



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

ESQUEMA MARKOVIANO PARA LA DETECCIÓN REMOTA DE DERRAMES PETROLEROS

T E S I S

QUE PARA OBTENER EL TÍTULO DE
INGENIERO EN TELECOMUNICACIONES

P R E S E N T A :

LUDWIN VENTURA LÓPEZ LÓPEZ

DIRECTOR DE TESIS:
DR. MIGUEL MOCTEZUMA FLORES





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 Dios, por el regalo de estar aquí y ahora.

A la familia Bohórquez por darme su mano durante este camino de cinco años. Gracias siempre.

A mi tía Agustina por su cariño. Infancia es destino.

Al triedro de ingenieros que fueron mis maestros. Gracias donde quiera que estén.

Dr. Miguel Moctezuma Flores gracias por su gran apoyo y guía invaluable.

Gracias al Ing. Alfredo Arenas por ser un excelente tutor y amigo durante todo este tiempo.

A mis amigos.)la vida se escribe entre paréntesis(.

Gracias UNAM por darme tanto desde el primer día de clases.

POR MI RAZA HABLARÁ EL ESPÍRITU

Para Fide, Xóchitl,
Hugo y Huguito.

ÍNDICE

AGRADECIMIENTOS	i
1. INTRODUCCIÓN	
1.1 OBJETIVO	1
1.1.1 <i>Fundamentación</i>	1
1.1.2 <i>Definición del problema</i>	2
1.1.3 <i>Aportación de la tesis</i>	3
2. FUNDAMENTOS DE PERCEPCIÓN REMOTA	
2.1 DESCRIPCIÓN GENERAL	5
2.2 PRINCIPIOS Y FUNDAMENTOS	6
2.3 APLICACIONES	16
2.3.1 <i>Fotografía aérea</i>	17
2.3.2 <i>Imágenes multiespectrales</i>	19
2.3.3 <i>Imágenes SAR</i>	20
2.4 FUNCIONALIDAD DE LAS FUENTES DE DATOS	23
2.5 TENDENCIAS EN LA ADQUISICIÓN DE DATOS REMOTOS	24
3. SATÉLITES Y CAPTORES	
3.1 ENVISAT	25
3.2 ERS 1 Y ERS 2	26
3.3 ASTER	29
3.4 LANDSAT	29
3.5 SPOT	35
3.6 IKONOS	37

4. ANÁLISIS TEXTURAL

4.1 ALTERNATIVAS EN EL ANÁLISIS DE TEXTURAS	39
4.1.1 <i>Aproximación estadística</i>	40
4.1.2 <i>Aproximación estructural</i>	48
4.1.3 <i>Aproximación espectral</i>	49
4.2 MATRIZ DE CO-OCURRENCIA DE NIVELES DE GRIS (GLCM)	52
4.2.1 <i>Descriptoros de textura</i>	54
4.2.2 <i>Parámetros en el cálculo de la GLCM</i>	56
4.3 ALGORITMOS RÁPIDOS PARA EL CÁLCULO DE LA GLCM	56
4.3.1 <i>Grey Level Co-occurrence Linked List (GLCLL)</i>	57
4.3.2 <i>Grey Level Co-occurrence Hibrid Structure (GLCHS)</i>	58

5. FUNCIONES DISCRIMINANTES ESTADÍSTICAS

5.1 GENERALIDADES	59
5.2 TEORÍA DE BAYES	61
5.3 FUNCIONES DISCRIMINANTES	64
5.3.1 <i>Función pérdida</i>	65
5.3.2 <i>Función discriminante de Bayes</i>	67
5.3.3 <i>Decisión de máxima probabilidad</i>	67
5.3.4 <i>Función discriminante óptima</i>	69
5.4 CASO BINARIO DE SEGMENTACIÓN	77
5.5 APROXIMACIÓN MÁXIMO A POSTERIORI- (MAP)	79
5.6 CASO APARTE: ALGORITMO DE K-MEDIAS	80

6. CAMPOS ALEATORIOS DE MARKOV

6.1 GENERALIDADES	83
6.2 TOPOLOGÍAS	86
6.2.1 <i>Vecindades</i>	86
6.2.2 <i>Cliques</i>	89
6.3 PROPIEDADES DE MARKOV	96
6.4 MODELOS DE MARKOV	98
6.4.1 <i>Modelo de Ising</i>	98
6.4.2 <i>Modelo de Potts</i>	99
6.5 OPTIMIZACIÓN	101
6.5.1 <i>Recocido Simulado</i>	101
6.5.2 <i>Esquemas de decremento de temperatura</i>	105

7. FUSIÓN DE DATOS

7.1 GENERALIDADES	107
7.2 ESQUEMA DE FUSIÓN SHAN-YU	113

8. ESQUEMA DE PROCESAMIENTO: APLICACIÓN EN LA DETECCIÓN REMOTA DE DERRAMES PETROLEROS DEL BUQUE "PRESTIGE"

8.1 ESQUEMA GENERAL	117
8.1.1 <i>Selección de sitios</i>	118
8.2 ANÁLISIS DE TEXTURA	122
8.3 SEGMENTACIÓN MARKOVIANA	134
8.3.1 <i>Segmentación imágenes SAR</i>	143
8.3.2 <i>Segmentación descriptores de textura</i>	146
8.4 FUSIÓN DE DATOS	148
8.4.1 <i>Fusión de Shan-Yu</i>	149
8.4.2 <i>Propuesta de Fusión</i>	155
8.5 EVALUACIÓN DE RESULTADOS	158

9. CONCLUSIONES	163
10. ANEXO	167
11. REFERENCIAS	171

1

INTRODUCCIÓN

1.1 OBJETIVO

El objetivo de la presente tesis es implementar un esquema de segmentación de imágenes basado en la teoría de Campos Aleatorios de Markov, realizando análisis estadístico de textura y fusionando las imágenes resultado de la segmentación con el propósito de incorporar la información en una sola entidad.

Este esquema es aplicado para la detección de derrames de petróleo en imágenes SAR de percepción remota. El resultado que se obtiene es binario y muestra en detalle la mancha de petróleo en la zona de análisis.

1.1.1 *Fundamentación*

Una característica de las imágenes adquiridas con el Radar de Apertura Sintética, es la naturaleza ruidosa. Aún cuando en la imagen no se cuenta con un contraste adecuado, se puede apreciar cierta característica textural. La textura es inherente a las superficies de los objetos que se observan en las imágenes de percepción remota y por lo tanto puede servir para discriminar entre objetos presentes de una misma escena.

Si se analiza una textura basándose únicamente en su histograma se presenta la desventaja de que no existe información sobre la ubicación espacial que guardan entre sí los pixels con determinado nivel de gris. La Matriz de Co-ocurrencia de Niveles de Gris (GLCM) es una matriz que salva este obstáculo. La GLCM se calcula en una ventana, a una distancia interpixel y orientación dada, proporcionando la probabilidad de ocurrencia de cualquier par de pixels con determinados niveles de gris.

Por otra parte, el modelo de Campo Aleatorio de Markov (CAM) incorpora una gran capacidad en el análisis para la segmentación de imágenes. El uso práctico de un modelo de Campo Aleatorio de Markov se atribuye a la equivalencia entre la distribución de probabilidad de un Campo Aleatorio de Markov y la distribución de probabilidad de Gibbs.

En el presente esquema de procesamiento, la pertenencia a una determinada clase está dada por la probabilidad Máxima A Posteriori (MAP). Para encontrar la configuración óptima, se utiliza un método que minimiza globalmente la función de energía *a posteriori*, tal algoritmo recibe el nombre de *Recocido Simulado*.

Las imágenes satelitales son bastante complejas y un algoritmo único no es suficiente para el análisis de éstas. Si una imagen es procesada con diferentes algoritmos, las imágenes resultado presentaran información complementaria entre sí, por lo que es necesario incorporar esta información para brindar un resultado mas completo.

Bajo esta consideración y como parte final del esquema de procesamiento, la fusión de datos se plantea como la integración de información presente en las diferentes imágenes procesadas con diferentes algoritmos. Esta integración brinda como resultado una unidad específica y comprensiva de datos con respecto a la entidad que se estudia.

1.1.2 Definición del problema

Para la observación de escenas de difícil acceso es necesario el uso de instrumentos embarcados en plataformas aéreas y/o satelitales, los cuales presentan todo un abanico de posibilidades para registrar datos en diferentes bandas del espectro electromagnético.

El estudio de las imágenes de percepción remota es importante, ya que el análisis de estas imágenes brindará información necesaria para su interpretación. Uno de los principales instrumentos usados en percepción remota, es el Radar de Apertura Sintética (SAR). Este instrumento presenta la ventaja de proveer su propia fuente de “iluminación”, permitiendo la adquisición de los datos bajo condiciones climatológicas de nubosidad y sin importar si es de día o de noche, su desventaja consiste en la presencia de ruido *speckle*, propio de sistemas (láser, radar y ultrasonido) que utilizan fuentes de iluminación coherente.

Las imágenes SAR se caracterizan por la variedad de texturas presentes en la escena. El estudio de las imágenes SAR con algoritmos de análisis de textura es indispensable para la discriminación de las diferentes clases que se encuentran presentes en la imagen. Existen varios algoritmos para el análisis de texturas. Entre estos algoritmos se encuentra la Matriz de Co-ocurrencia de Niveles de Gris y los descriptores de textura ligados a ésta, los cuales han mostrado ser una herramienta robusta para la discriminación de clases con texturas características.

En el presente esquema, el problema de etiquetado entre clases $\{\text{petróleo}, \text{nopetróleo}\}$ de los pixels de las imágenes SAR, se realiza como una segmentación basada en Campos Aleatorios de Markov y que se formulan en el marco Bayesiano. En este caso, la solución óptima es definida por el etiquetado MAP y se calcula minimizando la *energía a posteriori* por medio de un algoritmo de minimización global como lo es el Recocido Simulado.

La *probabilidad a posteriori* se define con la regla de Bayes, a partir de un modelo *a priori* y un modelo de probabilidad. El modelo de probabilidad se relaciona de cómo son observados los datos y es un problema dependiente del dominio. El modelo *a priori* depende de cómo se expresan ciertas restricciones *a priori*.

Los resultados de la teoría de Campos Aleatorios de Markov proporciona herramientas para codificar restricciones contextuales en la probabilidad *a priori*.

Del uso común que se hace de los Campos Aleatorios de Markov, la definición y solución de un modelo de Campo Aleatorio de Markov se resume en los siguientes pasos:

1. Plantear el problema de segmentación como un problema de etiquetado bayesiano en el que la solución óptima se define como la configuración de etiquetas MAP.
2. Usar la función de energía de Gibbs para caracterizar la distribución *a priori* de las configuraciones del etiquetado.
3. Caracterizar la función de energía conjunta.
4. Definir la función de energía *a posteriori*.
5. Usar un algoritmo (Recocido Simulado) para minimizar la función de energía *a posteriori* de la distribución de Gibbs y obtener así la configuración óptima de etiquetas MAP.

Para incorporar la información de las imágenes procesadas con los diferentes algoritmos es necesario la utilización de un algoritmo de fusión.

Las imágenes procesadas bajo diferentes algoritmos contienen información redundante y complementaria entre sí. Esta consideración es el punto de inicio para realizar la fusión. Para obtener un buen resultado, el algoritmo de fusión debe considerar dos elementos: las etiquetas presentes en la vecindad del *site* analizado de una sola imagen y las etiquetas presentes en las vecindades del *site* analizado, a través de todas las imágenes que sirven de entrada al algoritmo de fusión.

1.1.3 Aportación de la tesis

Los puntos que aporta esta tesis son:

1. Estudio sobre técnicas de procesamiento de imágenes y su aplicación a imágenes SAR.
2. Análisis de textura en imágenes SAR.
3. Estudio sobre segmentación por Campos Aleatorios de Markov.
4. Estudio sobre fusión binaria.
5. Aplicación del esquema markoviano propuesto para la detección de derrames de petróleo del buque “Prestige”.

2

FUNDAMENTOS DE PERCEPCIÓN REMOTA

2.1 DESCRIPCIÓN GENERAL

La percepción remota es la disciplina que se encarga de adquirir, procesar e interpretar imágenes y datos relacionados entre sí, obtenidos desde aviones o satélites, registrando la interacción entre materia y radiación electromagnética.

La percepción remota implica la medición de la energía que es reflejada o emitida por objetos, sin entrar en contacto con ellos. Esta energía electromagnética, viaja por el espacio como ondas electromagnéticas las cuales se diferencian entre sí por su longitud de onda.

Un sistema de percepción remota [CHUV90], en general, consta de los siguientes elementos:

a)Fuente de energía:

Supone el origen del flujo energético detectado por el sistema sensor. Puede tratarse de una fuente externa al sensor, en cuyo caso se refiere a percepción remota pasiva, o de un haz energético emitido por éste, siendo el caso de la percepción remota activa.

b)Superficie terrestre:

Formada por distintas masas de vegetación, suelo, agua, o construcciones humanas, que reciben la señal energética procedente de la fuente de energía.

c)Atmósfera

La atmósfera se comporta como un filtro para diversas longitudes de onda, por lo cual, puede afectar lo resultados, siendo necesario tomar en cuenta este efecto para considerar el tipo de sensor y longitud de onda a detectar.

d)Sistema sensor:

Compuesto por el sensor, propiamente dicho, y la plataforma que los soporta. Tiene como objetivos el captar la energía procedente de la cubierta terrestre, codificarla, grabarla o enviarla directamente al sistema de recepción.

e) Sistema de recepción y comercialización:

Aquí se recibe la información transmitida por la plataforma, se graba en un formato apropiado y, tras las oportunas correcciones se distribuye a quien realizará la interpretación.

f) Intérprete:

Es quien analiza la información, convirtiéndola en una clave temática o cuantitativa, orientada a facilitar la evaluación del problema de estudio.

g) Usuario final:

Encargado de analizar el documento fruto de la interpretación, así como de dictaminar sobre las consecuencias que de él deriven.

Las distancias que separan al sensor del objeto y el área de estudio varían desde unos pocos metros hasta miles de kilómetros. Los métodos más comunes para recoger información incluyen el uso de sensores colocados sobre aviones o sobre satélites.

Las técnicas de percepción remota son usadas ampliamente para recolectar información de zonas sobre la superficie terrestre. Los datos percibidos remotamente tienen muchos usos que van desde la vigilancia militar, planificación del uso de la tierra, hasta el estudio del suelo agrícola y cultivos.

2.2 PRINCIPIOS Y FUNDAMENTOS

Anteriormente se definió la percepción remota como aquella técnica que permite obtener información a distancia de los objetos situados sobre la superficie terrestre. Para que esta observación remota sea posible, es preciso que entre los objetos y el sensor exista algún tipo de interacción. Por ejemplo, nuestros sentidos perciben un objeto sólo cuando son capaces de interpretar la información que procede del mismo.

Dentro de los diferentes elementos que conforman un sistema de percepción remota vistos anteriormente, los tres principales son: sensor, objeto observado y la fuente de energía que permite poner a los primeros dos en relación mutua.

El flujo energético entre el objeto de estudio y el sensor constituye una forma de radiación electromagnética.

Las propiedades de la radiación electromagnética se han explicado por dos teorías: aquella que la concibe como un haz ondulatorio (Maxwell) y aquella otra que la considera como una sucesión de unidades discretas de energía, fotones o cuantos, con masa igual a cero (Planck).

Según la teoría ondulatoria, la energía electromagnética se transmite de un lugar a otro siguiendo un modelo armónico y continuo, a la velocidad de la luz, conteniendo dos campos ortogonales

entre sí: campo eléctrico y campo magnético. Las características de este flujo energético puede describirse por dos elementos: longitud de onda (λ) y frecuencia (f). La primera hace referencia a la distancia entre dos picos sucesivos de una onda mientras que la frecuencia designa el número de ciclos pasando por un determinado punto en una unidad de tiempo, estando relacionados de la siguiente manera:

$$c = \lambda f$$

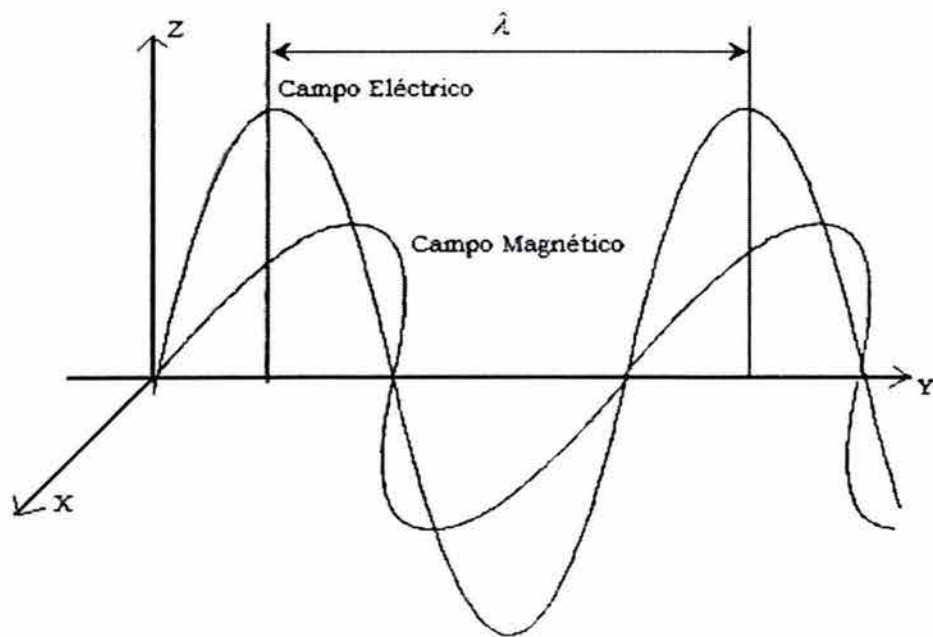
ec. 2.1

c : velocidad de la luz en el vacío (3×10^8 [m/s])

λ : longitud de onda [m]

f : frecuencia [Hz]

Modelo de propagación armónico y continuo



a)

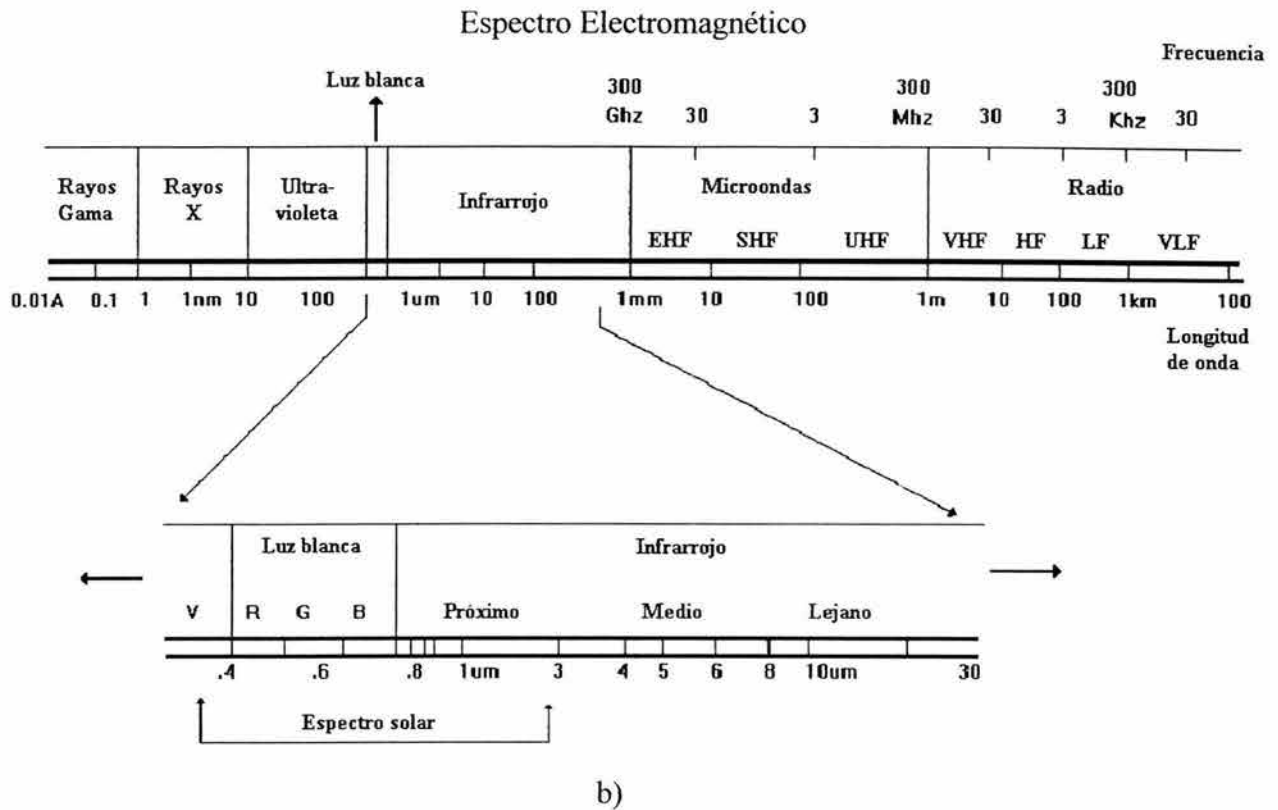


fig. 2.1

A mayor longitud de onda, menor frecuencia y viceversa, por lo que basta con un solo término para identificar propiamente el tipo de energía mencionada.

