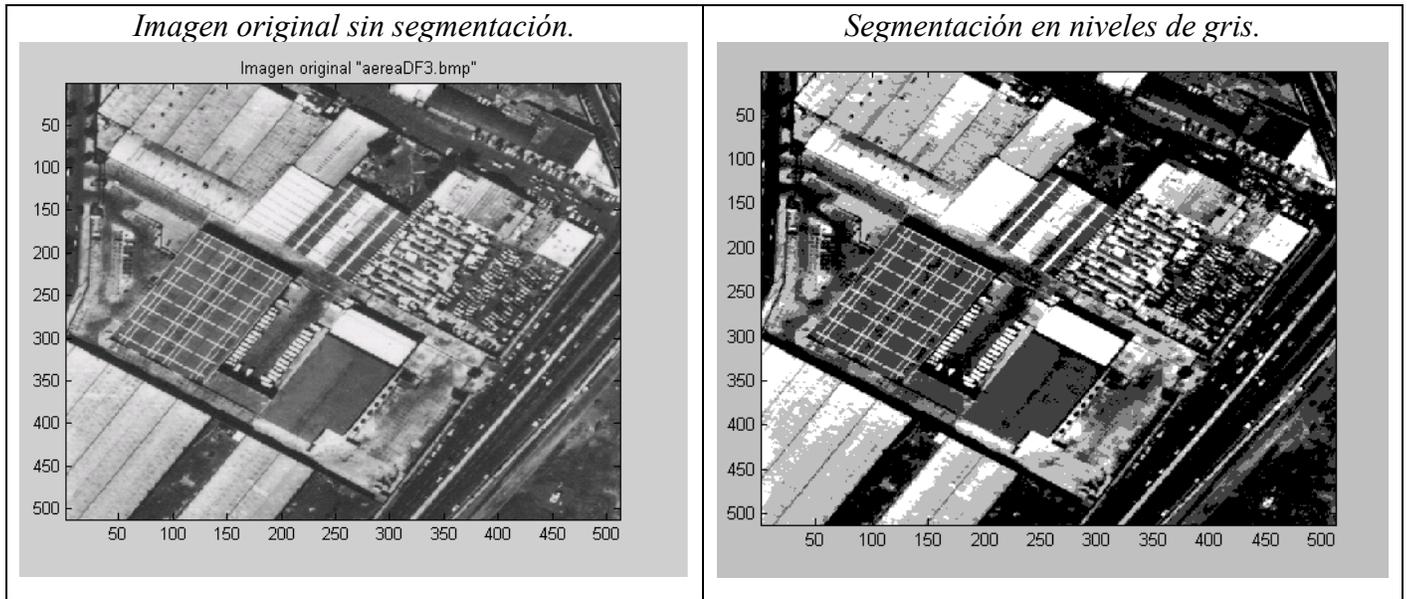
*Relajación determinística con 7 iteraciones.*



*Limpieza morfológica con 7 iteraciones.*

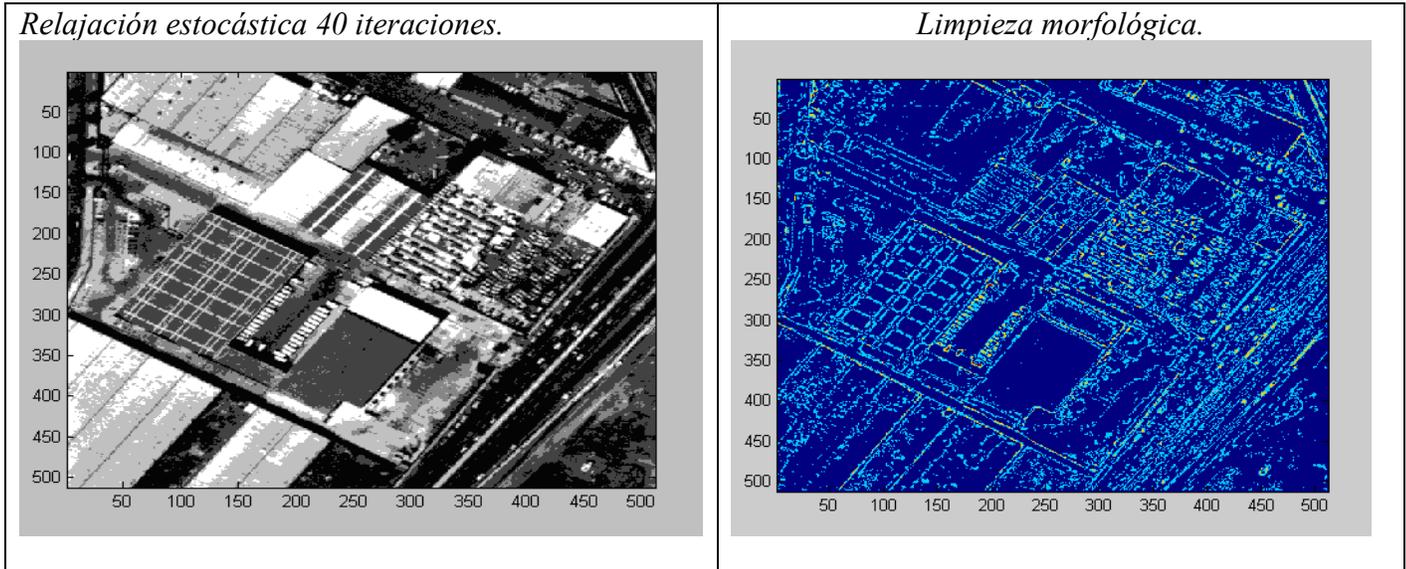


7.6.6 Relajación estocástica para la escena urbana 2, con diferentes configuraciones de vecindad y 40 iteraciones cada una



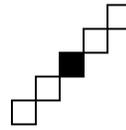
**Configuración 1:**

$U_2(y) = - \sum_{c_2 = \{i,j\}} \beta_1 d_{i=j}$	
---------------------------------------------------	---------------------------------------------------------------------------------------



**Configuración 2:**

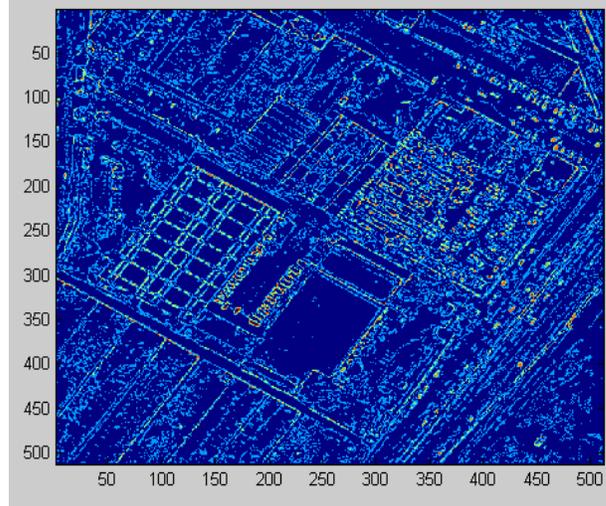
$$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{-i=j}$$