De la fórmula anterior, se deduce que podemos describir cualquier tipo de energía radiante en función de su longitud de onda o frecuencia. Aunque la sucesión de valores de la longitud de onda es continua, se suele establecer una serie de bandas en donde la radiación electromagnética manifiesta un comportamiento similar. La organización de estas longitudes de onda o frecuencia se denomina espectro electromagnético.

El espectro electromagnético consiste en todas las frecuencias de las ondas electromagnéticas. El espectro continuo se subdivide en algunos tipos familiares de energía electromagnética (fig. 2.1-b). Desde el punto de vista de la percepción remota, conviene destacar algunas bandas espectrales, que son las usadas comúnmente, que se dividen de la siguiente forma:

Ondas de radio: se caracteriza por las bajas frecuencias y grandes longitudes de onda. Las ondas electromagnéticas en esta porción 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.

Microondas: se sitúan en la franja de 1 [mm] a 30 [cm]. En esta porción de longitudes de onda se puede 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.

Infrarrojo: de gran importancia para la percepción remota. Engloba la radiación con longitudes de onda de 0,75 [μm] a 1,0 [mm]. La radiación infrarroja es fácilmente absorbida por la mayoría de las sustancias (efecto de calentamiento).

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 (380 a 750 [nm]), la cual es importante en la percepción remota, ya que las imágenes obtenidas en esta franja, generalmente presentan una excelente correlación con la experiencia visual del intérprete.

Ultravioleta: extensa franja del espectro (10 [nm] a 400 [nm]). Las películas fotográficas son más sensibles a la radiación ultravioleta, que a la luz visible. Uno de los grandes obstáculos para la utilización de esta región del espectro, es la fuerte atenuación atmosférica.

Rayos X: franja 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.

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 franja superior de frecuencias para la radiación, conocida como rayos cósmicos.

Por otra parte, todos los objetos emiten energía electromagnética (asumiendo que tienen temperaturas mayores a cero absoluto (-273°C)), lo que significa que todo emite radiación). Sin embargo, esta radiación se emite en longitudes de onda que están usualmente fuera de la banda del espectro visible. El ojo humano no puede percibir esta energía.

Cuando la energía electromagnética, como la proveniente del sol o una fuente diferente, pega en un objeto, tres diferentes fenómenos se pueden presentar:

1. *Reflexión.*
2. *Transmisión.*
3. *Absorción.*

En realidad, cuando esto sucede, más de uno de estos tres fenómenos se presenta al mismo tiempo. Algunas longitudes de onda pueden ser reflejadas, transmitidas y otras absorbidas. La energía absorbida es convertida en calor y emitida a una longitud de onda diferente.

La manera en que afecta un objeto cada longitud de onda que recibe depende de sus características y de los ángulos en que la luz le incide. La energía reflejada y emitida de los objetos es lo que se mide en la percepción remota.

Cada objeto o grupo de objetos, refleja un único espectro de longitudes de onda. Las diferencias entre el espectro reflejado es la base para distinguir un objeto de otro.

Los sistemas de percepción remota se sustentan en la energía reflejada o emitida por los objetos, es interesante definir el concepto de *reflectividad o reflectancia* que equivaldría al porcentaje de energía reflejada respecto a la energía incidente. Sin embargo, la reflexión de los objetos depende, en gran medida, de su naturaleza interna y del grado de rugosidad de la superficie, pudiéndose diferenciar dos tipos de reflexión extremos:

Reflexión especular. Se produce en superficies lisas y planas (donde las variaciones de altura de la rugosidad superficial sean menores que la longitud de onda). En estos casos el ángulo de reflexión es igual al de incidencia, pudiendo llegar a ser muy molesta en la percepción remota al saturar el detector.

Reflexión difusa: Cuando las variaciones de la rugosidad superficial son mayores que la longitud de onda, la energía es reflejada más o menos uniformemente en todas las direcciones.

Cuando la superficie refleja la energía uniformemente, con independencia del ángulo de incidencia recibe el nombre de "superficie Lambertiana" (arena, nieve...), dependiendo la reflectancia sólo de la longitud de onda. Sin embargo, lo predominante en la naturaleza es una mezcla de ambos tipos, originándose grandes diferencias de reflectancia entre los objetos debido a la variabilidad de las texturas superficiales existentes y a la movilidad con que el relieve modifica el ángulo de incidencia. Por otra parte, al mismo tiempo, cada uno de los objetos de la superficie terrestre emiten energía electromagnética (infrarrojo térmico) con una intensidad que está en relación con su temperatura, por lo que también puede ser utilizada para su identificación.

Todo ello nos lleva al concepto de *emisividad*, que es la proporción de energía emitida en relación con la energía incidente o, lo que es igual, la proporción entre la energía emitida por un objeto y la que emitiría un perfecto emisor, un "cuerpo negro", a su misma temperatura.

Una gran parte de la transmisión de energía se produce a través del cuerpo gaseoso que rodea nuestro planeta: la atmósfera. Sin embargo, en el tránsito de la radiación hacia el sensor, la atmósfera no se comporta como un cuerpo transparente, sino que impone un conjunto de modificaciones que es necesario conocer al utilizar cualquier sistema de percepción remota. Los efectos son muy variados, destacando por su influencia en la percepción remota las modificaciones en la intensidad, frecuencia y distribución espectral de la radiación impuestas por la dispersión y la absorción atmosféricas. Ambas dependen de la altura de la capa atmosférica que debe atravesar la radiación (que en el caso de los satélites es toda) antes de ser captada por el sensor, de su propia composición, del contenido en partículas en suspensión y de las longitudes de onda en las que opere el sensor.

En lo referente a la *dispersión atmosférica* (conocida con el término inglés "scattering"), este fenómeno aparece a causa de la difusión en todas las direcciones que las partículas en suspensión y las moléculas de los componentes de la atmósfera, producen sobre la radiación que interacciona con ellos. De forma general, este efecto de dispersión decrece a medida que aumenta la longitud de onda, aunque pueden distinguirse varios tipos:

Dispersión "Rayleigh": Este tipo de dispersión es producida por las moléculas y partículas atmosféricas que tienen un diámetro menor que la longitud de onda incidente, afectando especialmente a las longitudes de onda más cortas de espectro visible. Esta es la causa del "azul" del cielo en los días despejados.

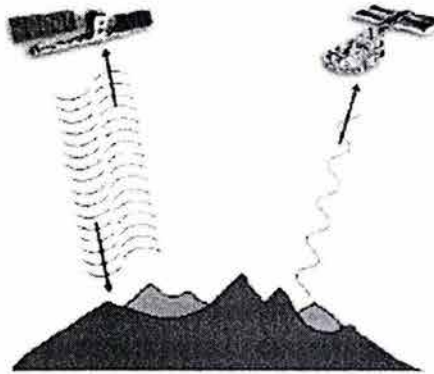
Dispersión "Mie": Este tipo se produce cuando las partículas y moléculas presentes en la atmósfera tienen un diámetro de igual tamaño que la longitud de la onda de la energía incidente, afectando singularmente a longitudes de onda más largas que la dispersión Rayleigh.

Dispersión "no selectiva": Este tipo de dispersión, independiente de la longitud de onda de la energía incidente, se produce cuando las partículas y moléculas atmosféricas presentan un diámetro de mayor tamaño que la longitud de onda de la radiación. Un caso típico es el color blanco de las nubes, ya que los elementos y partículas que las forman reflejan con igual intensidad todas las longitudes de onda dentro del visible (azul, verde, rojo).

Por su parte, la *absorción atmosférica* está ligada a la retención de energía por parte de los diferentes componentes de la atmósfera, sobre todo del vapor de agua (H_2O), el dióxido de carbono (CO_2), el oxígeno (O_2) y el ozono (O_3), aunque este fenómeno de la absorción es muy selectivo, limitándose a determinadas longitudes de onda dentro del espectro electromagnético.

Las regiones del espectro donde estos procesos son débiles, y por lo tanto la transmisión de energía se produce con facilidad, reciben el nombre de *ventanas espectrales* y se centran especialmente en el ultravioleta, el visible y el infrarrojo reflejado y térmico, además de las longitudes de onda superiores a 1 [mm], donde la transmisión atmosférica es total, siendo ésta una de las principales razones de su frecuente uso en la percepción remota (radar). Lógicamente, los principales sensores utilizados en los sistemas de percepción remota se centran, por razones obvias, en las regiones de las ventanas espectrales y en las longitudes de onda donde la emisión es mayor.

La temperatura también puede afectar el espectro de energía reflejada o emitida por los objetos. Todos los objetos físicos que tienen una temperatura superior a cero absoluto emiten radiación electromagnética. Objetos con mayor temperatura emiten la mayor parte de su radiación con menores longitudes de onda, en cambio objetos más fríos emiten con mayores longitudes de onda. La mayor parte de la energía emitida por objetos debida al calor está en la banda del infrarrojo (IR). Por eso, esta banda de energía se conoce normalmente como banda termal. Esta es la razón por la que las cámaras infrarrojas y los sensores infrarrojos de temperatura pueden medir el calor emitido por los objetos, y de esta forma estimar su temperatura.



sistemas de percepción activa y pasiva respectivamente

figura 2.2

En general, los sistemas de percepción remota se pueden clasificar en dos tipos:

1.-*Los sistemas de percepción activa*

Son aquellos que generan una señal, rebota en un objeto, y miden las características de la señal reflejada. Un ejemplo de este tipo de percepción es el radar. Las ondas de radio son emitidas por un transmisor y son recogidas por un receptor después de haber sido reflejadas por un objeto. Las señales reflejadas pueden ser usadas por los sistemas de radar para determinar la distancia y dirección de los objetos percibidos. También se pueden crear imágenes de los objetos.

Los sistemas activos, como el radar, proveen su propia fuente de energía, que emiten hacia los cuerpos, y reciben la señal de retorno, por ello, estos sistemas pueden trabajar bajo cualquier condición atmosférica, tanto de día como de noche.

2.- *Sistemas de percepción pasiva*

Estos sistemas, simplemente reciben las señales emitidas y reflejadas por los objetos percibidos, es decir, se limita a recoger la energía electromagnética procedente de la cubierta terrestre, ya sea reflejada de los rayos solares, o emitida en virtud de su propia temperatura.

Puede establecerse una división en estos sensores en función del procedimiento que emplean para recibir la radiación procedente de los objetos, a saber quedaría:

1. Captores instantáneos de tipo fotográfico: en ellos, la visión analógica del terreno se produce a través de la utilización de emulsiones capaces de registrar la energía electromagnética reflejada en el espectro visible e infrarrojo próximo. Se les puede acoplar dispositivos de barrido electrónico con digitalización, facilitándose así la transmisión de los datos.
2. Sensores de barrido multiespectral (scanners): en este caso, el sensor "barre" o inspecciona secuencialmente sucesivas partes de terreno (pixel) para construir las líneas

que constituyen la imagen. La radiación captada en cada una de ellas pasa por una serie de filtros (espectrómetro) individualizando las longitudes de onda seleccionadas (generalmente dentro del visible e infrarrojo reflejado y térmico) y las dirige hacia los correspondientes detectores, los cuales emiten una señal eléctrica proporcional a la intensidad de la energía recibida para cada una de ellas, cuyo valor es almacenado o transmitido. Existen varios tipos:

a) De espejo oscilante ("Wiskbroom"): aquí es un espejo oscilante con una velocidad de rotación adaptada a la velocidad de la plataforma espacial quien barre secuencialmente la escena observada a través de líneas perpendiculares a la dirección de la órbita. Un ejemplo bastante conocido es el Multispectral Scanner (MSS) embarcado en los satélites LANDSAT.

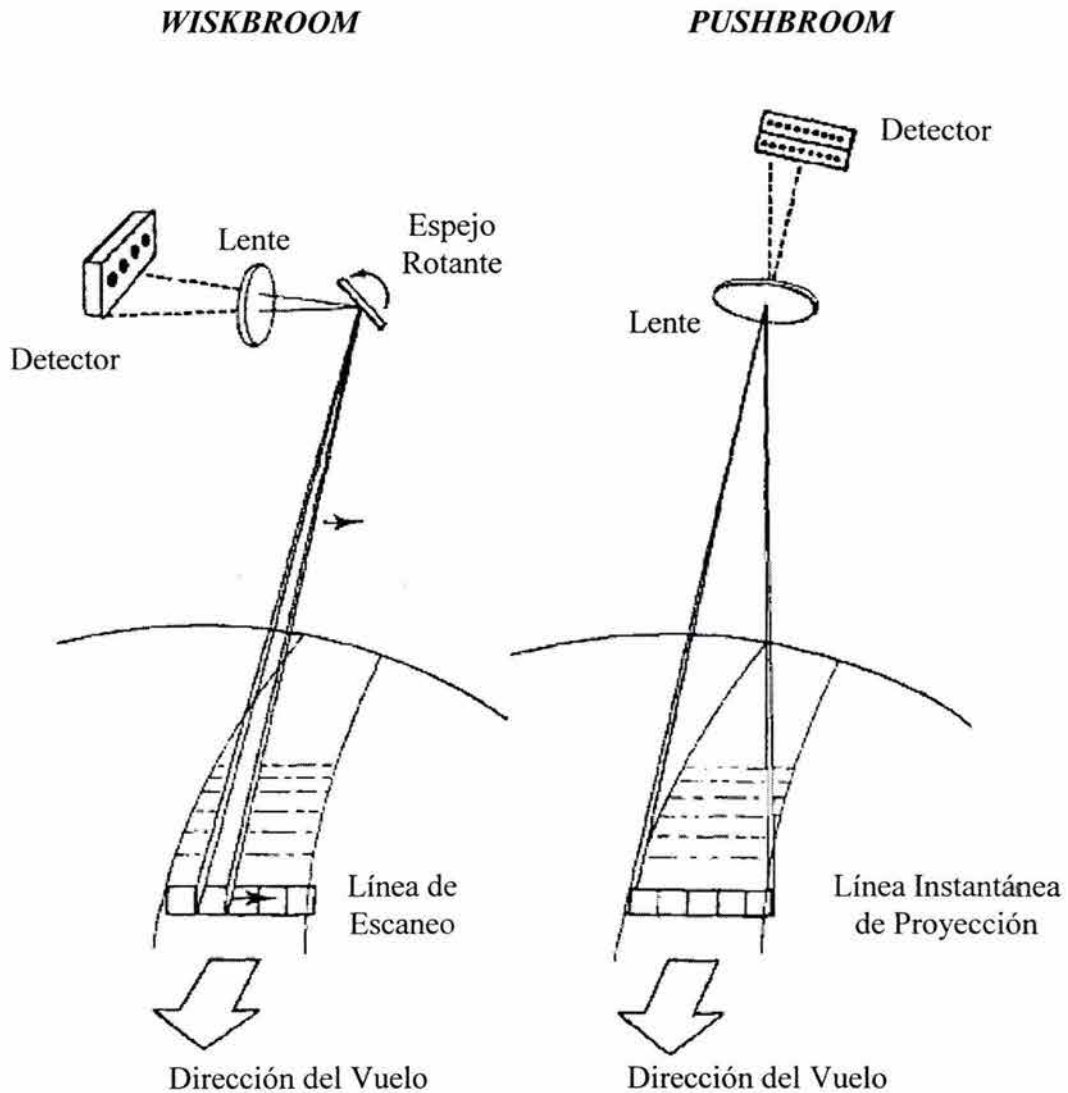
En otros casos, cuando el satélite permanece fijo sobre un punto concreto de la superficie terrestre, el barrido de la imagen se consigue por el giro del satélite sobre sí mismo. (METEOSAT).

b) Recientemente se han construido sensores de "matriz lineal", donde un conjunto importante de detectores sólidos captan, de una sola vez, cada una de las líneas de la imagen ("Pushbroom"), con lo que se eliminan problemas de calibración, inducidos por el giro inherente a los sensores de espejo oscilante. Este sistema equipa al sensor HRV del satélite SPOT.

La energía recibida por el sensor, depende de varios factores como lo son: la reflectividad de la cubierta terrestre, las condiciones atmosféricas, el emplazamiento ambiental y la geometría de la observación.

Pueden detectarse algunos rasgos geográficos como ríos, o carreteras de anchura inferior a la resolución espacial del sensor, siempre y cuando exista suficiente contraste radiométrico entre la señal proveniente de estos rasgos y sus vecinos.

Otro punto importante, es la resolución de un sensor. La resolución de un sensor es la capacidad para discriminar la información en detalle. Depende de la cadencia temporal, de la capacidad para distinguir variaciones en la energía que detecta, del detalle espacial del sensor, y del número y anchura de las bandas espectrales que es capaz de adquirir.



tipos de sensores de barrido multispectral

fig. 2.3

Existen varios tipos de resolución:

1.-Resolución Espacial:

Es el objeto mas pequeño que puede ser distinguido sobre la imagen. Define el tamaño del pixel, que es la distancia correspondiente al tamaño de la mínima unidad de información en la imagen. En sensores de antena, la resolución depende de la apertura, altura de la plataforma, y la longitud de la onda electromagnética. A mayor radio, menor altitud y menor longitud de onda, la resolución es mas detallada. Solo son discriminables en la imagen los elementos de tamaño superior a la resolución espacial (es decir, al tamaño del pixel).

2.-Resolución Espectral:

Es el número y anchura de las bandas espectrales que puede discriminar el sensor. El radar presenta la menor resolución espectral, ya que trabaja con un solo canal.

3.-Resolución Radiométrica

Es la sensibilidad del sensor, es decir, la capacidad para detectar variaciones en la radiancia espectral que recibe. Determina el número de niveles de gris adquiridos por el sensor, se expresa en niveles por píxel (64, 128, 256 ó 1024). A mayor resolución radiométrica, mejor interpretación de la imagen. El número de niveles más adecuado, depende del método que se siga en la interpretación, el ojo humano no puede detectar más de 64 niveles de gris.

4.-Resolución Temporal

Es la frecuencia de cobertura que proporciona el sensor, es decir, la periodicidad con que éste adquiere imágenes de la misma porción de la superficie terrestre. La cual está en función de las características orbitales de la plataforma (altura, velocidad e inclinación) y del diseño del sensor (ángulo de observación y ángulo de cobertura). A mayor resolución espacial, menor resolución temporal, menor resolución espectral y menor resolución radiométrica.

El sensor convierte una señal analógica (la radiancia recibida), en un valor digital por ello no se habla de fotografías, sino de imágenes digitales, en razón a su estructura interna. Estos valores se traducen de nuevo en niveles de radiancia conociendo los coeficientes de calibrado del sensor y las condiciones de adquisición. La conversión de digital a analógica obtiene los valores de radiancia original detectados por el sensor. Por ello, no solo facilitan una imagen, sino una medición de la superficie.

Los datos obtenidos directamente por los sensores empleados en percepción remota, sólo en pocas ocasiones, pueden utilizarse en estado original. Una de las primeras operaciones realizadas en los equipos de tratamiento de imágenes, previa recalibración de los datos, es corregir un conjunto de deformaciones o de distorsiones que habitualmente presentan estos datos originales.

Una parte importante de estas distorsiones están ligadas a la alteración que impone la atmósfera a la radiación reflejada o emitida en su tránsito hacia el sensor. Las operaciones utilizadas para ello son conocidas como *correcciones atmosféricas* y su aplicación exige conocer la composición de la atmósfera en el momento de la adquisición para corregir las modificaciones inferidas, y así, poder relacionar los datos registrados con la respuesta de los objetos en la superficie terrestre, a la vez que su corrección facilita la comparación con datos del mismo espacio para otras fechas distintas. A ellas habría que unir las correcciones de las distorsiones introducidas por errores en la toma de datos debidos al funcionamiento deficiente del sensor (correcciones radiométricas).

Otro grupo importante de estas deformaciones se vincula a las distorsiones geométricas ligadas al funcionamiento del dispositivo de adquisición de los datos (scanner), a los cambios de altitud de la órbita o a la rotación relativa de la Tierra respecto a la órbita del satélite, entre ellas cabe citar la distorsión producida en la escala tangencial debido a que el dispositivo de barrido se mueve a una velocidad constante y, por lo tanto, la distancia de superficie inspeccionada por unidad de

tiempo aumenta hacia los extremos. Sin embargo, la mayor parte de las mismas se pueden corregir o minimizar simplemente programando la velocidad del satélite para compensar este defecto, o bien utilizando, como ya se hace en algunos satélites, sistemas de matriz de líneas "pushbroom" de gran simplicidad mecánica, larga duración y fidelidad geométrica muy alta.

Ligadas directamente al sistema de barrido, podemos incluir las deformaciones inducidas por el desplazamiento unidireccional o abatimiento del relieve o cualquier elemento vertical y aquellas producidas por la modificación de los parámetros de vuelo. Las primeras están directamente relacionadas con la altura del sensor, respecto al nivel de referencia y son, obviamente, importantes en los sensores aerotransportados. No obstante, al ser el abatimiento similar al producido en una fotografía aérea, en las plataformas espaciales juega un papel importante a su favor la altura de la órbita (800 a 36000 [Km]), presentando esta deformación unos valores insignificantes. En el segundo caso, las deformaciones se deben a la inestabilidad de la plataforma aérea, ligada a su irregular desplazamiento por la siempre dinámica capa atmosférica. Obviamente, la estabilidad de las plataformas espaciales alejadas de las turbulencias atmosféricas es, en este sentido, significativamente mayor que en los aviones.

Sin embargo, a pesar de las matizaciones hechas con antelación que, en cierta medida, justifican la utilización de los datos originales para muchas aplicaciones temáticas, las distorsiones antes aludidas (escala tangencial, movimientos de la plataforma, desplazamientos horizontales...), pueden ser corregidas fácilmente para obtener productos de mayor fidelidad geométrica. Este conjunto de técnicas reciben el nombre de *correcciones geométricas* y su objetivo es aplicar a los datos originales una transformación tal que el espacio geográfico resultante sea perfectamente ajustable a una cartografía general de referencia.

Generalmente el proceso se basa en la elección de un conjunto de puntos de control homólogos en la imagen original y en el documento cartográfico utilizado, a partir de los cuales se obtienen las funciones de transformación polinómicas que permiten convertir el espacio geométrico de la imagen original en otro espacio igual al de la cartografía de referencia.

Son muy numerosos los métodos de corrección geométrica, implicando siempre, una reubicación de los píxeles que componen una imagen y una modificación de su valor radiométrico influenciado por los píxeles de su entorno.

2.3 APLICACIONES

Los comienzos de la percepción remota se confunden con los de la fotografía aérea, aunque, sin lugar a dudas, los primeros vuelos espaciales facilitaron una nueva visión de la Tierra y de sus recursos naturales hasta ese momento desconocida, inicialmente en forma fotográfica y posteriormente a través de registros digitales. Estas fotografías e imágenes obtenidas desde satélites en misiones experimentales hacían vislumbrar una nueva era en cuanto a las posibilidades de reconocimiento de los recursos del planeta a escala global, a la vez que los avances en óptica y electrónica daban como resultado el desarrollo de nuevos sensores que permitían escrutar la respuesta de los diferentes elementos de la superficie terrestre en el espectro electromagnético.

Todo comenzó a finales de los años 40, cuando se obtuvieron las primeras fotografías en color desde alturas en torno a los 100 [Km]. Poco después, se pusieron en marcha programas de meteorología desde el espacio, que son los auténticos precursores de la percepción remota. Así, en 1960 se puso en marcha el programa TIROS-1 que llegaría a transmitir 23 000 imágenes de la Tierra y continuó con el programa NIMBUS que ya permitía cubrir la totalidad de la Tierra con imágenes espaciales.

En la actualidad, las aplicaciones más comunes de la percepción remota son:

- a) Inventario regional del medio ambiente para preparar estudios de impacto ambiental.
- b) Cartografía geológica para la exploración mineral y petrolífera, así como el estudio de nuevos depósitos volcánicos.
- c) Inventario de agua superficial.
- d) Cartografía térmica de la superficie del mar.
- e) Análisis de corrientes marítimas.
- f) Cartografía de la cobertura vegetal del suelo.
- g) Seguimiento del ritmo de crecimiento de la población.
- h) Selección de rutas óptimas para nuevas vías de comunicación.
- i) Predicción del rendimiento de cultivos.

La fotografía aérea, imágenes espectrales e imágenes SAR son las herramientas utilizadas para llevar cabo estas tareas.

2.3.1 Fotografía aérea

Las fotografías aéreas son normalmente tomadas de 10:00 a.m. y 3:00 p.m., ya que el sol forma un ángulo lo suficientemente alto con respecto al horizonte para que las sombras tengan una mínima extensión. Este tipo de previsión es necesaria, ya que fotografías que serán analizadas para estudiar la topografía de un lugar, no pueden tener sombras.

La fotografía aérea ha sido trabajada tradicionalmente a alturas menores de 6000 [m], resultando en escalas de 1:40 000, a partir del año de 1970, se dió el mejoramiento tecnológico de las cámaras y las películas fotográficas, mejorando la rapidez y resolución.

Por varios años, la NASA, por medio del programa NASA High-Altitude Photographs, se ha dado a la tarea de fotografiar desde una altura de 18 [Km], por medio de aviones como el U2 y el RB-57, con cámaras de 152 [mm] de longitud focal y un formato de película de 23x23 [cm²], cubriendo un área de aproximadamente 840 [Km²] a una escala de 1:120 000, como en cualquier

tipo de tarea de esta índole, se puede utilizar película B/N, color o IR, muchas misiones utilizan dos tipos de cámaras utilizando diferentes tipos de película.

Otro programa de fotografía aérea es el de la U.S. Geological Survey, llamado NHAP –National High Altitude Program-, en el cual se toman fotografías a escala de 1:80 000 para B/N y 1:58 000 para color.

Sin duda, una forma común de la percepción remota es la fotografía aérea, de la cual, podemos mencionar alguna de las ventajas que ofrece:

- 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.
- Gran flexibilidad ya que por el hecho de usar plataformas aéreas para portar 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).
- Costo relativamente bajo.

Algunas desventajas son:

- Se necesita buen clima para la toma de fotografías.
- La información es registrada de manera analógica, por lo que para su posterior procesamiento es necesario digitalizar tal información.



fotografía aérea del oriente de México, D.F.

fig. 2.4

2.3.2 Imágenes multispectrales

Una imagen multispectral es un arreglo de imágenes simultáneamente adquiridas y registradas – de un mismo objeto- a diferentes bandas espectrales. El carácter multispectral de las imágenes resulta básico de cara a su interpretación. La posibilidad de observar un fenómeno en diversas bandas del espectro amplía notablemente la capacidad de reconocerlo. Los sistemas multispectrales utilizan sistemas ópticos separados para poder capturar cada una de las imágenes a determinada banda de manera independiente.

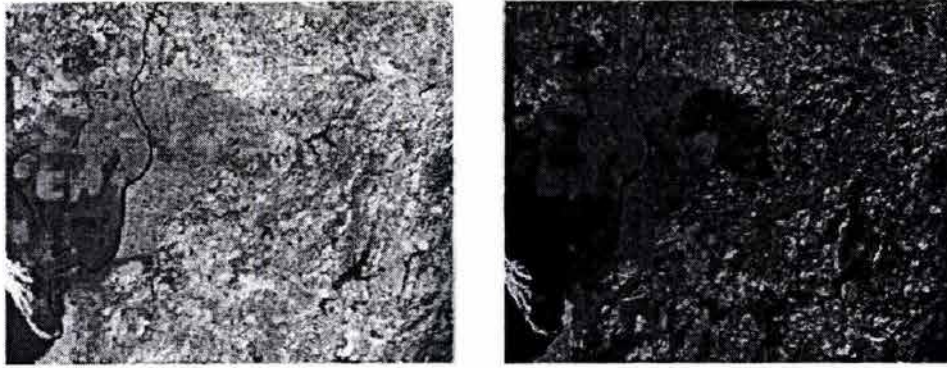
Cabe recordar que los sensores fotográficos se concentran en capturar la energía electromagnética en forma de luz que reflejan los objetos, por otra parte, los sensores termales se encargan de capturar la energía electromagnética (IR) que emiten los cuerpos. Los sistemas multispectral scanner (MSS) son capaces de realizar ambas cosas, ya que están diseñados de tal manera que pueden capturar la información proveniente de varias bandas espectrales.



banda 1



banda 3



banda 2

imagen en falso color (bandas 1, 2, 3)

imagen multiespectral del sensor TM-LANDSAT

fig. 2.5

2.3.3 Imágenes SAR

Las imágenes adquiridas por el sistema SAR son muy similares a las fotografías aéreas tradicionales aunque tienen diferencias muy significativas, por ejemplo, una imagen SAR no presenta vegetación ya que árboles y rocas reflejan las ondas radar de forma similar. En la mayoría de los casos, 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 radar están formadas a partir de una frecuencia radar simple y se presenta casi siempre en blanco y negro.

Los sensores remotos que actúan en la zona espectral de las microondas se basan en el principio del radar. El sensor provee su propia fuente de energía para “iluminar” la superficie que se estudia –sistema de percepción activa-, a diferencia de los llamados sensores pasivos que operan en el visible e infrarrojo utilizando como fuente de energía al sol.

Las características principales que se destacan en estos sistemas son:

- 1.- Las microondas tienen la capacidad de penetrar la atmósfera en diferentes condiciones meteorológicas durante todo el día, esto lo convierte en un sistema para todo tiempo.
- 2.- Para la interpretación de las imágenes de radar, se debe tener en cuenta que las propiedades reflectivas o emisivas en el rango espectral de las microondas no tienen el mismo comportamiento espectral sobre los objetos que en las regiones espectrales del visible e infrarrojo.
- 3.- Los primeros sensores de radar fueron aerotransportados y utilizaban una antena que se fijaba longitudinalmente en el avión, ejecutando un barrido lateral; a estos sistemas se los denomina SLAR (radar de apertura real de visión lateral). Se desarrollaron para propósitos de reconocimiento militar en la década del 50. Luego, al considerarse su

capacidad de uso civil, se convirtieron en una poderosa herramienta para la adquisición de información de recursos naturales.

En el SLAR, para conseguir una buena resolución en la dirección de vuelo (resolución azimutal) se necesita una antena muy larga, siendo esto una desventaja importante por las dificultades de instalación y operatividad.

Para una dada longitud de onda, el ancho del haz puede ser controlado por dos medios diferentes:

1. *Por la longitud física de la antena.*
2. *Sintetizando la longitud de la antena.*

A los sistemas que responden al primer caso se los denomina de apertura real siendo de diseño simple y de fácil obtención de los datos; pero, debido a problemas de resolución, su operación se ve restringida a distancias cortas, a baja altitud y al uso de longitudes de onda corta; estas restricciones limitan el área de cobertura y generan mayor atenuación y dispersión atmosférica.

Dichas deficiencias han sido superadas en los radares de alta resolución transportados en satélites o aviones con el uso de otro método: los sistemas de radar de apertura sintética SAR. La antena del SAR es físicamente corta, pero, a través de la modificación en el registro de datos y en las técnicas de procesamiento, simula o "sintetiza" el efecto de una antena grande. El resultado de este modo de operación es un haz efectivo angosto, aún a largas distancias, sin recurrir a una antena físicamente grande. Esta síntesis lograda electrónicamente, se basa en el efecto Doppler, que establece que la frecuencia o longitud de onda de una fuente vibratoria, aumenta o disminuye si el receptor o la fuente están en movimiento relativo entre sí.

El radar de apertura sintética SAR es un radar activo que emite en el intervalo de frecuencias de microondas en un periodo pequeño de tiempo y recibe los ecos provenientes de reflexiones de la señal en los objetos dando lugar a una apertura sintética, es decir, debido a la gran velocidad de desplazamiento del vehículo o plataforma espacial (7.5 [Km/s] aproximadamente), la antena del dispositivo SAR se convierte en una antena virtual de mayor tamaño. El blanco permanece en el haz de la antena durante unos instantes y es observado por el radar desde numerosos puntos a lo largo de la trayectoria del satélite, lo que es equivalente a prolongar la longitud real de la antena.

El procesamiento de datos obtenidos por SAR es complicado debido a un gran volumen de información correspondiente a cada imagen. Al final de este complejo procesado se obtienen imágenes en tonos de gris de 100×100 [Km²] con una resolución de 25 [m] o entre 12.5 [m] y 200[m] por píxel dependiendo del tipo de radar y de las necesidades de usuario.

Los sistemas de SAR emiten ondas cortas de unos centímetros. La intensidad de la señal recibida es debida principalmente a la rugosidad a pequeña escala de la superficie. Cuando la superficie monitoreada mediante radar es rugosa y contiene pendientes pequeñas cuyo tamaño es comparable con la longitud de onda de la señal de radar, entonces se observa el fenómeno de difracción.

En las imágenes SAR marítimas se pueden observar áreas con altos tonos de gris, casi blanca. En el caso contrario, cuando en la superficie del mar no hay olas pequeñas, no se produce la resonancia de la señal y en la imagen se observan zona superficiales oscuras. Las trazas de materia orgánica (como lo son productos de crudo, aceites minerales, desechos urbanos, etc...) en la superficie marina, con propiedades tensioactivas, producen una disminución local de la rugosidad de la superficie del mar debida principalmente al viento y a olas, que se reflejan en una menor reflectividad de la señal emitida y detectada por los sensores SAR. Estas manchas o penachos oscuros en las imágenes SAR, son independientes de las condiciones de iluminación solar y de cobertura por nubes, sin embargo el posible contraste depende fuertemente de la velocidad del viento, a saber: cuando la velocidad del viento es mayor de 2-3 [m/s] y menor de 10-12 [m/s] las condiciones de la detección de las manchas mediante el SAR son mas favorables.

Las imágenes de los radares activos SAR, ayudan a detectar en alta mar y en las proximidades de las costas, los derrames de residuos oleosos procedentes de los buques y de otras fuentes de contaminación las cuales se aprecian estelas detectadas respectivamente mediante fotografía aérea y SAR.

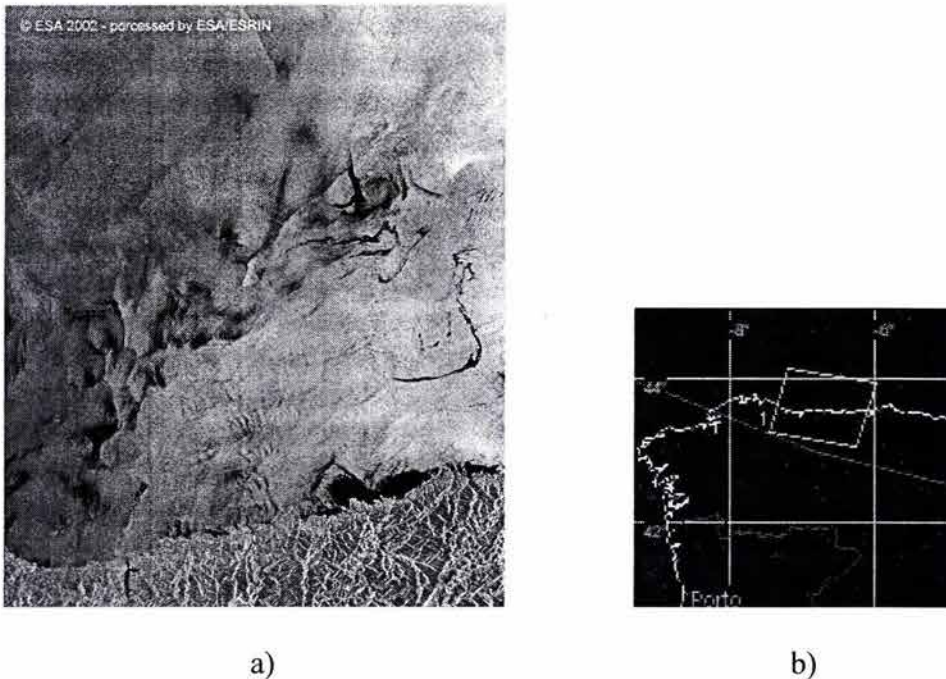


imagen SAR de un derrame petrolero sobre la región de Galicia y Asturias.

fig. 2.6

El ruido *Speckle* es uno de los principales factores que degradan la calidad de las imágenes SAR. El *Speckle* es un ruido multiplicativo proporcional a la intensidad de la señal recibida. El efecto visual de este ruido origina una textura granulosa que puede dificultar la interpretación de las imágenes de radar, reduciendo la separabilidad entre clases de uso del suelo, tipos litológicos, entre otros.

Existen dos métodos para disminuir el ruido *Speckle*: el filtraje y el procesamiento "multi-look". El ruido *Speckle* está siempre asociado a sistemas que utilizan fuentes de iluminación coherentes, tales como los obtenidos por microondas, láser y ultrasonografía.

2.4 FUNCIONALIDAD DE LAS FUENTES DE DATOS

La percepción remota ha probado ser una fuente valiosa de información ambiental sobre la atmósfera, continentes y océanos.

En 1960 los Estados Unidos de Norteamérica, lanzaron los primeros satélites para el monitoreo meteorológico, los cuales portaban pequeñas cámaras de baja resolución, trabajando en el espectro visible, actualmente se ha dado paso a el uso de sensores que trabajan en la región del IR. Debidas a estas mejoras, ahora los satélites también se usan para recolectar datos oceanográficos, hidrológicos y vegetación.

Las fotografías convencionales tomadas desde el espacio no se diferencian en nada de las realizadas por cualquier cámara fotográfica, salvo por la calidad técnica de las cámaras utilizadas. La fotografía espacial es pues, idéntica a la efectuada desde aviones, si bien la altura del satélite permite la observación de grandes extensiones de nuestro planeta. Sin embargo, su realización está condicionada al regreso de la plataforma espacial y, por consiguiente, su uso se limita casi exclusivamente a los vuelos espaciales tripulados.

Las "restituciones" fotográficas en B/N proceden del procesamiento de los datos captados por los sensores digitales. Estos, como ya hemos visto, captan la respuesta espectral de los objetos para diferentes longitudes de onda del espectro electromagnético (azul, rojo, infrarrojo térmico...) pero, en este caso, no sensibilizan una película fotográfica sino que son convertidas en señales eléctricas según su intensidad, las cuales serán almacenadas o transmitidas de forma digital para construir las imágenes (pixel a pixel). Esta información, que puede visualizarse en los centros de tratamiento (consolas, impresoras...), también puede utilizarse para sensibilizar películas fotográficas (a posteriori), que recojan analógicamente el resultado de la visualización.

El resultado de este proceso de manipulación informático de los datos lo convierte en una "imagen", donde la restitución se realiza punto a punto y en líneas sucesivas que se corresponden con el barrido del sensor sobre la superficie terrestre. Cuando se trata de sensores multispectrales este proceso de restitución puede realizarse sobre cualquiera de las bandas registradas.

Los datos de los sensores no sólo puede visualizarse en blanco y negro, sino que estos valores numéricos también pueden restituirse en color, utilizando tres de las bandas registradas por el sensor a las que se les asignan las tablas de color creadas por la computadora para los colores primarios azul, verde y rojo. Como el resultado visual es una imagen en color donde estos no tienen por qué corresponder con la realidad, se les denomina imágenes en "falso color", aunque la manipulación con la computadora y la elección de las bandas adecuadas permiten también obtener visualizaciones muy semejantes a las apreciadas por nuestros ojos. De la misma forma que en el caso anterior estas visualizaciones pueden restituirse en productos fotográficos (papel, negativos...), los cuales constituyen unos de los productos más difundidos entre los proporcionados por los satélites de reconocimiento de recursos naturales. En ellos, la cantidad de

información se multiplica respecto a las imágenes en B/N, dado que, en estas, las posibilidades de tonalidades de gris no pasan de 256, mientras los tonos de color se pueden elevar, según el procedimiento empleado, hasta 16 000 000.

Los productos pueden ser adquiridos directamente en los principales centros de distribución de los datos originales (NASA, ESA, SPOTIMAGE), los cuales ofrecen una gran versatilidad en la elección, por parte del usuario, de los canales, las combinaciones de bandas, las escalas y el tipo de soporte. Generalmente estos productos son sometidos a pequeños tratamientos para mejorar su calidad visual y su fácil acceso y la posibilidad de aplicarles técnicas de interpretación visual, les proporciona un significativo interés. Por otra parte, las imágenes están ligadas a la disponibilidad de un equipo informático adecuado, y para su obtención se necesita una especialización en las técnicas de tratamiento y clasificación de imágenes, por lo que su difusión es más restringida, aunque en los últimos años está sufriendo un espectacular despegue debido a la comercialización de una mayor variedad de equipos de tratamiento a coste más asequible para numerosos centros de investigación.

2.5 TENDENCIAS EN LA ADQUISICIÓN DE DATOS REMOTOS

Por varias décadas nuevas tecnologías ha sido introducidas para la fotografía aérea y nuevas cámaras han sido desarrolladas. Películas con mayor resolución espacial y mayores velocidades. Pero dos últimos cambios tecnológicos en la captura de imágenes aéreas marcan el actual desarrollo tecnológico.

Las fotografías aéreas han sido digitalizadas y distribuidas por medio de CD-ROM, facilitando su análisis, procesamiento y reproducción. Esta tecnología reemplazará poco a poco a las películas fotosensibles.

El segundo punto de desarrollo tecnológico ha sido la sustitución de las cámaras fotográficas por cámaras CCD (Charge Couple Device), el cual es un arreglo bidimensional de pequeños detectores sensibles a la luz, siendo la imagen almacenada de manera digital, facilitando el procesamiento posterior.