*Relajación estocástica 40 iteraciones.*

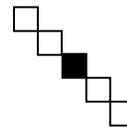


*Limpieza morfológica.*

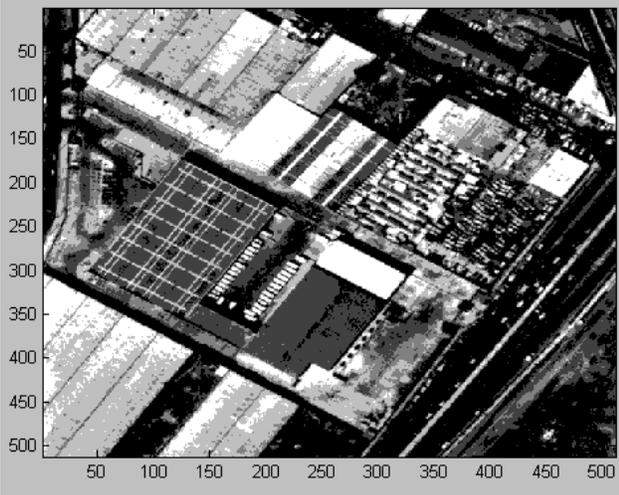


**Configuración 3:**

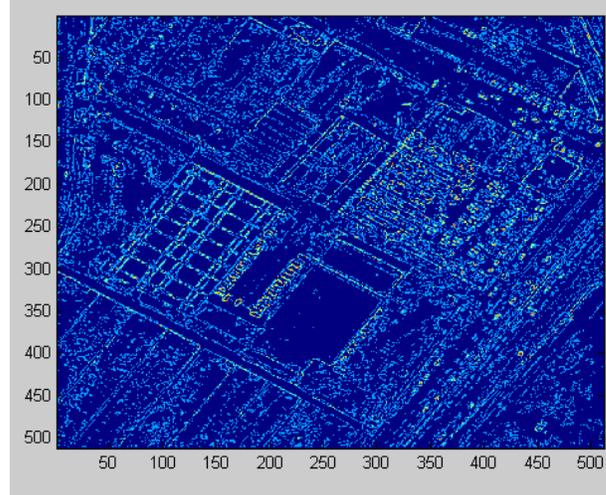
$$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i=-j}$$