A un lado de los avances tecnológicos incorporados a los satélites y captores, la percepción remota ha tenido una aplicación más intensiva en sondas espaciales que envían fotografías de la superficie de otros planetas, siendo los mismos los captores que van a bordo de estas plataformas espaciales, los utilizados para la percepción remota de la tierra.

3

SATÉLITES Y CAPTORES

3.1 ENVISAT

ENVISAT es el satélite mas poderoso que se haya lanzado al espacio aplicado a la percepción remota. *ENVISAT* tendrá un gran número de instrumentos que estudiarán: la tierra, océanos, atmósfera y capas de hielo polares.

Auxiliará a los científicos de todo el mundo para el estudio del cambio climático y como se relaciona éste con otros sistemas ambientales. *ENVISAT* será la punta de lanza para el estudio de nuestro planeta durante toda esta década.

Muchos de los instrumentos que el *ENVISAT* tiene, son mejoras y nuevas versiones de los satélites ERS-1 y 2 de la década pasada, por lo que la comparación de las mediciones actuales y las adquiridas a lo largo de la década pasada permitirán fijar un referente sobre el medio ambiente de manera global.

ENVISAT tiene una inclinación de 98 grados, altitud de 800 [Km] la duración de una órbita es de 101 minutos y con ciclos de 35 días. El costo fue de 2 billones de euros y tiene los siguientes instrumentos:

<i>INSTRUMENTO</i>	<i>OBJETIVO</i>
<i>Global Ozone Monitoring by Occultation of Stars (GOMOS)</i>	Medir concentración de ozono en la troposfera.
<i>Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY)</i>	Medición de gases atmosféricos y concentraciones de aerosoles en la atmósfera.
<i>Michelson Interferometer for Passive Atmospheric Sounding (MIPAS)</i>	Medición de procesos químicos y físicos en la estratósfera, como el efecto de la concentración del ozono en el futuro.

<i>Medium Resolution Imaging Spectrometer (MERIS)</i>	Medición de la radiación en 15 bandas de frecuencias diferentes, para proveer información sobre oceanobiología, calidad del agua marina, vegetación terrestre, nubes y vapor de agua.
<i>Advanced Synthetic Aperture Radar (ASAR)</i>	Imagen de radar –día y noche-
<i>Advanced Along Track Scanning Radiometer (AATSR)</i>	Medición de la temperatura superficial del agua, el cual es un parámetro clave para medir el efecto invernadero.
<i>Radar Altimeter (RA-2)</i>	Medición de la distancia de la tierra al satélite, por lo que puede medir la superficie terrestre.
<i>Microwave Radiometer (MWR)</i>	Permite la corrección para los datos de RA.
<i>Doppler Orbitography and Radio Positioning Integrated by Satellite (DORIS)</i>	Proporciona la información referente a la posición actual de ENVISAT, dentro de un rango de centímetros, lo cual es de suma importancia para la corrección e interpretación de los datos.
<i>Laser Retro-Reflector (LRR)</i>	Este instrumento refleja un haz laser a las estaciones de seguimiento en tierra para determinar exactamente la posición donde se encuentre.

tabla 3.1

3.2 ERS-1 Y ERS-2

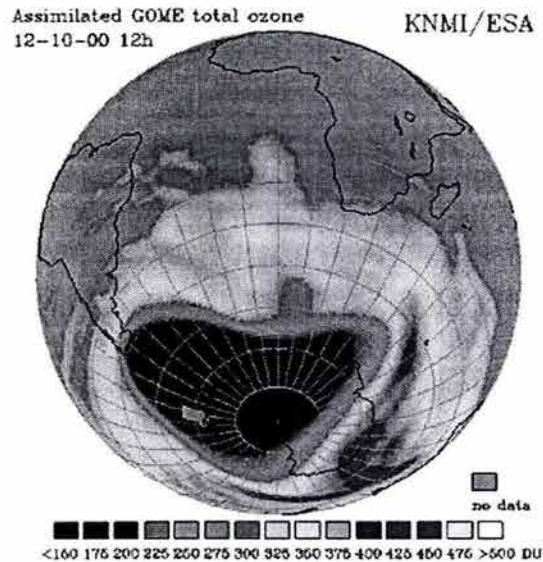
La ESA desarrolló el *European Remote Sensing Satellite*, cuya misión primaria es realizar la percepción remota de los océanos, masas polares de hielo y regiones costeras. Los satélites proporcionarán medidas globales sistemáticas, de la velocidad y dirección del viento, temperaturas superficiales, la altitud superficial, zonas cubiertas por nubes, y los niveles atmosféricos de vapor de agua.

Los datos de ERS-1 serán compartidos con la NASA según los términos de un acuerdo recíproco de Seasat y Nimbus 7. Los satélites serán controlados del centro de control de Darmstadt de la ESA.

Esta nave es estabilizada por 3 ejes, con un solo panel solar de 11.7 [m] x 2.4 [m]. Potencia máxima suministrada de 2600 W, con un banco de baterías para 2650 [Whr]. La antena del SAR es 10 m x 1 m, las comunicaciones se dan en la banda-S para el uplink de 2 kbps el downlink en la banda-X a 105 Mbps y 15 Mbps, con capacidad de almacenar hasta 6.5 Gbits.

La carga útil del Satélite ERS-1 incluye:

AMI (Active Microwave Instrument) que incluye un *SAR (Synthetic Aperture Radar)*, *Wave Scatterometer* y *Wind Scatterometer*; *Radar Altimeter (RA)*; *Scanning Radiometer and Sounder (ASTR-M)*; *Laser Reflector (LRR)*; *Precision Ranging Equipment (PRARE)*

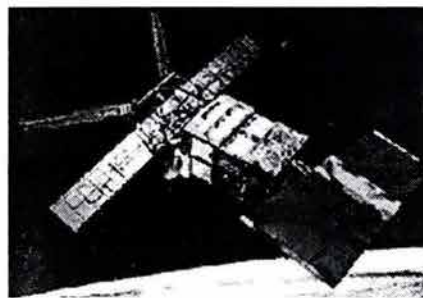


distribución del ozono en la atmósfera

fig. 3.1

Carga útil del satélite ERS-2:

Incluye los mismos instrumentos que el satélite ERS-1 mas un nuevo instrumento, el *Global Ozone Measuring Equipment (GOME)* el cual mide la distribución del ozono en la atmósfera exterior (fig. 3.1).



satélite ERS 1

fig. 3.2

*ERS 1 y ERS 2**Elementos de la Órbita:*

Sincronizada con el sol
 Altura: 985 [Km]
 Inclinación: 99°
 Recurrencia: 3 días

Año de lanzamiento: 1991 y 1994

Sensor	Espectro	IFOV	Cubrimiento
<i>Modo SAR</i>		30 [m]	100 [Km]
<i>Modo Onda</i>	5.3 [Ghz]	5 [Km]	-
<i>Modo Viento</i>		50 [Km]	500 [Km]
RA <i>Radar Altimetro</i>	13.8 [Ghz]	-	-
ATSR/M <i>Radiómetro Barredor a lo Largo de la Trayectoria/ Resonancia de Microondas</i>	1.6 [μ m] 3.7 11 12	1 [Km]	500 [Km]
<i>Radiómetro IR Termal</i>			
<i>Resonancia de Microondas</i>	23.8 [Ghz] 36.5	22 [Km]	-

tabla 3.2

Entre los diferentes sensores a bordo del satélite, se tiene el **AMI** (Active Microwave Instruments), constituido por un **SAR** (Radar de Abertura Sintética) y un escaterómetro (para

medición de vientos). Las imágenes adquiridas por el SAR, suministran datos de una faja de 100 [Km] x 100 [Km], con una resolución espacial de 30 [m].

Una antena de 10 [m] x 1 [m] emite y recibe un haz de microondas en la región de 5,3 [Ghz] (banda C), con polarización VV y un ángulo de incidencia de 23 grados.

3.3 ASTER

Japón desarrolló la tecnología *ASTER* (*Advance Space-borne Termal Emisión and Reflection Radiometer*), con el fin de obtener datos geológicos detallados y estudiar fenómenos que pudieran producir un impacto ambiental a nivel global, como por ejemplo la actividad volcánica.

ASTER es un sensor satelital de imágenes de alta resolución, transportado por el satélite *TERRA*, lanzado el 18 de diciembre de 1999. El Ministerio de Industria y Comercio Internacional de Japón (*MITI*) designó a *JAROS* (*Japan Resources Observation Systems*) para el desarrollo del sensor y a *ERSDAC* (*Earth Remote Sensing Data Análisis Center*) para el desarrollo de sistemas de procesamiento y aplicaciones de los datos.

Este sensor tiene 3 subsistemas que registran datos simultáneamente de las regiones del espectro correspondiente al visible e infrarrojo cercano (*VNIR*), infrarrojo medio o de onda corta (*SWIR*) y del infrarrojo termal (*TIR*). El *VNIR* tiene 3 bandas con 15 [m] de resolución; *SWIR* registra 6 bandas con 30 [m] de resolución; y *TIR* tiene 5 bandas con 90 [m] de resolución.

Además de alta precisión geométrica y radiométrica, es posible obtener pares estereoscópicos a lo largo del camino orbital. El ancho de barrido de *ASTER* es de 60 [Km]. La revisita es cada 16 días, sin embargo, en caso de emergencias tales como vulcanismo, incendios o inundaciones, se pueden obtener datos cada 2 días en virtud de la posibilidad de reorientar el sensor hasta +/- 24° para registrar imágenes en las órbitas adyacentes.

Este tipo de tecnología permite, entre sus mayores ventajas, identificar la distribución precisa de diferentes minerales de alteración, rocas con distintos contenidos de sílice y también producir un modelo digital de terreno.

3.4 LANDSAT

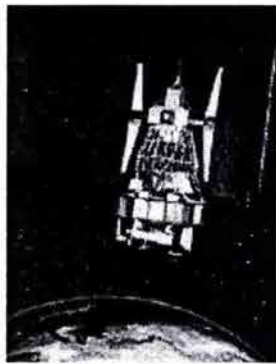
El programa *Landsat* era originalmente un proyecto de investigación experimental patrocinado por el departamento de la defensa de los Estados Unidos de Norteamérica y la NASA. Su propósito era determinar la capacidad de una nave espacial orbitando la tierra para proporcionar datos útiles sobre el uso actual de la tierra. Desde el inicio de la llegada de los datos proporcionados por *Landsat 1* fue evidente que tales datos podrían ser extremadamente valiosos para un gran número de usuarios, y todos los datos de *Landsat* fueron puestos a disposición el público en general.

En 1984, el gobierno de los Estados Unidos de Norteamérica fijó las tareas de la operación y distribución de los satélites a una entidad comercial, EOSAT (Earth Observation Satellite Company). Cuyos miembros son Hughes y RCA, las cuales tienen en funcionamiento las naves Landsat 4 y Landsat 5, vendiendo los datos proporcionados, EOSAT es también responsable del desarrollo de Landsat 6 y 7.

La estación de recepción principal es el centro de datos del EROS de USGS en Sioux Falls, Dakota del Sur. Muchos países obtienen los datos de Landsat directamente de los satélites, vía las estaciones de tierra dedicadas de Landsat con una cuota anual.

Las primeras 3 misiones de Landsat fueron conocidas como *Earth Resources Technology Satellite (ERTS)*. Las imágenes de la misión Landsat 1 mostraron gran utilidad para los datos de percepción remota sobre uso de la tierra, recursos acuíferos, agronomía, bosques, el movimiento de las capas de hielo y cartografía. Cuando Landsat 1 fue retirado de servicio en 1978 había enviado alrededor de 300 000 imágenes. Landsat 2 retirado de servicio en 1982 y Landsat 3 retirado de servicio en 1983. Las imágenes de estas misiones están disponibles con EOSAT.

El diseño de la nave espacial tomó elementos significativos del programa de Nimbus (fabricado por GE). Con potencia pico de 1000 [W]. Baterías de NiCd con paneles solares gemelos. La comunicación se da en banda-S, con downlink de 15Mbps. Sistema de la propulsión por hidracina.



Satélite Landsat-1
fig. 3.3

Landsat 1 y Landsat 2

Elementos de la Órbita:

Sincronizada con el sol
 Altura: 915 [Km]
 Inclinación: 99°
 Recurrencia: 18 días
 Año de lanzamiento: 1972 y 1975

Sensor	Espectro [μ m]	IFOV	Cubrimiento
MSS <i>Sistema Barredor Multiespectral</i>	0.5-0.6	80 [m]	185 [Km]
	0.6-0.7		
	0.7-0.8		
	0.9-1.1		
RBV <i>Camara Vidicon de Haz de Retorno</i>	0.475-0.575	80 [m]	185 [Km]
	0.580-0.680		
	0.690-0.830		

tabla 3.3

Landsat 3

Elementos de la Órbita:

Sincronizada con el sol
 Altura: 915 [Km]
 Inclinación: 99°
 Recurrencia: 18 días
 Año de lanzamiento: 1978

Sensor	Espectro [μ m]	IFOV	Cubrimiento
MSS <i>Sistema Barredor Multiespectral</i>	0.5-0.6	80 [m]	185 [Km]
	0.6-0.7		
	0.7-0.8		
	0.9-1.1		
RBV <i>Camara Vidicon de Haz de Retorno</i>	0.505-0.750	38 [m]	98 [Km]

tabla 3.4



satélite Landsat 4
fig. 3.5

El Landsat 4 y 5 fueron colocados en órbitas más bajas que sus antecesoras. Landsat 4 de manera prematura presentó una serie de fallas, por lo que se vio limitado su desempeño programado. Esto obligó al lanzamiento del Landsat 5.

El sistema de propulsión funciona con hidracina. Un solo panel solar, dos baterías de NiCd que pueden proporcionar 100 [Ahr]. Los sistemas de comunicaciones usan las bandas S, X, L y Ku. Poco después de su lanzamiento Landsat 4 perdió la mitad de su panel solar. Landsat 5 perdió dos de sus sistemas de comunicaciones primarios (downlink de banda-X y un transpondedor de banda-Ku) siendo activados los sistemas de respaldo correspondientes.

La carga útil consistía:

Multi-Spectral Scanner (**MSS**) Thematic Mapper (**TM**). El TM provee 7 bandas y el MSS tiene 4 bandas. El MSS cubre de 0.5 [μ m] a 12.6 [μ m] y provee 80 [m] de resolución con 185 [Km] de ancho de barrido. El TM cubre de 0.45 [μ m] a 12.5 [μ m] con resolución de 30 [m] en la banda de VIS/IR y 120 [m] en la banda termal y banda infrarroja.

Landsat 4 y Landsat 5

Elementos de la Órbita:

Sincronizada con el sol
 Altura: 705 [Km]
 Inclinación: 98°
 Recurrencia: 17 días
 Año de lanzamiento: 1982 y 1984

Sensor	Espectro [μ m]	IFOV [m]	Cubrimiento
MSS <i>Sistema Barredor Multiespectral</i>	0.5-0.6	80	185 [Km]
	0.6-0.7		
	0.7-0.8		
	0.9-1.1		
TM <i>Mapeador Temático</i>	0.45-0.52	30	185 [Km]
	0.52-0.60		
	0.63-0.69		
	0.75-0.90		
	1.55-1.75		
	2.08-2.35		
	10.40-12.5	120	

tabla 3.5

Landsat 6 fue diseñado para continuar el programa Landsat y llevó una versión mejorada de instrumentos. Sin embargo, no pudo alcanzar órbita durante lanzamiento, y forzó la operación de Landsat 4 y 5.

Contaba con un sistema de propulsión de hidracina. Un panel solar capaz de proveer 1430 [W], dos baterías de NiCd de 100 [Ahr]. Downlink directo a 85 Mbps.

El *Mapeador Temático Mejorado (ETM)* con 15 [m] de resolución pancromática y 30 [m] en banda multiespectral.

Landsat 6 y Landsat 7

Elementos de la Órbita:

Sincronizada con el sol
 Altura: 705 km.
 Inclinación: 98°
 Recurrencia: 17 días
 Año de lanzamiento: 1993 y 1999

Sensor	Espectro μm	IFOV	Cubrimiento
ETM+ Mapeador Temático Mejorado Plus	0.5-0.9	13 m × 15 m	185 km
	0.45-0.52	30 m	
	0.52-0.60		
	0.63-0.69		
	0.75-0.90		
	1.50-1.75		
	2.08-2.35	60 m	
	10.40-12.50		

tabla 3.6

Los sensores de estos satélites fueron:

RBV (Return Beam Vidicon): es un sistema semejante a una cámara de televisión y permite el registro instantáneo de un área del terreno. La energía proveniente de toda la escena sensibiliza la superficie fotosensible del tubo de la cámara y durante un determinado tiempo, la entrada de energía es interrumpida por un obturador, para que la imagen del terreno sea barrida por un haz de electrones. La señal de vídeo puede ser transmitida teleméricamente.

MSS (Multispectral Scanner): es un sistema sensor que permite la captura de líneas del terreno en una faja de 185 [Km] perpendicularmente a la órbita del satélite. El barrido del terreno es realizado con ayuda de un espejo que oscila perpendicularmente al desplazamiento del satélite. Durante la oscilación del espejo, la imagen del terreno, a lo largo de la faja, es captada por una matriz de detectores. La dimensión de cada detector que compone la matriz de detectores, es responsable por el campo de visión instantáneo (área de la superficie de la Tierra observada por cada detector). La energía registrada por cada detector es transformada en una señal eléctrica y posteriormente es transmitida para las estaciones en la Tierra.

TM (Thematic Mapper): es un sistema avanzado de barrido multiespectral concebido para proporcionar: resolución espacial más fina, mejor discriminación espectral entre objetos de la superficie terrestre, mayor fidelidad geométrica y mejor precisión radiométrica con relación al sensor MSS.

La energía proveniente de la escena alcanza el espejo de barrido que oscila perpendicularmente a la dirección de desplazamiento del satélite en sentido este-oeste y oeste-este. La señal atraviesa un telescopio y un conjunto de espejos, cuya función principal es corregir la señal colectada por el espejo de barrido. De esta manera, la señal detectada en cada matriz de detectores de cada banda es transferida para un amplificador y convertida en una señal digital a través de un sistema A/D (analógico/digital). Los datos de salida son entonces transmitidos vía telemetría.

3.5 SPOT

El *SPOT* (*Satellite Probatoire d'Observation de la Terre*) es el programa civil patrocinado por el gobierno francés en colaboración con Suecia y Bélgica. Proporcionando cobertura completa a la tierra cada 26 días. Los imágenes son manejadas por SPOT-Image Corp.



Satélite SPOT 1,2,3

fig. 3.6

Instrumento:

HRV (*High Resolution Visible*) con una resolución de 20 [m] en modo multiespectral y 10 [m] en modo pancromático. *HRV* cubre las siguientes bandas espectrales: 0.5 [μ m] - 0.59 [μ m], 0.61 [μ m] - 0.68 [μ m], 0.79 [μ m] - 0.89 [μ m], y la banda pancromática: 0.51 [μ m]- 0.73 [μ m]. El ancho de campo es de 4.13° en ambos modos con 3000 pixels por línea multiespectral y 6000 por línea pancromática. Con un barrido de 60 [Km] de ancho.

SPOT 1, SPOT 2 Y SPOT 3

Elementos de la Órbita:

Sincronizada con el sol

Altura: 832 [Km]

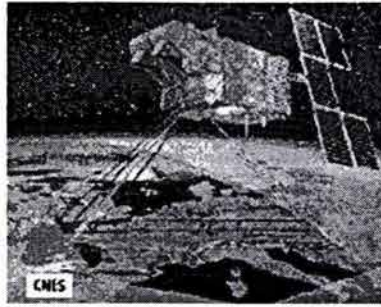
Inclinación: 99°

Recurrencia: 26 días

Año de lanzamiento: 1986, 1990 y 1993(perdido)

Sensor	Espectro [μ m]	IFOV [m]	Cubrimiento [Km]
HRV <i>Alta Resolución Visible</i>	0.5-0.59	20	60
	0.61-0.68		
	0.79-0.89		
	0.51-0.73	10	

tabla 3.7



satélite SPOT-4

fig. 3.7

Instrumentos del SPOT-4:

Dos instrumentos *HRVIR* (*High Resolution Visible - Infrared*), derivados del *HRV* de las misiones anteriores SPOT 1-3. Este sistema obtiene imágenes de 10 [m] de resolución en modo pancromático y 20 [m] en banda multispectral. El *HRVIR* incluye una banda especial para IR con la capacidad de análisis de vegetación y bosques. Cada *HRVIR* tiene un ancho de barrido de 60 [km]. El instrumento *Vegetation Monitoring* tiene una resolución de 1 [Km] en las mismas bandas que el *HRVIR*.