*Relajación estocástica 40 iteraciones*

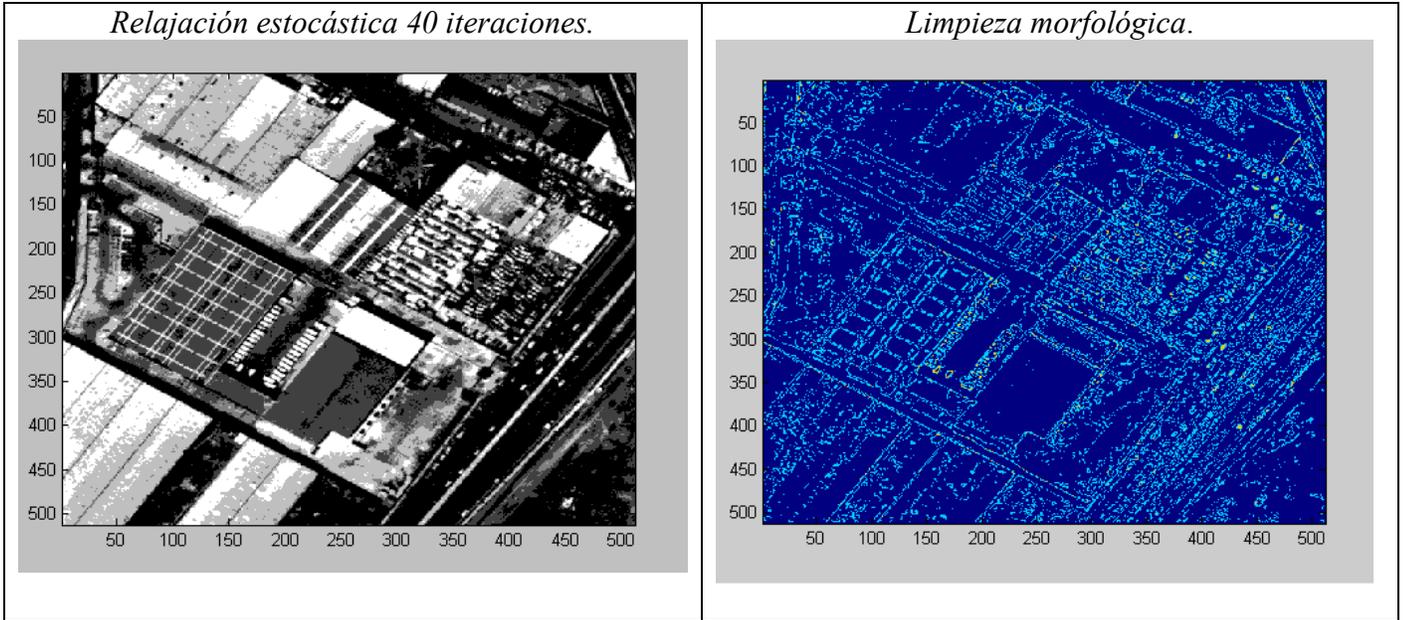


*Limpieza morfológica.*



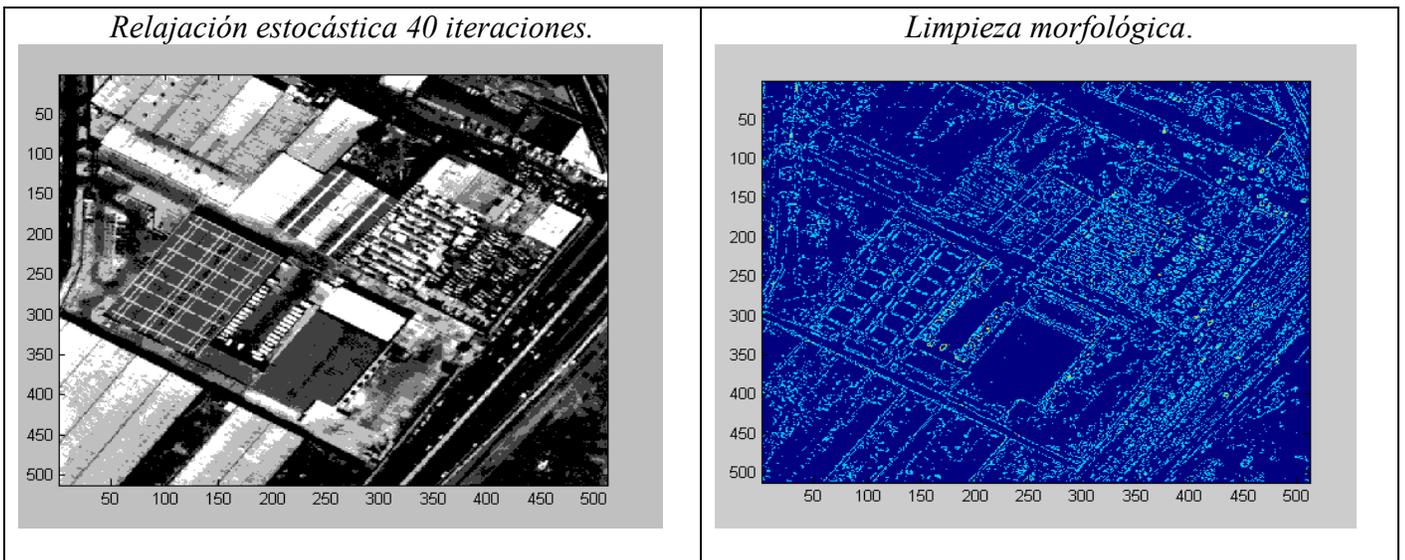
Configuración 4:

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{0j}$	
------------------------------------------------	-------------------------------------------------------------------------------------



Configuración 5:

$U_2(Y) = - \sum_{c_2=\{i,j\}} \beta_1 d_{i0}$	
------------------------------------------------	---------------------------------------------------------------------------------------



Posteriormente de tomar el resultado de la segmentación de la función discriminante para el caso 2, como muestra para el proceso de relajación con los tres métodos analizados, se pueden observar los resultados de cada uno de ellos a distinto número de iteraciones.

Se considera que el mejor resultado es el implementado mediante la relajación probabilística, tomando en cuenta que el número de iteraciones es menor en comparación con los otros dos métodos, resultando un menor costo computacional en un tiempo aproximado de 7.4591 minutos en cada iteración, al mismo tiempo se observa la calidad al tomar píxeles aislados, regularizando el etiquetado de la imagen de una manera más homogénea.

En el método de relajación probabilística, el algoritmo funciona de manera local, es decir la manera en que trabaja su implementación es a partir de las zonas locales o individuales, definidas por el número de clases. La eficiencia de lograr una imagen más suave desde el punto de vista de procesamiento digital y el bajo costo computacional derivado, hacen que ambos logros determinen que el método sea lo bastante conveniente para regularizar lo más exactamente posible una imagen, sin perder de vista que no se debe de perder información.

El algoritmo de la relajación determinística, usando el concepto de la energía mínima que se produce en base a la distancia de los píxeles, e introduciendo un elemento aleatorio dentro del esquema de iteración, para tratar de eliminar la sobre-segmentación, originando la unión o fusión (por llamarlo de alguna forma) de píxeles aislados. Al calcular la configuración de la energía mínima en 8-conexidad se observa que a pesar de la introducción de la aleatoriedad en el análisis de todos y cada uno de los píxeles al barrer la imagen, se logró obtener en un número de iteraciones adecuado, aunque con un tiempo considerable a la hora de arrojar el resultado, lo que implica un resultado óptimo, aunque quizás con el inconveniente del costo computacional en un tiempo aproximado de 12.7854 para cada iteración, debido seguramente de que al ser aleatorio puede volver a dejar sin fusionar algunos píxeles que en un número menor de iteraciones lo había obtenido, pero esto se corrige a medida que el número de iteraciones se va incrementando. No obstante proporciona un rápido avance en la decisión del etiquetado para la regularización.

Otro de los algoritmos presentados es el algoritmo de la relajación estocástica, implementado base al recocido simulado con la ayuda de la distribución de Gibbs. De acuerdo a las distintas configuraciones, se observa claramente que la más óptima es la configuración 1; debido a que su vecindad cubre un mayor número de píxeles respecto a las otras cuatro configuraciones, esto conlleva a que la configuración 1 minimice de mejor manera la función de energía a posteriori, propiciando así un resultado aceptable. Este algoritmo de relajación, más bien visto como una segmentación pero desde un enfoque de ayudar a precisar el menor número posible de píxeles aislados en la imagen, muestra cómo se obtiene una regularización homogénea fusionando los píxeles aislados de las regiones conexas, ayudando así a evitar en la sobresegmentación. El costo computacional en cada una de las configuraciones fue de 6.7483min para cada iteración,

por lo que se considera un tanto alto, si se toma en cuenta que el algoritmo se implemento para 40 iteraciones en cada una.

Se pretende mostrar el grado de regularización con las imágenes que muestran la limpieza morfológica llevada a cabo en cada uno de los algoritmos. Dicho de otra forma, esta limpieza morfológica, es el resultado del grado de homogeneización de los píxeles aislados, para generar de esta forma bloques o zonas homogéneas. El resultado que deriva de limpieza pareciera que solamente es un contorno de la original, pero tomando en cuenta la discontinuidad de la imagen, se puede apreciar cómo gradualmente se va obteniendo mayor número de píxeles aislados conforme el número de iteraciones se incrementa. Es por esto que se puede apreciar más en la discontinuidad de los contornos, ya que en estas zonas de las imágenes es donde reside la importancia de la regularización.

---

---

## 8 Conclusiones

Como ya se ha mencionado, la función de la percepción remota ayuda al análisis de las imágenes captadas por sensores remotos. Así mismo los métodos y algoritmos utilizados en el procesamiento digital de imágenes ayudan a distinguir de una mejor manera la información obtenida a través de estos sensores, y con la ayuda de una pc para procesar los datos, se logra el reconocimiento de información para asignar un significado a un conjunto de regiones u objetos reconocidos.

Para el desarrollo de este trabajo, en los primeros capítulos se trato de introducir la teoría básica, así como los métodos de procesamiento para el tratamiento de la imagen pancromática aérea a procesar, con un tamaño de 512x512 píxeles, adjudicando 256 niveles de gris.

El esquema general presentado al final de este capítulo, donde se incluye diferentes métodos de análisis de segmentación, proporciona una temática a seguir, para lograr el objetivo de conseguir con un menor recurso de programación el proceso de homogeneización de una manera más eficiente para las imágenes con las cuales se trabajó. Estos algoritmos proporcionan una base para el estudio de detección y localización de diversas estructuras irregulares contenidas en una imagen.

La homogeneización llevada a cabo, resultó en un etiquetado de una forma más concisa y precisa; para lograr identificar todas las zonas representadas por cada una de las diferentes clases.

En este trabajo se emplearon algoritmos tanto de segmentación bayesiana como de segmentación markoviana, la primera se refiere a una segmentación no contextual, es decir que no toma en cuenta el entorno vecinal; en tanto que la segunda en contraparte es una segmentación contextual. En ambos casos los criterios más importantes que se tomaron en cuenta fueron: primeramente el grado de homogeneidad de la escena y en segundo término la discontinuidad de regiones en la escena. Es por ello que en este trabajo de tesis se determinó utilizar un procedimiento de segmentación lo menos tosco posible, para que así no se dilatará la solución satisfactoria del problema, es decir evitar que se utilicen demasiados recursos de la computador a la hora de la programación, procurando al mismo tiempo no implementar un algoritmo de segmentación débil, que garantizaría seguramente algún fallo en el etiquetado de la imagen.

Se escogió la función discriminante de Bayes, implementándola como un método de decisión teórica para el reconocimiento, encontrando el número de funciones de decisión, en relación directa con el número de clases con las que se trabajó ( $k = 5$  en total); considerando la pertenencia de un píxel dado a una cierta clase. Además con esto se pudo optimizar de una manera más exacta la frontera de decisión, que permitió separar una clase

de otra, mediante la resta de las funciones discriminantes, así como de las raíces o **puntos discriminantes** que se obtuvieron, con la ayuda de etiquetar cada punto discriminante a ciertos valores de umbrales predeterminados.

Se utilizó el resultado de las medias obtenidas en la función discriminante, tanto para el desarrollo del algoritmo de recocido simulado en la Segmentación markoviana, así como para el proceso de la relajación estocástica. No obstante se tiene el algoritmo de K-medias para la obtención de las medias, pero se prefirió optar por las obtenidas en la función discriminante, tomando como antecedente la manera en que se obtuvieron; asumiendo que los resultados de estas medias fueron lo más acertado posible, y que de modo que nos ayudarían al siguiente proceso. Tratando de llevar una secuencia lógica en base a resultados.

El siguiente paso fue el de implementar un algoritmo para la segmentación markoviana, en este caso el de recocido simulado, mismo que serviría de base en la implementación de la relajación estocástica. Con la teoría de los CAM, se pudo obtener de manera acertada, además del proceso de la formación de imagen de un modo iterativo, el proceso de etiquetado de diferentes regiones, con la ayuda del decremento de una temperatura inicial, y de la función de energía que permite el etiquetado de cada píxel de la imagen original, decidiendo a que clase pertenecen todos y cada unos de estos píxeles considerando la dependencia contextual de los píxeles que componen la imagen, mediante el criterio de Potts.

Definiendo previamente una segmentación razonablemente correcta, se procedió al análisis de los métodos de relajación propuestos. En este trabajo de tesis, se presentan tres diferentes algoritmos de relajación; los cuales fueron el algoritmo de relajación probabilística, determinística y estocástica. Cuyos resultados experimentales son derivados de la teoría mencionada que se encuentra dentro del capítulo 6; y los cuales se ilustran y comparan dentro del capítulo 7 de resultados.

Empleando el método del proceso de la relajación probabilística, puede ser implementado para obtener un cierto nivel fino de resolución, para un número pequeño de iteraciones. La segmentación automática de la imagen que se analizó, mejora el nivel de segmentación, para la localización del proceso de etiquetado en la imagen. El mejoramiento consistió en la regularización del etiquetado, con esto se logró proporcionar un método estadístico que logra determinar con mayor precisión el número de clases dentro de una imagen.

El algoritmo I.C.M (modo condicional iterativo) implementado para la relajación determinística, puede mencionarse que es adecuado llegando a un mínimo local de energía, con la configuración de partida, aunque oscile entre varios mínimos locales tomando en cuenta que la energía de estos es similar, cabe mencionar que resulta efectivo para la regularización de la imagen.

El algoritmo de la relajación estocástica representado por un CAM con la ayuda de la distribución de Gibbs pero desde un enfoque o modelo no causal, el cual puede ser visto

como una aproximación de los otros dos, se puede inferir que ambos conceptos pueden adaptarse de forma muy concreta y flexible, para la representación de las dependencias de la intensidad de niveles de gris de píxeles cercanos en una imagen. Además, de mencionar que la degradación es independiente de la información que se tiene de la imagen, y que más bien depende del número de iteraciones que se efectúa el proceso. Con las cuales se pretende tener una distribución posterior, par que de esta forma se alcance con una estructura de vecindad razonable como lo es asumido en los CAM. De esta manera, se examina que el proceso sea adaptado de acuerdo a aproximaciones contextuales de la imagen a procesar, pretendiendo lograr que el algoritmo sea un sistema dinámico.

El resultado de la simulación presentada muestra que el algoritmo para la relajación estocástica supera al algoritmo ICM significativamente; aún cuando se comience desde una buena condición inicial. Esto es debido a que se manejaron distintas configuraciones de vecindad lo cual implica que el análisis sea más exhaustivo y concreto. Pero a pesar de llevar a cabo este análisis, no supera al algoritmo de relajación probabilística.

De a cuerdo a los resultados mostrados se puede intuir que el etiquetado en el proceso de relajación es un método formal muy útil para regularizar la información de una imagen que contiene un bajo nivel de información contextual. Con esto se puede mencionar que el etiquetado en el proceso de relajación, provee de una dependencia contextual adecuada, con prioridad de la información de los píxeles y su entorno, generando bloques o zonas homogéneas.

Cuando se usa el proceso de relajación basada en la segmentación, con ayuda de esta se conoce automáticamente de acuerdo a la cercanía o lejanía de un píxel dado a cada una de las regiones, a partir del valor de sus medias, tomando una medida simple de la diferencia absoluta entre el nivel de gris del píxel y la media característica de la región, proporcionada por el valor de las medias de los píxeles que la componen. He aquí la importancia de este valor estadístico para el proceso.