Para el satélite SPOT 5 se planea incluir el instrumento de *HRG* (*High Resolution Energy*). Cada nave espacial llevará 3 *HRG*'s, uno que "mira" adelante, uno que mira abajo, y uno que mira a popa. El sistema proporcionará resolución de 5 [m] en modo pancromático y 10 [m] en banda multispectral

Una de las características principales de los instrumentos a bordo del SPOT es la posibilidad de observación "off-nadir". El sensor podrá ser direccionado de modo de observar escenas laterales a la órbita en la que se encuentra el satélite en un momento determinado. Esta posibilidad de observación "off-nadir" aumenta los medios de obtener un aumento en el recubrimiento repetitivo de determinadas áreas. Otra ventaja de la vista "off-nadir" es la posibilidad de obtener pares estereoscópicos de determinadas áreas.

La luz proveniente de la escena alcanza un espejo plano, que puede ser controlado a partir de las estaciones terrestres variando en ángulos de +/- 0,6 hasta 27° con relación al eje vertical.

La energía que alcanza el espejo plano es captada por una matriz lineal de detectores del tipo CCD (*Charge-Coupled Device*). Cada matriz consiste en 6000 detectores ordenados linealmente, formando lo que se denomina "push-broom scanner" o sistema de barrido electrónico. Este sistema permite el imageamiento instantáneo de una línea completa en el terreno, perpendicularmente a la dirección de desplazamiento del satélite en su órbita.

SPOT 4

Elementos de la Órbita:

Sincronizada con el sol

Altura: 822 [Km]

Inclinación: 99°

Recurrencia: 26 días

Año de lanzamiento: 1986

Sensor	Espectro [μ m]	IFOV	Cubrimiento
HRVIR <i>Alta Resolución</i> <i>Visible e Infrarrojo</i> <i>Medio</i>	0.5-0.59	20 [m]	60 [Km]
	0.61-0.68		
	0.79-0.89		
	1.55-1.75		
Pancromático	0.61-0.68	10 [m]	
VMI <i>Instrumento de</i> <i>Monitoreo de</i> <i>Vegetación</i>	0.43-0.49	1 [Km]	200 [Km]
	0.50-0.59		
	0.61-0.68		
	0.79-0.89		
	1.58-1.75		

tabla 3.8

3.6 IKONOS

El satélite *IKONOS*, es el primer satélite comercial de *alta resolución*. Toma simultáneamente imágenes pancromáticas (B/N) de 1 [m] de resolución e imágenes multispectrales (color) de 4 [m] de resolución (4 bandas espectrales).

IKONOS toma imágenes de la Tierra desde una altura de 680 [Km] y vuela a una velocidad de 7 [Km] por segundo. La órbita del *IKONOS* es circular, casi polar y helio-sincrona (es decir, sigue la superficie de la Tierra iluminada por el sol). El satélite da 14 vueltas por día alrededor de la Tierra, es decir, una vuelta cada 98 minutos.

La cámara del satélite puede distinguir objetos en la superficie terrestre de aproximadamente 1 [m] cuadrado, por lo tanto, se pueden “ver” autos, camiones, árboles, casas, barcos, aviones, u cualquier objeto mayor de 1 [m]. Las imágenes *IKONOS* tienen una amplia gama de utilidades como elaboración de mapas, agricultura, vegetación,

planificación de las ciudades, medio ambiente, seguridad nacional, redes de transporte, seguros, y evaluación de los desastres naturales

IKONOS

Elementos de la Órbita:

Sincronizada con el sol
 Altura: 681 [Km]
 Inclinación: 98.1°
 Recurrencia: 2.9 días con 1 m. de resolución
 Año de lanzamiento: 1999

Sensor	Espectro [μ m]	IFOV	Cubrimiento
<i>Pancromático</i>	0.45-0.90	1 [m]	13 [Km]
<i>Multiespectral</i>	0.45-0.52 0.52-0.60 0.63-0.69 0.76-0.90	4	

tabla 3.9

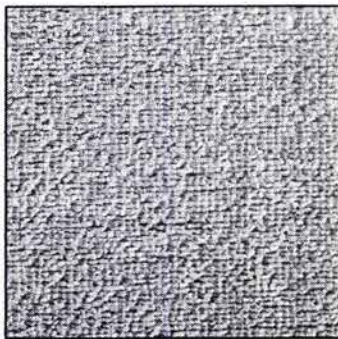
4

ANÁLISIS TEXTURAL

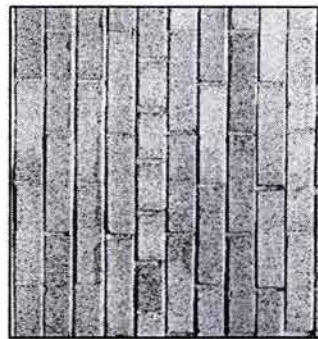
4.1 ALTERNATIVAS EN EL ANÁLISIS DE TEXTURAS

La textura es una propiedad innata de la materia de la cual se puede obtener información importante para determinar la relación que guardan las superficies y el medio que las rodea, la textura de las superficies puede ser usada para la discriminación de éstas.

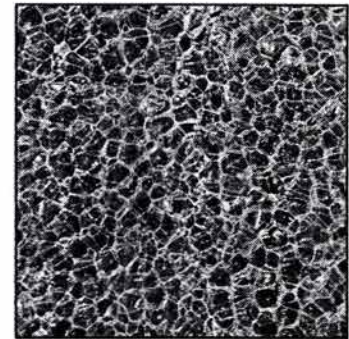
Por textura se entiende al patrón bidimensional de una región determinada en la cual se presentan variaciones de los diferentes niveles de gris, por lo que no se puede definir en un punto.



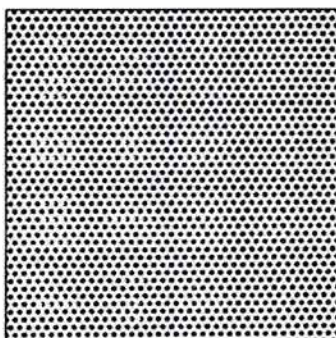
papel
a)



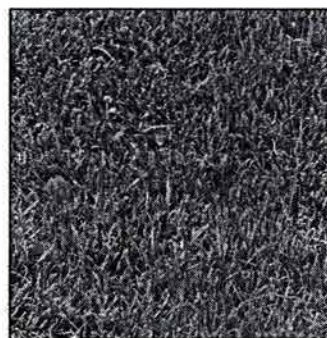
ladrillo
b)



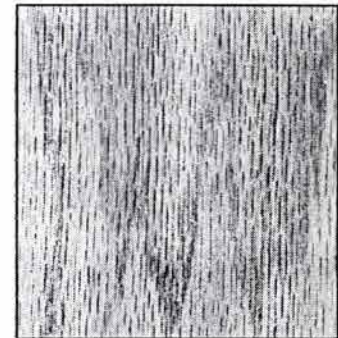
burbujas
c)



malla
d)



pasto
e)
ejemplos de texturas
fig 4.1



madera
f)

Lo que distingue a las imágenes anteriores es lo que llamamos textura, es decir, una disposición geométrica espacial específica de los niveles de gris.

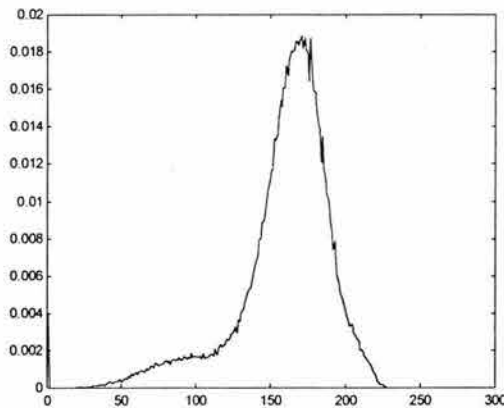
El tono y la textura siempre están presentes en una imagen y alguna de ellas puede dominar, cuando una pequeña área de la imagen tiene una variación pequeña entre los niveles de gris de los pixels que la componen, la propiedad predominante es el tono. Por otro lado, cuando en una pequeña área de una imagen se tiene un amplio rango de niveles de gris, la propiedad predominante es la textura.

Las tres principales aproximaciones usadas en el procesamiento digital de imágenes para describir la textura de una región son:[GONZ73]

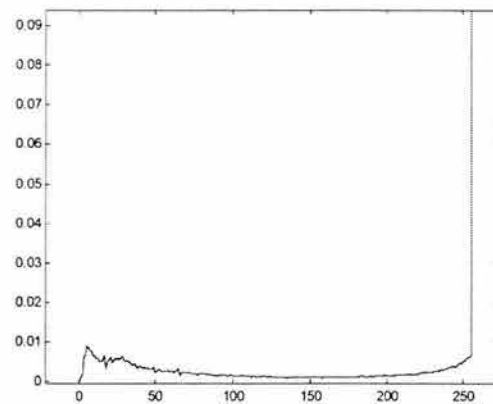
1. *Estadística*
2. *Estructural*
3. *Espectral*

4.1.1 Aproximación estadística

El histograma de una imagen digital con niveles de gris en $G \in \{1, 2, \dots, N_g\}$, es una función discreta $p(r_G) = \frac{n_G}{n}$, donde n es el número total de pixels en la imagen y n_G es el número de pixels pertenecientes al nivel de gris G , por lo que $p(r_G)$ es la probabilidad de ocurrencia del nivel de gris r_G .



histograma de la fig. 4.1 b)



histograma de la fig. 4.1 d)

fig. 4.2

Una de las aproximaciones más simples para descripción de textura es por medio del uso de momentos aplicados al histograma de los niveles de gris de una imagen o determinada región de interés. Sea z la variable aleatoria que denota el nivel discreto de las intensidades y sea $p(z_i)$, el correspondiente histograma, con L igual al número de intensidades de gris, el n -ésimo momento de z con respecto a la media es:

$$m_n(z) = \sum_{i=1}^L (z_i - \mu)^n p(z_i)$$

ec. 4.1

donde la media m queda definida como:

$$\mu = \sum_{i=1}^L z_i p(z_i)$$

ec. 4.2

De la ecuación 4.1, para $n=0$, $m_0 = 1$ y para $n=1$, $m_1 = 0$.

La varianza denotada como σ^2 (segundo momento con respecto a la media) es de particular importancia en la descripción de textura, la cual es una medida del contraste de una región que puede ser usado para establecer, de manera intuitiva, un descriptor de suavidad, por ejemplo:

$$R = 1 - \frac{1}{1 + \sigma^2}$$

ec. 4.3

para un área con valores de gris más o menos iguales, el valor de σ^2 se aproxima a cero, lo cual se ve reflejado en el valor de $R=0$, por otra parte, si los valores de gris en el área de estudio se encuentran bastantes dispersos con respecto al valor central de la media, se obtiene de forma aproximada que $R=1$.

El tercer momento del histograma con respecto a la media, es una medida llamada curtosis del histograma, mientras que el cuarto momento es una medida de la planicidad del mismo. Los momentos de mayor grado no están tan relacionados con la forma del histograma, por lo que no proporcionan información cuantitativa de éste.

La figura 4.3 contiene 50% de pixels negros y 50% de pixels blancos, siendo su distribución de intensidades el mismo, es decir, el histograma de los tres casos es idéntico, mas su textura es diferente. Como se puede observar, la limitante en el trabajo del histograma es la carencia de información sobre la posición espacial que se guarda entre píxel y píxel. Una manera para introducir ésta información en el proceso de análisis textural es considerar no solo la distribución de intensidades, sino también la posición entre píxel y píxel.

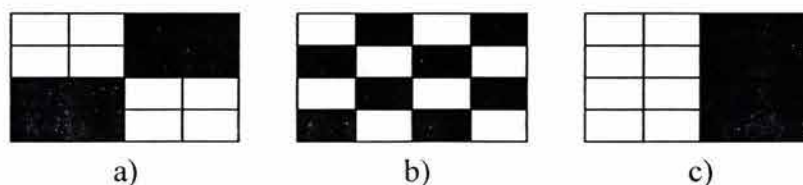


fig 4.3

La Matriz de Co-ocurrencia de Niveles de Gris es una herramienta que se propuso en [HARA73]. De manera breve consiste en lo siguiente. Sea P el operador posición y A una matriz $k \times k$ elementos con a_{ij} es el número de veces que un píxel con nivel de gris z_i ocurre (en la posición especificada por P) relativo con los píxels con nivel de gris z_j .

Por ejemplo, si consideramos la siguiente imagen que contiene 3 niveles de gris; $z_1=0$, $z_2=1$ y $z_3=2$, como sigue:

0	0	0	1	2
1	1	0	1	1
2	2	1	0	0
1	1	0	2	0
0	0	1	0	1

fig 4.4

Definiendo la posición del operador P como : “*píxel a la derecha y píxel abajo*” (-45°), bajo ésta regla se obtendría la siguiente matriz:

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

a)

$$C = \begin{bmatrix} 0.25 & 0.125 & 0.0625 \\ 0.125 & 0.1875 & 0.125 \\ 0 & 0.125 & 0 \end{bmatrix}$$

b)

fig 4.5

donde el elemento a_{11} (superior izquierda), corresponde con el número de veces que el punto con nivel $z_1=0$ aparece un píxel a la derecha y debajo de un píxel con el mismo nivel de gris, el elemento a_{13} (superior derecha) es el número de veces que el punto con nivel $z_1=0$ aparece un píxel abajo y a la derecha de un píxel con valor con nivel de gris $z_3=2$. Como puede observarse, la dimensión de esta matriz cuadrada es igual al número de niveles de gris que contiene la imagen.

Llamemos n a la suma de los elementos pertenecientes a la matriz A (fig 4.5-a) los cuales han sido obtenidos por medio de aplicar el operador P , entonces, la matriz C , será el resultado de dividir todos los elementos de la matriz A , entre n . Los elementos $c_{ij} \in C$ representan la probabilidad conjunta del par de píxels que satisfacen el operador P y que tienen por valor los niveles de gris z_i y z_j . La matriz es llamada: *Gray-Level Cooccurrence Matrix* (fig. 4.1-b). La presencia y una mejor descripción de la textura puede ser obtenida por medio de una elección correcta de la posición de los píxels que involucra el operador. La aproximación estadística es bastante útil cuando el área en estudio es lo bastante pequeña (microtextura) que puede conservar la información de esta textura específica.

El manejo de la Matriz de co-ocurrencia implica un gran costo computacional, debido a que para cada píxel, una matriz nueva debe ser formada y los respectivos descriptores de textura ser calculados.

Existe el método de suma y diferencia de histogramas [UNSE86], en el cual, el histograma S depende de los desplazamientos d_x y d_y , por ejemplo, el valor de gris en el píxel con la posición

(i,j) será sumado con el que se encuentra en la posición $(i+d_y, j+d_x)$, para una imagen en niveles de gris de 8 bits, la suma estará comprendida entre 0 y 512. El histograma diferencia D se calcula de manera análoga y sus valores estarán comprendidos entre -255 y 255. Posteriormente los histogramas S y D son normalizados por lo que los valores en ellos, se consideran como probabilidades.

Los descriptores de textura propuestos son:

Media:

$$\mu_s = \sum_i i \cdot S(i)$$

ec. 4.4

Energía:

$$E = \sum_i S(i)^2 \cdot \sum_j D(j)^2$$

ec. 4.5

Contraste:

$$C = \sum_j j \cdot D(j)$$

ec. 4.6

Homogeneidad:

$$H = \sum_j \frac{1}{1+j^2} \cdot D(j)$$

ec. 4.7

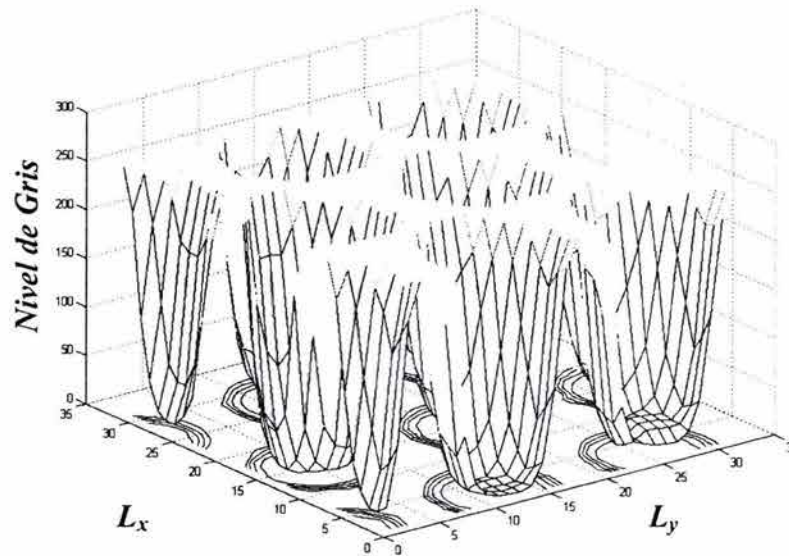
Entropía:

$$E_m = -\sum_i S(i) \cdot \log(S(i)) - \sum_j D(j) \cdot \log(D(j))$$

ec. 4.8

La velocidad para el cálculo de estos descriptores, puede llegar a ser del orden de veinte veces menor con respecto a las usadas por la GLCM.[PARK96]

También existen aproximaciones estadísticas que consideran el nivel de gris como superficie tridimensional, donde el nivel de gris es la tercera dimensión.



representación en tres dimensiones de la fig. 4.1-d)
para una región de 32x32 pixels
fig. 4.6

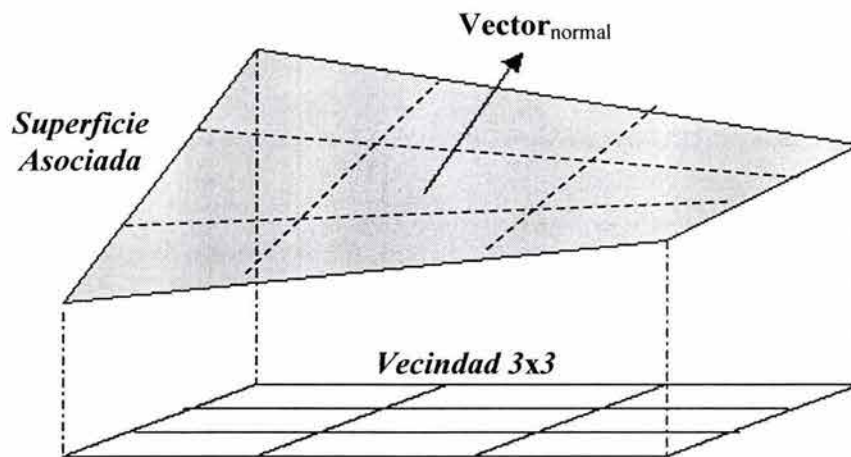


fig. 4.7

Las figura anterior muestra una región que comprende una vecindad de 3 x 3 niveles de gris y el vector normal correspondiente a la superficie relacionada a esa vecindad. El conjunto de 3 pixels horizontales y 3 pixels verticales van formando diferentes planos que se van uniendo sin solaparse entre sí, por lo que el plano resultante está representado por una ecuación lineal en dos dimensiones, la cual puede escribirse de la siguiente forma:

$$I(i, j) = \alpha i + \beta j + \gamma$$

ec. 4.9

para la imagen I , que se encuentra sobre el área de 3×3 pixels. Los coeficientes son fáciles de encontrar usando el ajuste de mínimo error cuadrático para un plano, el cual se representa como:

$$\alpha = \sum_{i=-1}^1 \sum_{j=-1}^1 i \cdot I(i, j)$$

ec. 4.10

$$\beta = \sum_{i=-1}^1 \sum_{j=-1}^1 j \cdot I(i, j)$$

ec. 4.11

$$\gamma = \sum_{i=-1}^1 \sum_{j=-1}^1 I(i, j)$$

ec. 4.12

Los coeficientes pueden ser tratados como vector. El vector normal es perpendicular al plano y el vector normalizado de la i -ésima superficie de las vecindades asociadas es de la forma:

$$\begin{pmatrix} K_i \\ L_i \\ M_i \end{pmatrix} = \frac{1}{\sqrt{\alpha_i^2 + \beta_i^2 + \gamma_i^2}} \begin{pmatrix} \alpha_i \\ \beta_i \\ \gamma_i \end{pmatrix}$$

ec. 4.13

Para una vecindad de N superficies asociadas con sus respectivos vectores normalizados, se tiene el valor κ

$$\kappa = \frac{N-1}{N-R}$$

ec. 4.14

R es calculado

$$R^2 = \left(\sum_{i=1}^N K_i \right)^2 + \left(\sum_{i=1}^N L_i \right)^2 + \left(\sum_{i=1}^N M_i \right)^2$$

ec. 4.15

para valores de κ cercanos a cero se tiene una superficie rugosa y para valores grandes de κ se tiene una superficie suave, por lo que se puede calcular un valor de κ para cada píxel considerando su vecindad.

En el método de dispersión vectorial, la primera aproximación que se realiza es considerar que un plano pasa por la vecindad de pixels, en la presente aproximación se realiza la suposición que es una superficie de segundo grado la que mejor se aproxima.

El polinomio de segundo grado que define a la superficie de nuestro interés esta dado por:

$$z(x, y) = a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{10}x + a_{01}y + a_{00}$$

ec. 4.16

La curvatura de la superficie puede ser descrita como el cambio en la pendiente de la tangente a un punto de la superficie de nuestro interés.