A manera un tanto hipotética se podría mencionar que comúnmente se observa en los esquemas de relajación que después de aplicar un número pequeño de iteraciones el resultado obviamente es erróneo, ya que el píxel etiquetado puede moverse de una clase a otra, por lo tanto de a cuerdo al número de iteraciones, surge la dependencia del resultado, de como tiende a converger a una solución adecuada. Así mismo cabe mencionar que se debe tener cuidado de no sobrepasar el número de iteraciones.

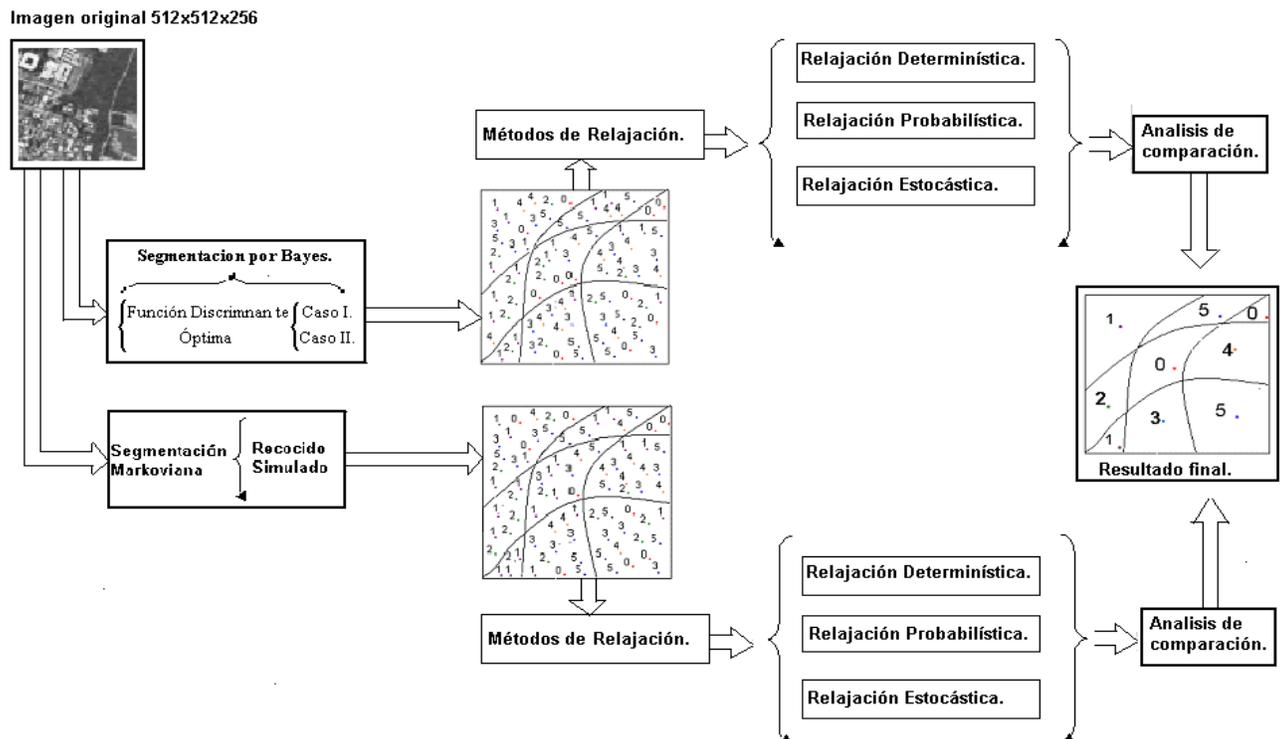
Si el número de iteraciones es grande beneficiara en algunas zonas porque elimina clases que no deberían de estar presentes en el contexto; pero en otras zonas donde existen clases significativas con información importante que no se puede omitir, un número grande de iteraciones dará como resultado la perdida de información que pudiera ser relevante dentro del análisis de la imagen, dando como consecuencia que será menor el número de segmentos pequeños en los que puede incluir zonas características en la imagen.. En otras palabras, cuando se somete la imagen a los métodos de relajación debe de tenerse cuidado de tomar un número de iteraciones adecuado, de manera que no se pierda la resolución de

la imagen y con esto conlleva a un resultado de binarización, lo cual sale del objetivo de este trabajo. Propiciando que se tenga un resultado amorfo.

Es por lo anterior que se asume la dependencia de con respecto a qué es lo que se quiere comparar o analizar en la imagen, para que de esta forma se pueden utilizar estos métodos, y así poder decidir exactamente que es lo relevante en la información contenida en la imagen o qué parte o partes interesa analizar con más detalle.

Tomando en consideración que la manera de trabajar de estos algoritmos es de una forma heurística, es decir que en la práctica funcionan de manera aceptable con los recursos de programación implementados, pero que no se cuenta con un método exhaustivo de comprobación del porqué sus resultados son aceptables. Es por esto que en este trabajo, se sugiere hacer un análisis de los tres métodos a manera de comparación entre uno y otro, manifestándolo como una forma de comprobación.

Por último se presenta el siguiente esquema que representa la metodología empleada en este trabajo de tesis. Del cual se determino que resultado presentaba una mejor solución para el análisis que se pretendió, y el cual dio pauta a presentar el esquema que se encuentra en el capítulo de introducción (capítulo 1).



### **A**

**Absorción.-** Llamada la energía de un haz de radiación electromagnética que es transformada en otras formas de energía. Es una atenuación selectiva observada en varios constituyentes, tales como vapor de agua, ozono y monóxido de carbono. En muchos casos, la absorción puede ser despreciada por ser muy pequeña.

**Angulo acimut.-** Angulo medido en el sentido de las manecillas del reloj a partir del norte, comprendido entre el meridiano de un lugar y el plano vertical en que este el objeto.

**Altitud.-** Altura de un punto de la tierra con relación al mar, desde un punto de referencia.

**Ancho de banda.-** Es la medida del intervalo de frecuencias en que opera una señal.

**Ancho de barrido.-** Es la dimensión de alcance de la escena o imagen que se desea analizar, medido del alcance terrestre o del alcance inclinado, según sea el caso.

**Angulo de vista.-** Es el ángulo que permite observar el objeto en el punto nodal de incidencia de este con la normal de la superficie de separación en el objeto y el sensor.

**Atenuación.-** Perdida de fuerza o amplitud de las señales de transmisión a través de la distancia, razón principal de varias restricciones entre los sistemas de telecomunicaciones.

### **B**

**Banda espectral** es el intervalo entre dos longitudes de onda, en el espectro electromagnético.

**Borde.-** Extremo u orilla de algún objeto o zona que se distingue por tener una tonalidad marcada.

### **C**

**Cenit.-** Punto de intersección entre la vertical de una referencia y la esfera celeste superior al horizonte.

**Clasificación.-** Se refiere a las técnicas de diferenciar la información expuesta para su mejor tratamiento, estas pueden ser divididas en: clasificación supervisada (por píxel) y clasificación no supervisada (por regiones).

**Clique.-** Conjunto de píxeles que viene determinado por su tamaño, dirección y orientación.

**Contraste.-** Procedimiento que hace una diferencia palpable entre las tonalidades de un objeto, haciendo resaltar los tonos más vivos respecto a sus tonos complementarios. Definido como la razón entre sus niveles de gris medios.

**Covarianza.-** En el cálculo de probabilidades es el valor que corresponde a la mayor o menor correlación o dependencia mutua que existe entre dos variables aleatorias

representando la media del producto de las desviaciones de dos variables con relación a su media. Busca comparar los resultados en diferentes grupos de una variable cuantitativa, pero “corrigiendo” las posibles diferencias existentes entre los grupos en otras variables que pudieran afectar también al resultado (covariantes).

**Cluster.-** Conjunto o grupo de píxeles con características de similitud.

## ***D***

**Dispersión.-** La energía de un haz de radiación electromagnética es removida por cambios de dirección. Al interactuar con la atmósfera, por el proceso de dispersión, generará un campo de luz difusa, que se propagará en todas las direcciones.

**Dispersión molecular o Rayleigh.-** Producida esencialmente por las moléculas de los gases de la atmósfera. Este tipo de dispersión se caracteriza por el hecho de que su intensidad es inversamente proporcional a la cuarta potencia de la longitud de onda ( $E = 1/\lambda^4$ ), lo que explica la coloración azul del cielo.

**Dispersión MIE.-** Ocurre cuando el tamaño de las partículas que causan la dispersión es similar a la longitud de onda de la radiación.

**Dispersión no selectiva.-** Ocurre cuando los diámetros de las partículas son mayores que la longitud de onda. La radiación de diferentes longitudes de onda será dispersa con igual intensidad. La apariencia blanca de las nubes es explicada por este proceso.

## ***E***

**Efemérides.-** Conjunto de parámetros que describen las orbitas de los satélites, recalculados con sus variaciones, sobre las orbitas predeterminadas y programadas en los sistemas de navegación de los satélites.

**Energía absorbida.-** Es la energía derivada cuando la energía incidente se almacena y espárce en todas direcciones, siendo una atenuación selectiva observada en varios constituyentes como vapor de agua, ozono y monóxido de carbono.

**Energía incidente.-** Es la energía linealmente polarizada que se reemite en formas de ondas que oscilan en un mismo plano.

**Energía reflejada.-** Es la energía incidente que regresa desde la superficie formando un ángulo con la normal a la superficie igual al ángulo de incidencia, variando con respecto a la energía incidente.

**Energía transmitida.-** Es la energía que causa un cambio en la velocidad de la radiación electromagnética generalmente con longitudes de onda más largas.

**Espectro de onda.-** Es producido por la dispersión de la radiación y su división en las frecuencias que la componen.

## ***F***

**Función discriminante.-** Función de decisión teórica utilizada en el reconocimiento de patrones, modelados estadísticamente.

## ***H***

**Heurístico.-** Procedimiento relativo a la investigación a base de proponer soluciones predeterminadas. Algoritmo que consta de utilizar pruebas o aproximaciones para llegar a dar con una solución. De esta forma, sin conocer unos datos base exactos, se puede llegar a una solución o resultado final.

## ***I***

**Imagen Pancromática.-** Imagen que se caracteriza por tener una sola banda que cubre gran parte del espectro visual.

**Intensidad.-** La intensidad o brillo es la medida de la energía total en todas las longitudes de onda, siendo por lo tanto responsable por la sensación de brillo de esa energía incidente sobre el ojo.

**Infrarrojo.-** De gran importancia para la Percepción remota. Engloba la radiación con longitudes de onda de 0,75  $\mu\text{m}$  a 1,0 mm. La radiación infrarroja es fácilmente absorbida por la mayoría de las sustancias (efecto de calentamiento).

## ***L***

**Longitud de onda.-** es la distancia entre dos crestas (parte más alta) sucesivas de la onda.

**Latitud.-** Distancia que hay desde un punto de la superficie terrestre al ecuador, contada por los grados de su meridiano.

## ***M***

**Microondas.-** Se sitúan en la faja de 1 mm a 30 cm o 3 X 10<sup>11</sup> a 3 X 10<sup>9</sup> Hz. En esta franja se pueden construir dispositivos capaces de producir haces de radiación electromagnética altamente concentrados, llamados radares. La poca atenuación por la atmósfera o por las nubes, hace propicio un excelente medio para uso de sensores de microondas en cualquier condición climática.

**Matiz.-** El matiz o color de un objeto es la medida de la longitud de onda media de la luz que se refleja o se emite y que define por lo tanto, el color del objeto.

**Media.-** Es el cociente que resulta de dividir la suma de las variables continuas o discretas por el número de ellas, dando como resultado una centralización o valor promedio.

## ***N***

**Nadir.-** Punto de la esfera celeste opuesto al cenit, situado en la vertical del observador en el plano vertical.

**Niveles de gris.-** Es el valor designado en tonalidades de la escala de gris, representado por la radiancia media de un área relativamente pequeña en una escena. En una imagen se cuenta con un intervalo definido entre 0 (negro) y 255 (blanco).

## ***O***

**Ondas de radio.-** Las ondas electromagnéticas en esta faja son utilizadas para la comunicación a larga distancia, ya que, además de ser poco atenuadas por la atmósfera, son reflejadas por la ionosfera, propiciando una propagación de largo alcance.

**Orbita heliosíncrona.-** Orbita que rota mismo tiempo que el sol formando un ángulo constante con la recta tierra-sol.

## ***P***

**Píxel.-** Es el elemento de representación más pequeño o unidad mínima en una imagen digital.

**Preprocesamiento.-** se refiere al procesamiento inicial de los datos brutos para la calibración radio-métrica de la imagen, la corrección de distorsiones geométricas y la eliminación de ruido.

**Polarización.-** Es la propiedad que poseen las ondas transversales después de sufrir la reflexión o la refracción de producir vibraciones localizadas desigualmente alrededor de la onda incidente.

## ***R***

**Radiación electromagnética.-** Se le conoce también como energía radiante. Tiene características de ondas y de partículas. Un ejemplo de un tipo de radiación electromagnética es la luz visible.