Existen diferentes medidas de curvatura que proporcionan información interesante sobre la superficie en la cual se calcula, de manera general el procedimiento para obtener esta información se siguen tres pasos:

1. Ajustar por mínimos cuadrados el polinomio para el mejor ajuste de la superficie.
2. Calcular las derivadas de la superficie en el punto indicado.
3. Usar tales derivadas –la pendiente de la tangente-, para calcular la curvatura.

Este procedimiento[PARK96] ya se encuentra hecho, por lo que los valores de los coeficientes para la superficie centrada en $K(i, j)$, encuentran definidos como:

$$a_{20} = \frac{A_2}{2} - \frac{A_1}{3}$$

ec 4.17

$$a_{11} = \frac{A_6}{4}$$

ec 4.18

$$a_{02} = \frac{A_5}{2} - \frac{A_1}{3}$$

ec 4.19

$$a_{10} = \frac{A_2}{6}$$

ec 4.20

$$a_{01} = \frac{A_3}{6}$$

ec 4.21

$$a_{00} = \frac{5A_1}{9} - \frac{A_4}{3} - \frac{A_5}{3}$$

ec 4.22

donde

$$A_1 = \sum_{n=-1}^1 \sum_{m=-1}^1 I(i+n, j+m)$$

ec. 4.23

$$A_2 = \sum_{n=i-1}^{i+1} I(n, j+1) - \sum_{n=i-1}^{i+1} I(n, j-1)$$

ec. 4.24

$$A_3 = \sum_{m=j-1}^{j+1} I(i-1, m) - \sum_{m=j-1}^{j+1} I(i+1, m)$$

ec. 4.25

$$A_4 = \sum_{n=i-1}^{i+1} I(n, j-1) - \sum_{n=i-1}^{i+1} I(n, j+1)$$

ec. 4.26

Los siguientes parámetros son llamados la primera forma (mayúsculas) y segunda forma (minúsculas) fundamental de la superficie:

$$E = 1 + a_{10}^2$$

ec. 4.27

$$F = a_{10}a_{01}$$

ec. 4.28

$$G = 1 + a_{01}^2$$

ec. 4.29

$$e = \frac{2a_{20}}{\sqrt{EG - F^2}}$$

ec. 4.30

$$f = \frac{2a_{11}}{\sqrt{EG - F^2}}$$

ec. 4.31

$$g = \frac{2a_{02}}{\sqrt{EG - F^2}}$$

ec. 4.32

La curvatura mínima en el punto $I(i,j)$ queda establecida de la siguiente forma

$$k_1 = \frac{gE - 2Ff + Ge - \sqrt{(gE - 2Ff + Ge)^2 - 4(eg - f^2)(EG - F^2)}}{2(EG - F^2)}$$

ec. 4.33

la curvatura máxima es

$$k_1 = \frac{gE - 2Ff + Ge + \sqrt{(gE - 2Ff + Ge)^2 - 4(eg - f^2)(EG - F^2)}}{2(EG - F^2)}$$

ec. 4.34

la curvatura gaussiana esta definida

$$k_g = k_1 \times k_2$$

ec. 4.35

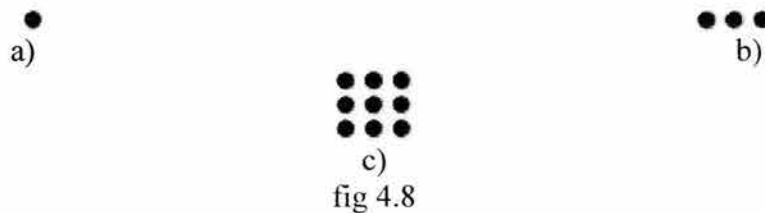
si el término $eg - f^2 > 0$ en el punto $I(i,j)$ es elíptico, si el término $eg - f^2 = 0$ en el punto $I(i,j)$ es parabólico y si el término $eg - f^2 < 0$ en el punto $I(i,j)$ es conocido como silla de caballo.

4.1.2 Aproximación estructural

La segunda clasificación para la descripción de texturas, está basada en conceptos estructurales. Supongamos que se tiene un regla de la forma $S \rightarrow aS$, la cual indica que el símbolo S , será reescrito como aS (tres aplicaciones de la regla arrojarán como resultado $aaaS$). Por ejemplo, si a , es la estructura de la figura 4.8-a) y la declaración “ n -repeticiones de la estructura a la derecha” se asigna a la cadena de longitud n de forma $aaa...aS$, entonces se permitirá la generación del patrón de la figura 4.8-b).

Los métodos estructurales suponen que la textura está formada por una estructura que se repite, denominada primitiva. Se intenta detectar esa primitiva y estudiar su distribución.

Algunos tipos de texturas poseen una unidad estructural que se repite llamada *texel*, pero en general a veces no es fácil encontrarlo y en gran cantidad de ocasiones ni siquiera existe, por esta razón no existe un mejor método de caracterización de texturas y depende del tipo de aplicación el método a seleccionar.



Supongamos ahora que se asignarán nuevas reglas al ejemplo anterior, las cuales son: $S \rightarrow cS$ y $S \rightarrow bS$, donde la presencia de b es “estructura abajo” y la presencia de c se interpretaría como

“estructura a la izquierda”. Ahora la secuencia $aaabccbaaS$ sería representada por una matriz de 3×3 estructuras (fig. 4.8-c). La generación de estas estructuras no se limita a la forma rectangular.

La idea básica de esto es que la generación de texturas complejas puede describirse por otras más sencillas, por lo que el estudio de las reglas que las generan es importante.

4.1.3 Aproximación espectral

Las transformada de Fourier también permite caracterizar texturas. A partir de la transformada de Fourier de la región que posee la textura se calculan dos parámetros: La suma de valores dentro de un anillo (fig. 4.9-a) o dentro de cierto ángulo (fig. 4.9-b).

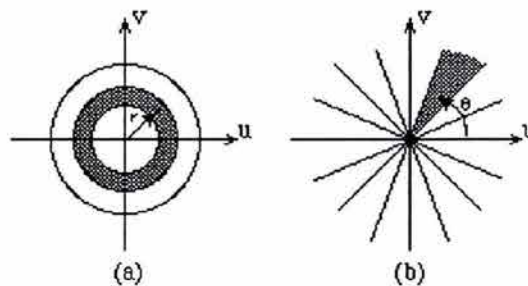


fig. 4.9

Para patrones texturales periódicos o casi periódicos, el espectro de Fourier es una herramienta que permite la descripción de estos patrones. Estos patrones globales de textura relativamente grande, son difíciles de describir por medio de métodos espaciales, ya que estos analizan la imagen en partes. Las características que permiten el estudio textural por medio de esta aproximación son:

1. Los picos prominentes del espectro dan la dirección principal de los patrones texturales, ya que la localización de los picos en el plano de la frecuencia proporciona el periodo espacial de los patrones texturales.
2. Si se elimina cualquier componente periódica por medio del filtrado, se dejan los elementos no periódicos, los cuales pueden ser descritos por medio de estadísticas.

Para el estudio de estas características del espectro se considera como una función bidimensional $S(r, \theta)$, donde S es la función del espectro y r, θ son variables en este sistema de coordenadas.

Para cada dirección θ , $S(r, \theta)$ se considera como una función unidimensional: $S_\theta(r)$, de manera similar para cada frecuencia r , $S_r(\theta)$ es una función unidimensional. Una descripción más global, es obtenida por medio de la suma de las diferentes variables como a continuación se presentan:

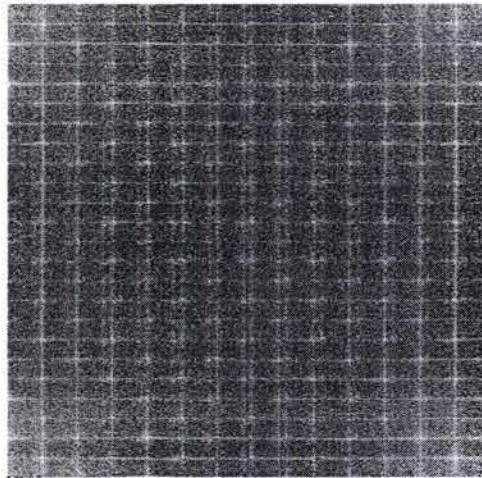
$$S(r) = \sum_{\theta=0}^{\pi} S_\theta(r)$$

ec. 4.36

$$S(\theta) = \sum_{r=1}^R S_r(\theta)$$

ec. 4.37

donde R es el radio del círculo con centro en el origen, para un espectro de $N \times N$, típicamente se selecciona $R = \frac{N}{2}$. Los resultados de las ecuaciones 4.4 y 4.5, proporcionan un par de valores $[S(r), S(\theta)]$ para cada par de coordenadas (r, θ) , al variar estas coordenadas se puede estudiar el comportamiento del espectro de la imagen o textura bajo estudio.



espectro de la fig. 4.1 d)
fig. 4.10

El estudio de imágenes con texturas usando la función de Gabor ha sido una técnica bastante exitosa para caracterizar texturas, en gran medida está relacionado con las propiedades métricas de la transformada de Gabor.

Una wavelet madre bidimensional es un oscilador armónico formado por una onda senoidal plana a una determinada frecuencia, orientado en dirección del eje x , la cual modula a una función gaussiana. La función de Gabor está totalmente determinada con la frecuencia (ω), y el ancho de banda (σ_x, σ_y). La wavelet madre bidimensional compleja está definida de la siguiente forma

$$G(\omega, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \exp(2\pi i \omega x)$$

ec. 4.38

en general la parte real se toma:

$$G(\varpi, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cos(2\pi\varpi x)$$

ec. 4.39

el banco de filtros de Gabor es obtenido al dilatar y rotar la wavelet madre:

$$G_{nm}(x, y) = a^{-m} G(x', y')$$

ec. 4.40

$$x' = a^{-m} (x \cos \theta_n + y \sin \theta_n)$$

ec. 4.41

$$y' = a^{-m} (-x \sin \theta_n + y \cos \theta_n)$$

ec. 4.42

donde $a > 1$, n y m son enteros, $\theta = \frac{n\pi}{K}$, K es el número total de orientaciones.

La localización del filtro de Gabor en el plano (ϖ_x, ϖ_y) y la forma de dos de estos filtros se muestra en la gráfica siguiente

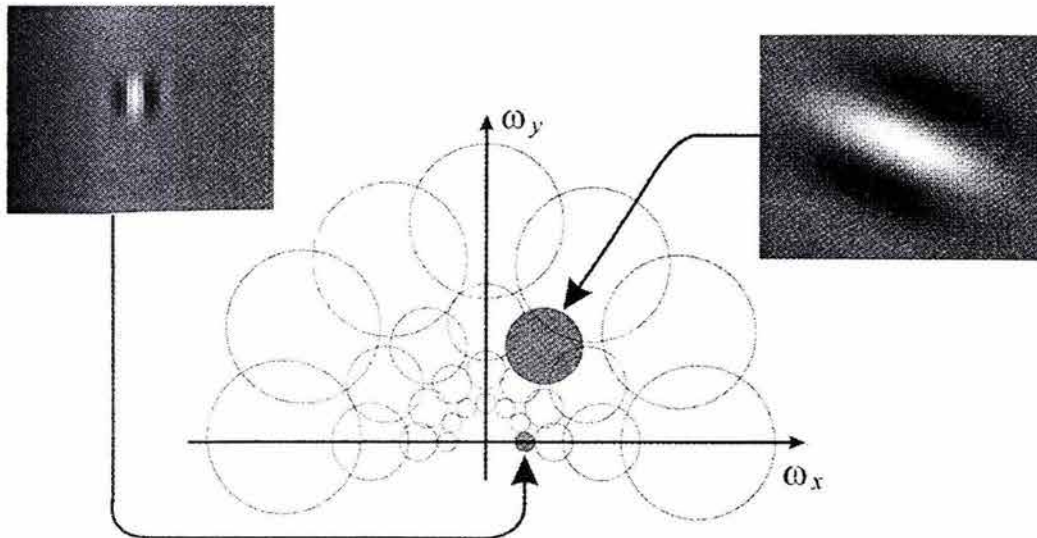


fig. 4.11

La imagen $I(x,y)$ es analizada por medio de la convolución con el banco de filtros, obteniendo una serie de funciones respuesta $w_{mn}(x,y)$

$$w_{mn}(x,y) = \sum_{a,b} G_{mn}(x-a,y-a)I(a,b)$$

ec. 4.43

las matrices $w_{mn}(x,y)$ no son usadas para representar directamente la imagen, para cada $w_{mn}(x,y)$, se calcula la media y la desviación estándar, los cuales son tomados en forma de vector para su uso como descriptores.

El desempeño del uso de filtros de Gabor es bueno, el problema consiste en encontrar la combinación adecuada entre escala y dirección de los filtros que mejor sirvan para extraer información de las texturas.

4.2 MATRIZ DE CO-OCURRENCIA DE NIVELES DE GRIS (GLCM)

Se considera que una imagen I tiene N_x pixels en la dirección horizontal y N_y pixels en dirección vertical. Se tiene que cada elemento perteneciente a este arreglo, ha sido cuantizado en alguno de N_g valores. Sea $L_x=\{1,2,\dots,N_x\}$ sea el dominio espacial horizontal y $L_y=\{1,2,\dots,N_y\}$ el dominio espacial vertical y $G=\{1,2,\dots,N_g\}$ el conjunto de N_g niveles cuantizados de gris. La imagen I puede ser representada como una función que asigna algún nivel de gris de G a cada par de pixels con coordenadas $L_x \times L_y$; $I: L_x \times L_y \rightarrow G$.

Se ataca el problema de clasificar texturas, obteniendo descriptores estadísticos de textura, basándose en el hecho de que una textura es esencialmente una interacción estadística de píxeles, un campo aleatorio homogéneo bidimensional.

La GLCM consiste en la matriz que provee la probabilidad conjunta de todas las combinaciones de pixels existentes en una ventana perteneciente a la imagen, definida para el cálculo.

La entrada sin normalizar $P(i,j)$ de la matriz, representa el número de ocurrencias de dos pixels vecinos, localizados en las coordenadas (k,l) y (m,n) , con $((k,l),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x)$, y que se encuentra localizados en la ventana W sobre la cual se calculará la GLCM, los valores de gris respectivos son: $I(k,l)=i$ y $I(m,n)=j$. Los dos pixels están separados una distancia δ y con una dirección α . [HARA73]

Para ángulos múltiples de 45° las frecuencias sin normalizar de co-ocurrencia quedarían como:

$$P(i,j,\delta,0^\circ) = \# \{ ((k,l),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x) \mid k-m=0, |l-n|=\delta \}$$

ec. 4.44

$$P(i,j,\delta,45^\circ) = \# \left\{ ((k,l),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x) \mid \begin{array}{l} k-m=\delta, l-n=-\delta \\ k-m=-\delta, l-n=\delta \end{array} \right\}$$

ec. 4.45

$$P(i, j, \delta, 90^\circ) = \#\{(k, l), (m, n) \in (L_y \times L_x) \times (L_y \times L_x) \mid |k - m| = \delta, l - n = 0\}$$

ec. 4.46

$$P(i, j, \delta, 135^\circ) = \#\left\{ ((k, l), (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) \mid \begin{array}{l} k - m = \delta, l - n = \delta \\ k - m = -\delta, l - n = -\delta \end{array} \right\}$$

ec. 4.47

es el número de veces que los pixels con coordenadas (k, l) y (m, n) son vecinos. La matriz es simétrica, es decir: $P(i, j; d, a) = P(j, i; d, a)$

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

$$L_x = \{1, 2, 3, 4\} \quad L_y = \{1, 2, 3, 4\} \quad N_g = \{0, 1, 2, 3\}$$

a)

(N_g, N_g)	0	1	2	3
0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

b)

$$P_{0^\circ} = \begin{bmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad P_{45^\circ} = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$P_{90^\circ} = \begin{bmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix} \quad P_{135^\circ} = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

c)

fig 4.12

Para calcular la GLCM se puede tomar el ejemplo de la figura 4.12-a), la cual representa una imagen con 4 niveles de gris. La figura 4.12-b) es la forma general de cualquier matriz GLCM. Las matrices de co-ocurrencia de niveles de gris con $d=1$ y ángulos 0° , 45° , 90° y 135° se encuentran respectivamente en la figura 4.12-c). En caso de que el número total de niveles de gris fuera de 256, entonces la matriz sería de 256×256 .

La GLCM consiste en una matriz que proporciona la probabilidad conjunta de todos los pares de pixels dentro de una ventana definida para su cálculo. La co-ocurrencia de los niveles de gris representan la probabilidad que cualquier par de pixels con niveles de gris que satisfacen la distancia interpixel y orientación usados para el cálculo.

El algoritmo para el cálculo de la GLCM sería el siguiente:

1. Localizar el píxel con nivel de gris i y nivel de gris j que cumplan con la distancia interpixel y el ángulo de separación entre ellos.
2. Aumentar en la unidad el elemento de la i -ésima fila y j -ésima columna de la matriz $P(i,j)$.
3. Repetir hasta que se terminen todos los puntos.

4.2.1 Descriptores de textura

Se considera que toda la información de la textura está contenida en la GLCM, por lo que la aplicación de los descriptores de textura [HARA73] se hará directamente sobre ésta, a continuación se presentan los descriptores de textura:

1)Energía:

$$f_1 = \sum_i \sum_j p(i, j)^2$$

ec. 4.48

2)Contraste:

$$f_2 = \sum_i \sum_j (i - j)^2 \cdot p(i, j)$$

ec. 4.49

3)Correlación:

$$f_3 = \sum_i \sum_j \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} \cdot p(i, j)$$

ec. 4.50

4)Homogeneidad:

$$f_4 = \sum_{i=0} \sum_{j=0} \frac{1}{1 + (i - j)^2} \cdot p(i, j)$$

ec. 4.51

5) Entropía:

$$f_5 = -\sum_i \sum_j p(i, j) \cdot \log(p(i, j))$$

ec. 4.52

6) Autocorrelación:

$$f_6 = \sum_i \sum_j (ij) \cdot p(i, j)$$

ec. 4.53

7) Disimilaridad;

$$f_7 = \sum_i \sum_j |i - j| \cdot \log(p(i, j))$$

ec. 4.54

8) Sombra de Crupo:

$$f_8 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 \cdot p(i, j)$$

ec. 4.55

9) Prominencia de Grupo:

$$f_9 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 \cdot p(i, j)$$

ec. 4.56

10) Máxima Probabilidad:

$$f_{10} = \text{MAX}_{i,j} p(i, j)$$

ec. 4.57

donde:

$$\mu_x = \sum_i \sum_j i \cdot p(i, j)$$

ec. 4.58

$$\mu_y = \sum_j \sum_i j \cdot p(i, j)$$

ec. 4.59

$$\sigma^2_x = \sum_i \sum_j (i - \mu_x)^2 \cdot p(i, j)$$

ec. 4.60

$$\sigma^2_y = \sum_j \sum_i (j - \mu_y)^2 \cdot p(i, j)$$

ec. 4.61

4.2.2 Parámetros en el cálculo de la GLCM

Para el cálculo de la GLCM, existen varios parámetros a considerarse [SOH99]:

a) Cuantización: el número de niveles de gris es un importante factor para el diseño y cálculo de la GLCM. La decisión que se debe tomar es el número de niveles de gris necesarios para representar la textura con los descriptores. Entre más niveles de gris se involucren en el cálculo de la GLCM, más precisa será la información textural obtenida y como consecuencia, el costo computacional será más elevado. Algunos de los principales esquemas son:

1. Uniforme: este esquema es el más simple, los niveles de gris son cuantizados en segmentos con ancho de la misma magnitud entre sí, en este caso no se toma en cuenta la distribución de los niveles de gris, por lo que si en alguna zona del histograma algunos niveles de gris son mas frecuentes que otros, se desearía intervalos mas pequeños.
2. Gaussiana: se asume que la distribución del histograma es de naturaleza gaussiana, por lo que cada nivel de cuantización tiene la misma área bajo la curva, siendo el ancho del nivel en la parte central mas angosto, en comparación con el ancho de los niveles de cuantización en las colas, que son mas largos.

b) Desplazamiento: el parámetro desplazamiento o distancia δ es importante en el cálculo de la GLCM. Si a una textura fina se aplica un valor grande para el desplazamiento, la GLCM resultante no capturará adecuadamente la información textural y viceversa. El desplazamiento con valor igual al tamaño del elemento textural, es el más adecuado para el cálculo de la GLCM.

c) Tamaño de ventana: el tamaño de ventana comprende el área sobre el cual se calculará la GLCM, si la ventana es muy chica, la textura no podrá ser trabajada apropiadamente con los descriptores de textura y si es muy grande, los detalles de la imagen que sean menores que el tamaño de la ventana se perderán.

d) Orientación: el parámetro de orientación es relativamente menos importante comparado con los demás, la dependencia angular presentan un problema especial. Supongamos que la imagen A tiene los descriptores de textura a , b , c y d para los ángulos de 0° , 45° , 90° y 135° respectivamente, la imagen B es idéntica con A con la excepción de que B esta rotada 90° con respecto a A , entonces B tendrá los descriptores de textura respectivos c , d , a y b .