**Radiación visible.-** Es definida como la radiación capaz de producir la sensación de visión para el ojo humano normal. Presentan una pequeña variación de longitud de onda (380nm a 750 nm). Importante para la Percepción remota, ya que las imágenes obtenidas en esta franja generalmente, presentan una excelente correlación con la experiencia visual del intérprete.

**Radiación ultravioleta.-** Intervalo del espectro (10 nm a 400 nm). Las películas fotográficas son más sensibles a la radiación ultravioleta, que a la luz visible. Este intervalo es

utilizado para la detección de minerales por luminiscencia y polución marina. Uno de los grandes obstáculos para la utilización de esta región del espectro es la fuerte atenuación atmosférica.

**Radiancia.-** Se denomina como el flujo radiante que proviene de una fuente en una determinada dirección por unidad de área.

**Radiación ultravioleta.-** Es la radiación electromagnética con una longitud de onda menor a la de la luz visible mayor a la de los rayos x.

**Radiómetro.-** Instrumento para detectar y medir la intensidad de energía térmica radiante, en especial de rayos infrarrojos.

**Rayos gama.-** Son los rayos más penetrantes de las emisiones de sustancias radioactivas. No existe, en principio, límite superior para la frecuencia de las radiaciones gama, aunque se encuentre aún una faja superior de frecuencias para la radiación, conocida como rayos cósmicos.

**Rayos infrarrojos.-** La radiación infrarroja se localiza en el espectro entre  $3 \times 10^{11}$  hz. Hasta aproximadamente los  $4 \times 10^{14}$  Hz Se subdivide en tres regiones, infrarrojo lejano, medio y cercano (Lejana 50000 - 1mm, Intermedia 2500 - 50000 nm y a lo visible 780 - 2500 nm).

**Rayos X.-** Su intervalo es de 1 Å a 10 nm (1 Å =  $10^{-10}$  m). Son generados, predominantemente por la parada o frenada de electrones de alta energía. Por estar constituido por fotones de alta energía, los rayos X son altamente penetrantes, siendo una poderosa herramienta en la investigación sobre la estructura de la materia.

**Realce.-** Se refiere a las técnicas más comunes en PDI las cuales incluyen: realce de contraste, filtraje, operación aritmética, transformación IHS-RGB y componentes principales.

**Resolución espacial.-** Es el mínimo detalle espacial (píxel) que registra un sensor. Depende del sistema óptico del sensor y de la altitud de la plataforma.

**Resolución espectral.-** Es la cantidad de bandas y sus respectivos rangos espectrales con que se capta la energía electromagnética.

**Resolución radio métrica.-** Es la capacidad de un sensor para registrar pequeños cambios de energía. Se mide en cantidad de niveles de grises o de cuentas digitales.

**Resolución temporal.-** Es la frecuencia de observaciones del sensor sobre un objeto. No depende del sensor sino de la plataforma.

## S

**Saturación.-** expresa el intervalo de longitud de onda alrededor de la longitud de onda media, en el cual la energía es reflejada o transmitida. Un valor alto de saturación resulta en un color espectralmente puro, mientras que un valor bajo indica una mezcla de longitudes de onda que va a producir grisáceos neutros (apagados).

***T***

**Telemetría.-** Es una tecnología que permite la medición remota de magnitudes físicas y el envío posterior de la información generada en la medición hacia el operador del sistema de telemetría.

**Textura.-** Se refiere a la frecuencia de cambio y a la disposición de los tonos en una imagen y se le describe por términos de variación de contraste entre píxeles vecinos.

***U***

**Umbralaje.-** Se refiere a establecer un límite o separación entre objetos, respecto al nivel de gris que los componen.

***V***

**Varianza.-** Representa la variabilidad de una muestra, describiendo la dispersión de los valores alrededor de sus medias.

# **Bibliografía**

## *Libros de texto y notas [txt]*

[1] Tratamiento digital de imágenes.  
Rafael C. Gonzalez.  
Richard E. Woods.  
2ª Edición en español.  
Addison-Wesley / Diaz de Santos.  
Madrid 1995.

[2] Pattern recognition and image preprocessing.  
Sing-Tze Bow.  
Department of electrical engineering Northern Illinois University, Dekalb, Illinois.

[3] Random processes for image and signal processing.  
Edward R. Dougherty.

[4] Markov random field modeling in computer vision.  
2ª. Edición.  
S.z. Li.  
Springer-Verlag New York Inc 1995.

[5] Pattern recognition.  
Concepts, methods and applications.  
J. P. Marques de Sá.

[6] Matlab edición de estudiante  
Con tutorial de Duane Hanselman y Brice Littlefield.  
Versión 4 guía de usuario.  
Prentice Hall.

[7] Image processing toolbox for use with matlab.  
Users guide.  
Version 3.

[8] Notas de Análisis de Señales Aleatorias Semestre 2003-1 Dr. Miguel Moctezuma Flores.

***Direcciones electrónicas [p. e.]***

***Fundamentos de percepción remota.***

- [1] <http://www.teledet.com.uy/quees.htm>  
“Percepción remota satelital ”
- [2] <http://www.aeroespacio.com.ar/site/anteriores/529-537/531/sar.htm>  
“Procesamiento digital de imagines SAR”
- [3] [http://www.dpi.inpe.br/spring/usuario\\_spa/sensorr.htm](http://www.dpi.inpe.br/spring/usuario_spa/sensorr.htm)  
“Introducción a la percepción remota”
- [4] <http://www.profc.udec.cl/~gabriel/tutoriales/curso/>  
“Introduction to Satellital Oceanography”
- [5] [http://www-etsi2.ugr.es/depar/ccia/rf/www/tema1\\_00-01\\_www/node5.html](http://www-etsi2.ugr.es/depar/ccia/rf/www/tema1_00-01_www/node5.html)  
“Introducción al Reconocimiento de Formas”

***Principales satélites de percepción remota.***

***ENVISAT.***

- [7] <http://atc2.aut.uah.es/~mprieto/asignaturas/taes/trabajos/envisat/envisat.pdf>  
“ENVISAT 1 El vigilante de la tierra”
- [8] <http://www.ealuro.com/envisat.html>  
El ENVISAT

***ERS1-2.***

- [9] <http://com.uvigo.es/asignaturas/scvs/trabajos/curso9900/Teledeteccion/>  
“Teledetección espacial”.
- [10] [www.unex.es/geodesia/fotod/FDTema\\_03.pdf](http://www.unex.es/geodesia/fotod/FDTema_03.pdf)  
“Fotogrametría satelital ”.
- “Sistemas de adquisición de imágenes digitales”.
- [11] <http://earth.esa.int/services/pg/spgers1.xml>  
“European Remote Sensing Satellite-1”.
- [12] <http://earth.esa.int/ers/instruments/>  
“Los instrumentos del ERS”.
- [13] <http://flash.lakeheadu.ca/~tsapic/Satellites.htm>  
“Remote sensing satellites”.

***LANDSAT.***

- [14] <http://www.bgeo.com.py/landsat5.html> (imagenes)  
“LANDSAT”.
- “Ficha técnica resumida: LANDSAT 5 TM”.
- [15] <http://www.aeroterra.com/p-landsat.htm>  
“Aeroterra”.
- [16] <http://www.cartesia.org/data/apuntes/teledeteccion/landsat-analisis-visual.pdf>  
“El satellite LANDSAT, análisis visual de imágenes ”.
- [17] [www.ltid.inpe.br/selper/image/espanhol/landsat2.html](http://www.ltid.inpe.br/selper/image/espanhol/landsat2.html)  
“Satélites LANDSAT”.

## BIBLIOGRAFÍA.

---

[18] <http://flash.lakeheadu.ca/~tsapic/Satellites.htm>  
“Remote sensing satellites”.

### ***SPOT.***

[19] <http://flash.lakeheadu.ca/~tsapic/Satellites.htm>  
“Remote sensing satellites”.

[20] [http://www.econosur.com/notas/spot\\_5.htm](http://www.econosur.com/notas/spot_5.htm)  
“SPOT 5: el 200º satélite puesto en órbita desde Kourou”.

### ***IKONOS.***

[21] <http://flash.lakeheadu.ca/~tsapic/Satellites.htm#IKONOS%201>  
“Remote sensing satellites”.

[22] <http://www.bgeo.com.py/ikonos.html> (imagenes)  
“Ficha técnica resumida: IKONOS ”.

### ***Segmentación bayesiana.***

[23] <http://dns1.mor.itesm.mx/~emorales/Cursos/KDD03/node41.html>  
“Aprendizaje bayesiano”.

[24] [http://www.itch.edu.mx/academic/industrial/sabaticorita/\\_private/08Teorema%20de%20bayes.htm](http://www.itch.edu.mx/academic/industrial/sabaticorita/_private/08Teorema%20de%20bayes.htm)

“Teorema de Bayes”.

[25] Notas de Análisis de Señales Aleatorias Smestre 2003-1 Dr. Miguel Moctezuma Flores.

### ***Segmentación por campos de Markov.***

[26] <http://revc.uab.es/revista/04/0401-tut.pdf>  
“Aplicación de los modelos de campos aleatorios de Markov en vision artificial”

[27] <http://www.mor.itesm.mx/~esucar/Clases-incer/clase06.hmm.ps>  
[www.cs.umn.edu/~shekhar/research/mrf.pdf](http://www.cs.umn.edu/~shekhar/research/mrf.pdf)

“Image Processing and Computer Vision”.

### ***Metodos de relajacion.***

[29] <http://www.ma1.upc.edu/~susin/contingut/files/ActCubeReport.pdf>  
“Universidad Politécnica de Cataluña”.

“Trabajo de investigación: Contornos Activos, Mallas Activas y Cubos Activos”.  
(estocástica)

[30] <http://translate.google.com/translate?hl=es&sl=en&u=http://www.bmva.ac.uk/bmvc/1997/papers/009/paper009.html&prev=/search%3Fq%3Dprobabilistic%2Brelaxation%26hl%3Des%26lr%3D>

“Relajación Probabilística de niveles múltiples”.

“Universidad de Surrey”.

“Reino Unido”.

[31] [http://translate.google.com/translate?hl=es&sl=en&u=http://www.cs.cf.ac.uk/Dave/Vision\\_lecture/node43.html&prev=/search%3Fq%3Dprobabilistic%2Brelaxation%26hl%3Des%26lr%3D](http://translate.google.com/translate?hl=es&sl=en&u=http://www.cs.cf.ac.uk/Dave/Vision_lecture/node43.html&prev=/search%3Fq%3Dprobabilistic%2Brelaxation%26hl%3Des%26lr%3D)

“Etiquetado de la relajación.”  
-“Técnicas Estadísticas De la Relajación”.

***Artículos [art].***

- [1] Automatic segmentation of biomedical images.  
P.W. Fung, B.E., K.K.Ly, BE., Y. attikiouzel B. Sc., Ph. D.  
Department of Electrical & Electronic Engineering  
The University of Western Australia, Nedlands, Western Australia, 6009  
1988 IEEE.
- [2] Remote sensing images segmentation a mixed approach adapted for fine structure detection.  
J. A Franco, M. Moctezuma and F. Parmiggiani.  
Depfi-Unam.  
Septiembre 2000.
- [3] Adaptive segmentation on synthetic aperture radar images.  
S. I. Zavala, R. Cordero, J. I. Ferrara, M. Moctezuma and F. Parmiggiani.  
Mayo 2000.
- [4] Aproximación contextual a la segmentación de fotografía aérea.  
J. A Franco, M. Moctezuma.  
Depfi-Unam.  
1998.
- [5] Image Reconstruction by stochastic relaxation.  
Fathelalem F. Ali, Sachiko Yokegawa, Zensho Nakao, & Yen-Wei Chen.  
Dept. of EEE, University of the Ryukyus  
Okinawa Japan.  
1999 IEEE.
- [6] Coastline detection from SAR images.  
D.B. Zhang, L. Van & A. Osterlinck  
Katholieke University it Leuven.  
1994.
- [7] Information Fusion for Rural Land-Use Classification whit High-Resolution Satellite Imagery.  
Wanxiao Sun, Volker Heidt, Peng Gong, and Gang Xu.
- [8] Restoration of severely blurred high range images using stochastic and deterministic relaxation algorithms in compound Gauss-Markov random fields.  
Rafael Molina, Aggelos K. Katsaggelos, Jvier Mateos, Aurora Hermoso, C. Andrew Segall.  
Departemento de ciencias de la computación e I. A. Universidad de Granada España.

## **BIBLIOGRAFÍA.**

---

Marzo 1999.

[9] SAR Images Filtering and Segmentation: A multiresolution and Contextual Approach.  
J. A Franco, M. Moctezuma, M. E. Barrilla, B. Escalante and F. Parmiggiani.  
Depfi-Unam.  
20001.

[10] Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images.  
Stuart Geman and Donald Geman.  
Noviembre 1984.