4.3 ALGORITMOS RÁPIDOS PARA EL CÁLCULO DE GLCM

El cálculo para determinar los descriptores de textura por medio de la GLCM implican un gasto computacional intenso debido al tamaño de la matriz y que algunos elementos de la matriz son ceros, por lo que se llevan a cabo diversas operaciones innecesarias. Por una parte, una mejora del método de la GLCM es la *Grey-Level Co-ocurrence Linked List (GLCLL)* y la *Grey-Level Co-ocurrence Hybrid Structure (GLHS)*. [CLAU01]

4.3.1 Grey Level Co-occurrence Linked List (GLCLL)

La orientación entre pixels vecinos para el cálculo de la GLCM comúnmente se fija en los valores de 0° , 45° , 90° y 135° . Recordando la naturaleza simétrica de la matriz GLCM, solo los elementos $i \geq j$ de las parejas correspondientes se consideran. En la fig. 4.13-a) se muestra una imagen de 4 niveles de gris de 5×5 . En la fig. 4.13-b) se encuentra la GLCM, para $\theta = 0^\circ$ y 180° , $\delta = 1$, como matriz triangular inferior.

La GLCLL se utiliza para almacenar los elementos de la GLCM, para facilitar la búsqueda y almacenamiento sucesivo del número correspondiente de veces en que los pixels con nivel de gris i y j aparecen, la GLCLL debe estar ordenada.

La búsqueda comienza desde la cabeza de la lista buscando primero el nivel de gris i , si le encuentra, entonces el algoritmo busca el nivel de gris j . Una vez localizado el par de niveles de gris (i,j) se actualiza el valor contenido. En dado caso de que el nodo (i,j) no se encuentre, deberá ser creado y añadido a la lista, con la probabilidad correspondiente. Un punto a tomar en cuenta para facilitar el manejo de la lista, es que ésta debe estar ordenada. En caso de que las parejas de gris (i,j) , $i \leq j$, simplemente se intercambian entre sí los valores de i y j . La figura 4.13-c) se ilustra el esquema general para la GLCLL.

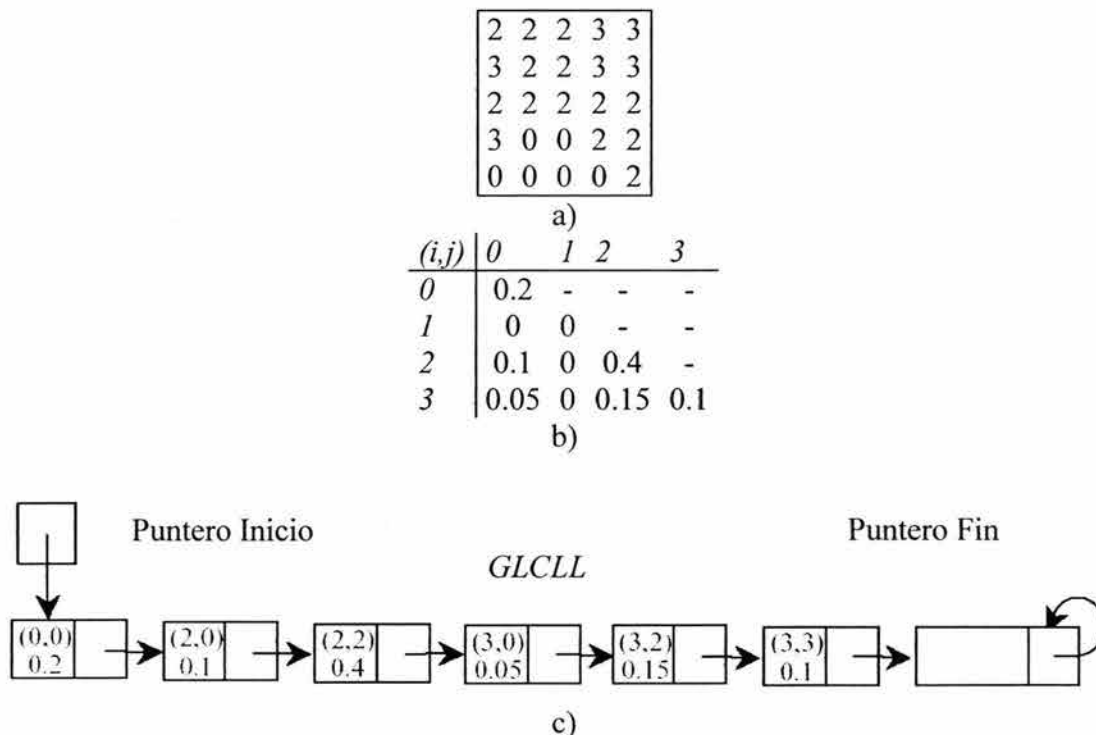


fig 4.13

4.3.2 Grey Level Co-occurrence Hybrid Structure (GLCHS)

Para mejorar la GLCLL, la condición de orden entre los nodos se eliminará, sustituyéndose por una tabla raíz y una lista doblemente ligada, basándose en esta combinación, la estructura se referirá como *Grey-Level Co-occurrence Hybrid Structure* (fig. 4.14).

Usando la GLCHS, la tabla raíz es creada para apuntar a los nodos de la lista. Los nodos de la lista doblemente ligada es usada para permitir la fácil inserción y borrado de los nodos. La tabla raíz permite el acceso rápido de cualquier nodo en la lista, si es que este nodo existe. La lista doblemente ligada permite el cálculo rápido del descriptor de textura, facilitando el recorrido de los nodos. Retomando el ejemplo de la fig. 4.13-a), los nodos con valor diferente de cero están indicados en color gris en la fig. 4.14.

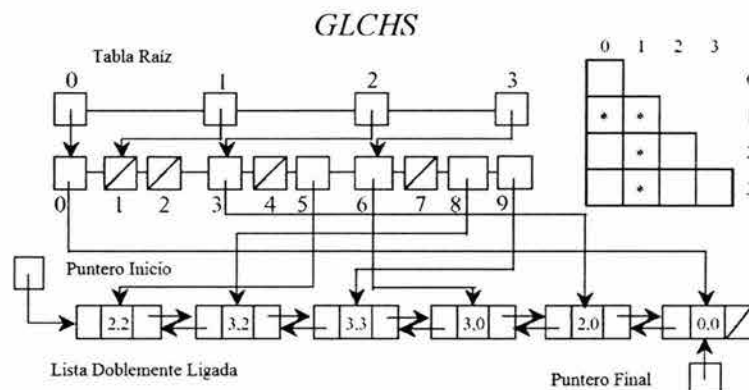


fig. 4.14

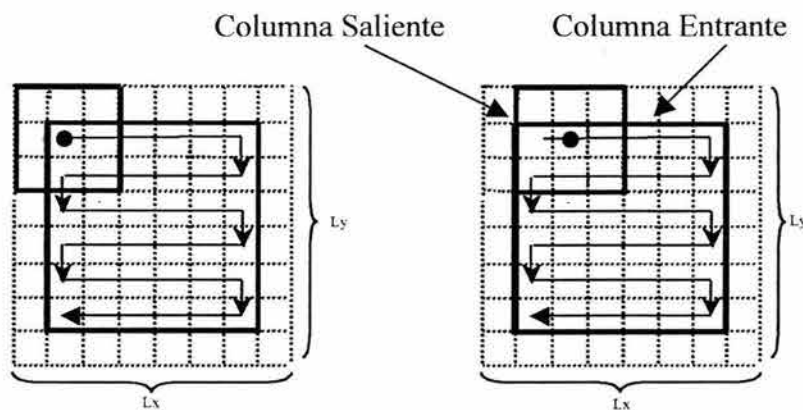


fig 4.15

Conforme la ventana va recorriendo los diferentes pixels de la imagen, para determinar los descriptores de textura asociados a ese píxel, muchas de la probabilidades dentro de la ventana se repiten, por lo que para disminuir la carga computacional se restan las probabilidades asociadas a la columna saliente y se suman las probabilidades asociadas con la columna entrante, la ventaja de este método se refleja para ventanas de gran tamaño. De manera análoga, el procedimiento anterior se aplicaría para los renglones al realizar un recorrido en zig-zag sobre la imagen (fig. 4.15), lo cual mejora el procedimiento.

5

FUNCIONES DISCRIMINANTES ESTADÍSTICAS

5.1 GENERALIDADES

Consideremos que en el universo de los objetos o entes de nuestro interés éstos poseen un conjunto de características reunidas en el vector $\mathbf{y} \in E$, donde las funciones discriminantes estadísticas asumen que las representaciones de los objetos en E *siguen ciertas leyes estadísticas*, cuyas formas funcionales pueden ser conocidas. Las funciones discriminantes estadísticas se definen en base a estas funciones y el aprendizaje consiste en la estimación de los parámetros que las describan.

Una regla de clasificación o regla de decisión consiste en que si \mathbf{y} es el vector de características de un objeto x ($\mathbf{y} = \mathbf{y}(x)$), x será asignado a una clase j tal que el valor de la función discriminante $g(\mathbf{y})$ cumpla con la siguiente condición:

$$g_j(\mathbf{y}) > g_i(\mathbf{y}) \quad \forall i \neq j$$

ec. 5.1.1

es decir,

$$x \in \varpi_j \mid \arg \max_{1 \leq j \leq C} g_j(\mathbf{y}(x))$$

ec. 5.1.2

un ejemplo de función discriminante es la probabilidad a posteriori:

$$g_\varpi(\mathbf{y}) = P(\varpi \mid \mathbf{y}) \quad 1 \leq \varpi \leq C$$

ec. 5.1.3

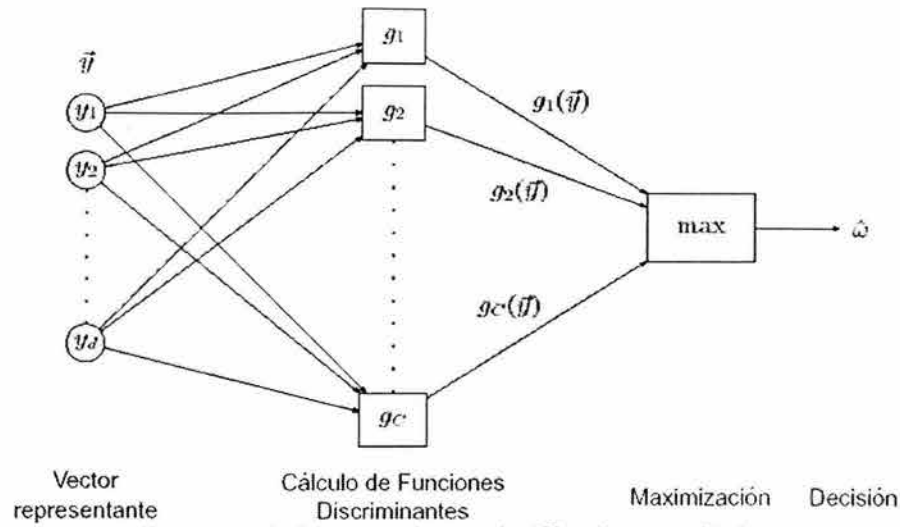


diagrama de bloques de un clasificador estadístico
fig. 5.1

Un clasificador divide el espacio de representación en C regiones de decisión, R_1, \dots, R_C :

$$R_j = \{ \mathbf{y} \in E \mid g_j(\mathbf{y}) > g_i(\mathbf{y}), i \neq j, 1 \leq i \leq C \}$$

ec. 5.1.4

La frontera de decisión entre dos clases i, j está dada por el lugar geométrico de los puntos $\mathbf{y} \in E$ para los que y en general son hipersuperficies definidas por las ecuaciones:

$$g_i(\mathbf{y}) - g_j(\mathbf{y}) = 0 \quad i \neq j \quad 1 \leq i, j \leq C$$

ec. 5.1.5

- Si $E = \mathbb{R}^3$ las fronteras son superficies.
- Si $E = \mathbb{R}^2$ las fronteras son líneas.
- Si $E = \mathbb{R}$ las fronteras son puntos

Las funciones discriminantes son lineales cuando son combinaciones lineales de las componentes del vector de características, de la forma:

$$g(\mathbf{y}) = \sum_{i=1}^d a_i y_i + a_0 = \mathbf{a}^T \mathbf{y} + a_0$$

ec. 5.1.6

donde \mathbf{a} se denomina vector de pesos y a_0 , peso umbral, las fronteras de decisión son hiperplanos.

Las funciones discriminantes cuadráticas son combinaciones cuadráticas de las componentes del vector de características:

$$g(\mathbf{y}) = \sum_{i=1}^d \sum_{j=1}^d a_{ij} y_i y_j + \sum_{i=1}^d a_i y_i + a_0 = \mathbf{y}^T \mathbf{A} \mathbf{y} + \mathbf{a}^T \mathbf{y} + a_0$$

ec. 5.1.7

donde \mathbf{A} es la matriz de pesos y las frontera de decisión son hipercuadráticas, por otra parte, se dice que dos clasificadores son equivalentes si

$$g_i(\mathbf{y}) > g_j(\mathbf{y}) \Leftrightarrow g'_i(\mathbf{y}) > g'_j(\mathbf{y})$$

$$\forall j \neq i, \forall \mathbf{y} \in E$$

ec. 5.1.8

Es decir, sea $f: \mathbb{R} \rightarrow \mathbb{R}$ cualquier función monótona creciente. Entonces los siguientes clasificadores son equivalentes (g_1, \dots, g_c) y $(f(g_1), \dots, f(g_c))$.

5.2 TEORÍA DE BAYES

La probabilidad de que un evento B se de cuando se sabe que algún otro evento A se ha presentado se llama *probabilidad condicional* y se le escribe $P(B|A)$. Esta expresión, por lo común se lee como: “la probabilidad de que B ocurra dado que ocurrió A ” o también, “la probabilidad de B , dado A ”.

La probabilidad condicional se define:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \text{ si } P(A) > 0$$

ec. 5.2.1

Además de que la probabilidad condicional permite una alteración de la probabilidad de un evento a la luz de mayor información, también da lugar a que se entienda mejor el importante concepto de independencia o, en el contexto actual, el de eventos independientes. Considerando la situación en la cual se tienen eventos A y B , donde B sucede, no tiene impacto en que se presente A . Es decir, la ocurrencia de A es independiente de la ocurrencia de B .

Dos eventos A y B son independientes, sí y sólo sí:

$$P(B|A) = P(B)$$

ec. 5.2.2

$$P(A|B) = P(A)$$

ec. 5.2.3

que también se puede expresar de la siguiente forma:

$$P(A \cap B) = P(A)P(B)$$

ec. 5.2.4

Al multiplicar la fórmula de la ec. 5.2.1 por $P(A)$, se obtiene la regla multiplicativa que permite calcular la probabilidad de que dos eventos ocurran, es decir, si en un experimento pueden ocurrir los eventos A y B, entonces:

$$P(A \cap B) = P(B \cap A) = P(A)P(A | B)$$

ec. 5.2.5

El Teorema de Probabilidad Total se enuncia de la siguiente manera:

Si los eventos B_1, B_2, \dots, B_k constituyen una división del espacio muestral S , de tal forma que $P(B_i) \neq 0$, para $i=1, 2, \dots, k$, entonces para cualquier evento A de S .

$$P(A) = \sum_{i=1}^k P(B_i \cap A) = \sum_{i=1}^k P(B_i)P(A | B_i)$$

ec. 5.2.6

Si los eventos B_1, B_2, \dots, B_k constituyen una división del espacio muestral S , de tal forma que $P(B_i) \neq 0$, para $i=1, 2, \dots, k$, entonces para cualquier evento A de S es tal que $P(A) \neq 0$ [WALL]

$$P(B_r | A) = \frac{P(B_r \cap A)}{P(A)} = \frac{P(B_r \cap A)}{\sum_{i=1}^k P(B_i \cap A)} = \frac{P(B_r)P(A | B_r)}{\sum_{i=1}^k P(B_i)P(A | B_i)}$$

ec. 5.2.7

para $r=1, 2, \dots, k$.

El teorema de Bayes aplicado y su aplicación al reconocimiento estadístico de clases, se puede expresar de la siguiente manera:

$$P(\alpha_i | \mathbf{x}) = \frac{P(\alpha_i)P(\mathbf{x} | \alpha_i)}{P(\mathbf{x})}$$

ec. 5.2.8

En donde cada uno de los elementos de esta igualdad tiene el siguiente significado:

$P(\alpha_i | \mathbf{x})$: Probabilidad de que un vector de características \mathbf{x} pertenezca a la clase α_i .

$P(\mathbf{x} | \alpha_i)$: Probabilidad de que cada α_i , el valor de la variable aleatoria sea, precisamente \mathbf{x} .

$P(\alpha_i)$: Probabilidad *a priori* de que se presente un elemento de la clase α_i .

$P(\mathbf{x})$: Probabilidad *a priori* de que se presente un objeto a clasificar con un vector de características igual a \mathbf{x} . Este último elemento puede despreciarse, ya que representa el mismo valor para un conjunto de N clases, $\alpha_1, \alpha_2, \dots, \alpha_N$. Se trata de un factor de escala, que no proporciona información discriminante.

El primer término de la ec. 5.2.8 aporta la solución al problema de clasificación. Dado un vector \mathbf{x} , ¿a qué clase pertenece?, la clasificación será la siguiente:

$$\mathbf{x} \in \alpha_j \text{ si } P(\alpha_j | \mathbf{x}) > P(\alpha_i | \mathbf{x})$$

$$\forall i \neq j \quad i=1,2,\dots,N$$

ec. 5.2.9

lo cual se puede ver gráficamente de la siguiente manera:

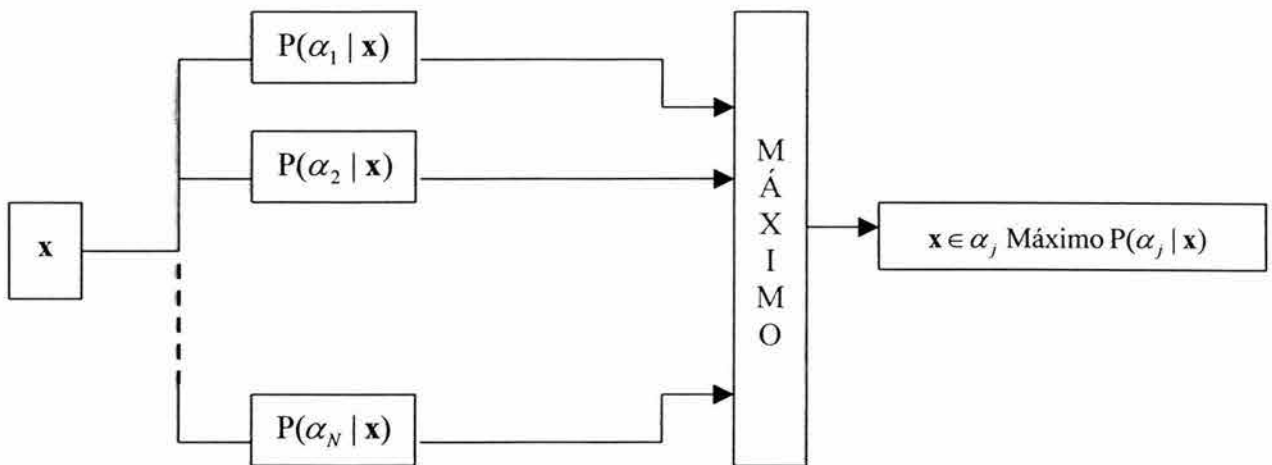


diagrama del clasificador estadístico *a posteriori*.

fig. 5.2

Si ahora nos basamos en el segundo miembro de la ec. 5.2.8 del teorema de Bayes (eliminando el factor de escala $P(\mathbf{x})$) tendremos una forma alternativa de clasificar un vector \mathbf{x} .

$$\mathbf{x} \in \alpha_j \text{ si } P(\alpha_j | \mathbf{x})P(\alpha_j) > P(\alpha_i | \mathbf{x})P(\alpha_i)$$

$$\forall i \neq j \quad i=1,2,\dots,N$$

ec. 5.2.10

lo cual se puede ver gráficamente de la siguiente manera:

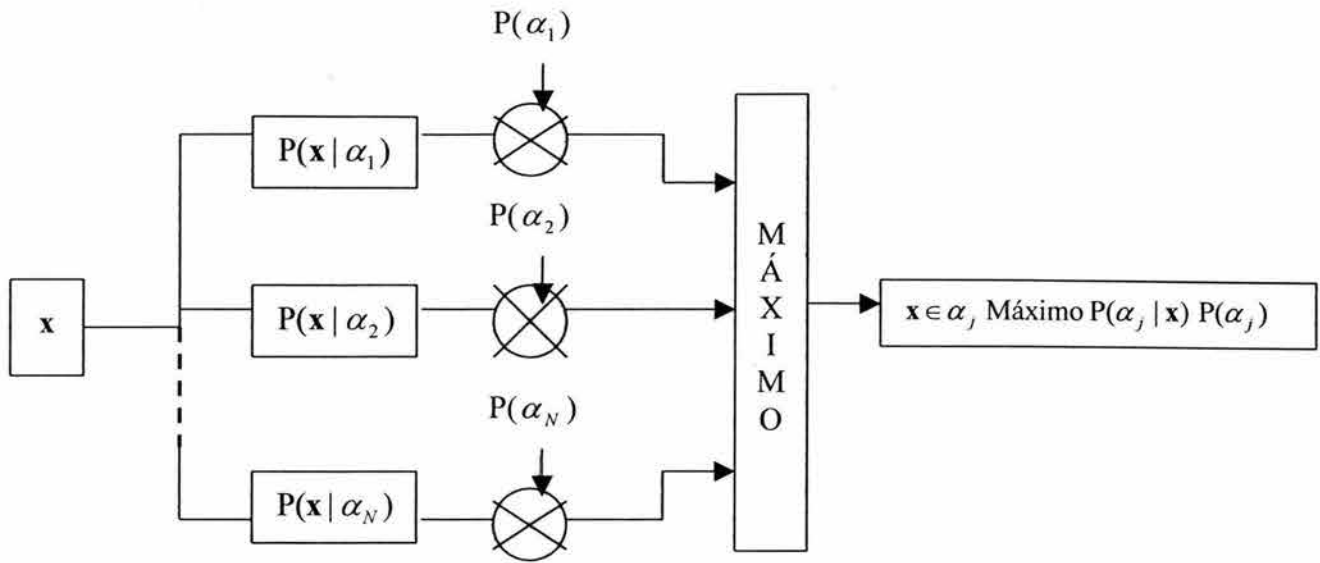


diagrama del clasificador estadístico *a priori*.
fig. 5.3

Hemos añadido el calificativo de *a posteriori* al clasificador estadístico basado en el primer miembro del teorema de Bayes y, análogamente, el calificativo de *a priori* al basado en el segundo miembro. La justificación es inmediata, en el primer caso se trata de calcular, o al menos estimar, la probabilidad *a posteriori* de que \mathbf{x} pertenezca a la clase α_i , mientras que en el segundo caso se trata de la probabilidad *a priori* de que cumpliéndose la hipótesis de pertenencia a la clase α_i el objeto sea \mathbf{x} .

La estimación de las probabilidades del clasificador estadístico *a posteriori* sólo se puede realizar mediante un proceso de aprendizaje $P(\mathbf{x} | \alpha_1), P(\mathbf{x} | \alpha_2), \dots, P(\mathbf{x} | \alpha_N)$ a partir de un conjunto de muestras o realizaciones físicas de las clases

En cuanto al otro procedimiento, el principal problema para el diseño del clasificador *a priori* es la estimación estadística de las funciones de probabilidad de las clases $\alpha_1, \alpha_2, \dots, \alpha_N$ consideradas como variables aleatorias.

5.3 FUNCIONES DISCRIMINANTES

El uso de funciones discriminantes estadísticas para la clasificación es ventajosa por dos razones, que a continuación se exponen:

1. El conocimiento que a lo largo del tiempo se ha adquirido en áreas como teoría de la decisión y comunicaciones es aplicable directamente a el reconocimiento de patrones.

2. La formulación estadística es particularmente ajustable para el problema del reconocimiento de patrones, ya que este tratamiento es de naturaleza estadística. En este problema es deseable el uso de toda la información *a priori* disponible.

5.3.1 Función pérdida

Antes de establecer la función pérdida, los siguientes puntos deben ser tomados en cuenta:

1. $P(\varpi_i)$ es conocida o puede ser estimada.
2. $P(\mathbf{x} | \varpi_i)$ es conocida o puede ser estimada por medio del entrenamiento del sistema.
3. $P(\varpi_i | \mathbf{x})$ no es conocida.

La función pérdida L_{ij} , definida como la pérdida, costo o multa, permitirá establecer el criterio para el cual $\mathbf{x} \in \varpi_i$, cuando de hecho $\mathbf{x} \in \varpi_j$. Similarmente, la pérdida condicional promedio o el riesgo condicional promedio $r_k(\mathbf{x})$ puede ser definido como:

$$r_k(\mathbf{x}) = \sum_{i=1}^M L_{ik} P(\varpi_i | \mathbf{x})$$

ec. 5.3.1

que es el promedio o la pérdida que se espera obtener al clasificar erróneamente \mathbf{x} perteneciente a ϖ_k , cuando de hecho, podría pertenecer a otra clase $\varpi_i, i=1,2,\dots,M, i \neq k$.

El trabajo del clasificador es entonces, encontrar la decisión óptima que minimice el riesgo o pérdida promedio. La regla de decisión consistirá de los siguientes pasos:

1. Calcular $r_i(\mathbf{x})$ para $\mathbf{x} \in \varpi_i, \forall i, i=1,2,\dots,M$.
2. Decidir si $\mathbf{x} \in \varpi_k$, siempre y cuando $r_k(\mathbf{x}) \leq r_i(\mathbf{x}), \forall i, i \neq k$.

La función discriminante estadística correspondiente es:

$$d_k(\mathbf{x}) = -r_k(\mathbf{x})$$

ec. 5.3.2

El signo negativo de $r_k(\mathbf{x})$ se escoge para hacer que $d_k(\mathbf{x})$ represente la clase más parecida. Entre menor sea el valor de $r_k(\mathbf{x})$, más acertado el que $\mathbf{x} \in \varpi_k$.

La matriz de pérdida se puede representar de la siguiente forma:

5.3.2 Función discriminante de Bayes

Recordando el Teorema de Bayes:

$$P(\varpi_i | \mathbf{x}) = \frac{P(\varpi_i)P(\mathbf{x} | \varpi_i)}{P(\mathbf{x})}$$

ec. 5.3.6

sustituyendo la ec. 5.3.6 en ec. 5.3.1:

$$r_k(\mathbf{x}) = \frac{1}{P(\mathbf{x})} \sum_{i=1}^M L_{ik} P(\mathbf{x} | \varpi_i) P(\varpi_i)$$

ec. 5.3.7

El término $P(\mathbf{x})$ es el mismo para todos los valores de k , por lo que se puede ignorar de la ec. 5.3.7 y buscar el siguiente mínimo:

$$\min_k r_k(\mathbf{x}) = \min_k \sum_{i=1}^M L_{ik} P(\mathbf{x} | \varpi_i) P(\varpi_i)$$

ec. 5.3.8

para obtener el mejor entre todos los posibles, alternativamente se puede decir que:

$$d_k(\mathbf{x}) = -r_k(\mathbf{x})$$

ec. 5.3.9

que es función discriminante de Bayes. El clasificador basado en esta minimización es llamado clasificador Bayesiano.

5.3.3 Decisión de máxima probabilidad

Para el caso de dos clases ($M=2$) la pérdida condicional promedio de que $\mathbf{x} \in \varpi_1$, (ec. 5.3.7) toma la siguiente forma:

$$r_1(\mathbf{x}) = L_{11}P(\mathbf{x} | \varpi_1)P(\varpi_1) + L_{21}P(\mathbf{x} | \varpi_2)P(\varpi_2)$$

ec. 5.3.10

de manera similar, en el caso de que $\mathbf{x} \in \varpi_2$

$$r_2(\mathbf{x}) = L_{12}P(\mathbf{x} | \varpi_1)P(\varpi_1) + L_{22}P(\mathbf{x} | \varpi_2)P(\varpi_2)$$

ec. 5.3.11

expresado de forma matricial

$$\begin{pmatrix} r_1(\mathbf{x}) \\ r_2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} L_{11} & L_{21} \\ L_{12} & L_{22} \end{pmatrix} \begin{pmatrix} P(\mathbf{x} | \varpi_1)P(\varpi_1) \\ P(\mathbf{x} | \varpi_2)P(\varpi_2) \end{pmatrix}$$

ec. 5.3.12

si $\mathbf{x} \in \varpi_1$, entonces

$$r_1(\mathbf{x}) < r_2(\mathbf{x})$$

ec. 5.3.14

$$L_{11}P(\mathbf{x} | \varpi_1)P(\varpi_1) + L_{21}P(\mathbf{x} | \varpi_2)P(\varpi_2) < L_{12}P(\mathbf{x} | \varpi_1)P(\varpi_1) + L_{22}P(\mathbf{x} | \varpi_2)P(\varpi_2)$$

ec. 5.3.15

rescribiendo la ecuación anterior de la siguiente manera

$$\frac{P(\mathbf{x} | \varpi_1)}{P(\mathbf{x} | \varpi_2)} > \frac{(L_{21} - L_{22})P(\varpi_1)}{(L_{12} - L_{11})P(\varpi_2)}$$

ec. 5.3.16

de donde se define la relación de probabilidad

$$l_{12} = \frac{P(\mathbf{x} | \varpi_1)}{P(\mathbf{x} | \varpi_2)}$$

ec. 5.3.17

y el valor de umbral

$$\theta_{12} = \frac{(L_{21} - L_{22})P(\varpi_1)}{(L_{12} - L_{11})P(\varpi_2)}$$

ec. 5.3.18

el criterio de decisión toma la siguiente forma

$$\mathbf{x} \in \varpi_1 \text{ si } l_{12} > \theta_{12}$$

ec. 5.3.19

la relación generalizada de probabilidad y el valor generalizado de umbral son respectivamente

$$l_{ki} = \frac{P(\mathbf{x} | \varpi_k)}{P(\mathbf{x} | \varpi_i)}$$

ec. 5.3.20

$$\theta_{ki} = \frac{(L_{ik} - L_{ii})P(\varpi_i)}{(L_{ki} - L_{kk})P(\varpi_k)}$$

ec. 5.3.21

por lo que el criterio de decisión es:

$$\boxed{} \text{ si } l_{ki} > \theta_{ki}$$

ec. 5.3.22

5.3.4 Función discriminante óptima

La función normal de densidad para M clases diferentes se representa de la siguiente forma:

$$P(\mathbf{x} | \varpi_k) = \frac{1}{\sqrt{2\pi^n} |\mathbf{C}_k|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mathbf{m}_k)^T \mathbf{C}_k^{-1}(\mathbf{x} - \mathbf{m}_k)\right]$$

ec. 5.3.23

$$P(\mathbf{x} | \varpi_k) = \mathcal{N}(\mathbf{x}, \mathbf{m}_k, \mathbf{C}_k) \quad k=1,2,\dots,M$$

$n = \text{dimensión del vector patrón}$
ec. 5.3.24

donde \mathcal{N} es la función de densidad normal, \mathbf{m}_k es el vector de medias y \mathbf{C}_k es la matriz de covarianza para la clase k .

$$\mathbf{m}_k = E_k[\mathbf{x}]$$

ec. 5.3.25

$$\mathbf{C}_k = E_k[(\mathbf{x} - \mathbf{m}_k)(\mathbf{x} - \mathbf{m}_k)^T]$$

ec. 5.3.26

La localización del centro del conjunto de vectores pertenecientes a una clase está determinado por el vector de medias y la forma del cluster está determinado por la matriz de covarianza. Sin pérdida de generalidad, para $n=2$, espacio bidimensional, podemos ver los siguientes ejemplos y la forma que toman sus respectivas matrices de covarianza:

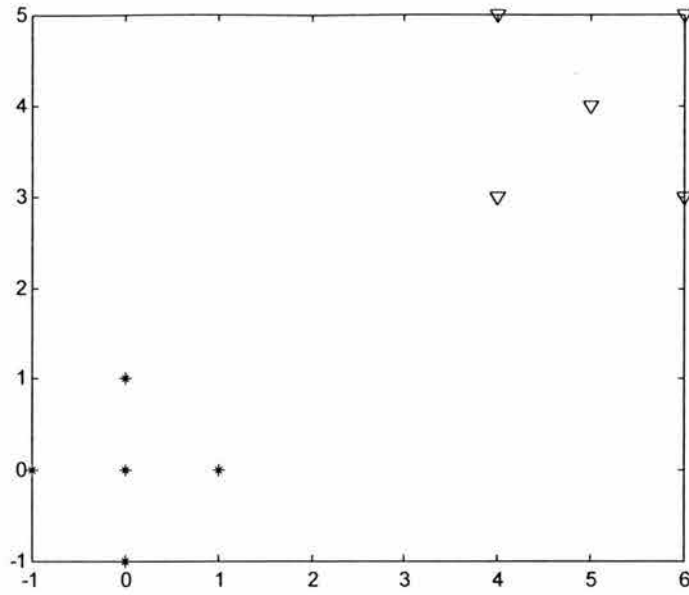


fig. 5.6

$$\begin{array}{l}
 \textit{clase 1} \\
 \left\{ \begin{array}{l} \mathbf{m}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \mathbf{C}_1 = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix} \end{array} \right.
 \end{array}
 \qquad
 \begin{array}{l}
 \textit{clase 2} \\
 \left\{ \begin{array}{l} \mathbf{m}_1 = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \\ \mathbf{C}_1 = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix} \end{array} \right.
 \end{array}$$

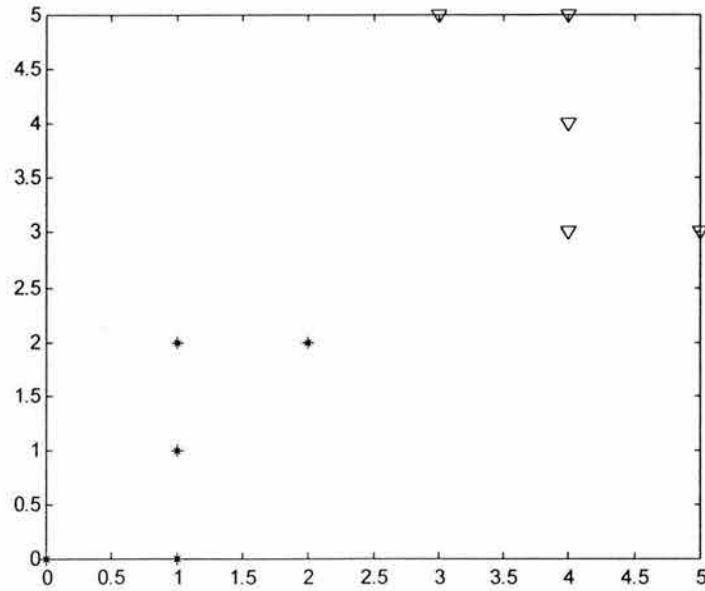


fig. 5.7

$$\begin{array}{l}
 \text{clase 1} \\
 \left\{ \begin{array}{l} \mathbf{m}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \mathbf{C}_1 = \begin{pmatrix} 0.4 & 0.4 \\ 0.4 & 0.8 \end{pmatrix} \end{array} \right.
 \end{array}
 \qquad
 \begin{array}{l}
 \text{clase 2} \\
 \left\{ \begin{array}{l} \mathbf{m}_1 = \begin{pmatrix} 4 \\ 4 \end{pmatrix} \\ \mathbf{C}_1 = \begin{pmatrix} 0.4 & -0.4 \\ -0.4 & 0.8 \end{pmatrix} \end{array} \right.
 \end{array}$$

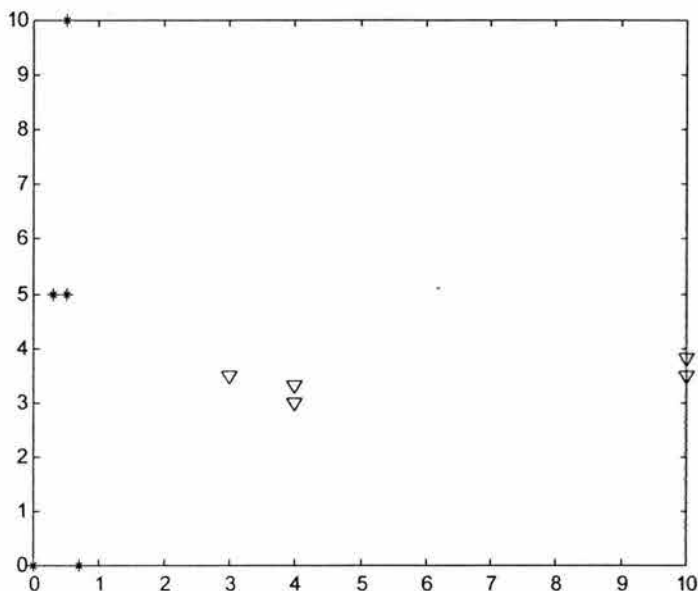


fig. 5.8

$$\begin{array}{l}
 \text{clase 1} \\
 \left\{ \begin{array}{l} \mathbf{m}_1 = \begin{pmatrix} 0.4 \\ 4 \end{pmatrix} \\ \mathbf{C}_1 = \begin{pmatrix} 0.056 & 0.2 \\ 0.2 & 14 \end{pmatrix} \end{array} \right.
 \end{array}
 \qquad
 \begin{array}{l}
 \text{clase 2} \\
 \left\{ \begin{array}{l} \mathbf{m}_1 = \begin{pmatrix} 6.2 \\ 3.42 \end{pmatrix} \\ \mathbf{C}_1 = \begin{pmatrix} 9.76 & 0.53 \\ 0.53 & 0.69 \end{pmatrix} \end{array} \right.
 \end{array}$$

Existe una medida de similitud usada comúnmente, que es la *distancia de Mahalanobis*, la cual, queda definida entre un patrón \mathbf{x} y el vector de medias \mathbf{m} , como:

$$r = (\mathbf{x} - \mathbf{m})^T \mathbf{C}^{-1} (\mathbf{x} - \mathbf{m}) \tag{5.3.27}$$

Por otra parte, la distancia de Mahalanobis entre dos clases puede expresarse como:

$$r_{ij} = (\mathbf{m}_i - \mathbf{m}_j)^T \mathbf{C}^{-1} (\mathbf{m}_i - \mathbf{m}_j) \tag{5.3.28}$$

Como se vió anteriormente, se retoma la siguiente ecuación discriminante:

$$d'_k = \log p(\mathbf{x} | \varpi_k) + \log p(\varpi_k)$$

ec. 5.3.29

$$P(\mathbf{x} | \varpi_k) = \frac{1}{\sqrt{2\pi^n} |C_k|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{m}_k)^T C_k^{-1} (\mathbf{x} - \mathbf{m}_k) \right]$$

$k=1, 2, \dots, M$
ec. 5.3.30

sustituyendo la ec. 5.3.30 en la ec. 5.3.29, la función discriminante d'_k toma la siguiente forma

$$d''_k = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |C_k| - \frac{1}{2} (\mathbf{x} - \mathbf{m}_k)^T C_k^{-1} (\mathbf{x} - \mathbf{m}_k) + \log p(\varpi_k)$$

5.3.31

el primer término del lado derecho de la ecuación es el mismo para todos los valores de k , por lo que se puede describir de la siguiente manera

$$d'''_k = -\frac{1}{2} (\mathbf{x} - \mathbf{m}_k)^T C_k^{-1} (\mathbf{x} - \mathbf{m}_k) + \left[\log p(\varpi_k) - \frac{1}{2} \log |C_k| \right]$$

ec. 5.3.32

escrita de una forma más compacta

$$d^{(iv)}_k = -\frac{1}{2} r + f(k)$$

ec. 5.3.33

donde

$$r = (\mathbf{x} - \mathbf{m}_k)^T C_k^{-1} (\mathbf{x} - \mathbf{m}_k)$$

ec. 5.3.34

$$f(k) = \log p(\varpi_k) - \frac{1}{2} \log |C_k|$$

ec. 5.3.35

donde r , es la distancia de Mahalanobis, vista anteriormente.

La discusión sobre esta función discriminante presenta dos casos especiales.

Caso 1: Cuando las matrices de covarianza son iguales para diferentes clases ($C_i = C_j = C_k = C$). El significado físico sería representada por clusters con la misma forma y tamaño, pero con localización en diferentes lugares.

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

ec. 5.3.36

El primero y último términos del lado derecho de la ecuación no proporcionan información discriminante al no cambiar su valor para diferentes valores de k .

rescribiendo la ec. 5.3.36

$$d_k(\mathbf{x}) = \mathbf{x}^T \mathbf{C}^{-1} \mathbf{m}_k + [\log p(\varpi_k) - \frac{1}{2} \mathbf{m}_k^T \mathbf{C}^{-1} \mathbf{m}_k]$$

ec. 5.3.37

Para el caso de dos clases, $M=2$:

$$d(\mathbf{x}) = d_1(\mathbf{x}) - d_2(\mathbf{x})$$

ec. 5.3.38

$$d(\mathbf{x}) = \mathbf{x}^T \mathbf{C}^{-1} (\mathbf{m}_1 - \mathbf{m}_2) + \log p\left(\frac{\varpi_1}{\varpi_2}\right) - \frac{1}{2} (\mathbf{m}_1^T \mathbf{C}^{-1} \mathbf{m}_1 - \mathbf{m}_2^T \mathbf{C}^{-1} \mathbf{m}_2)$$

5.3.39

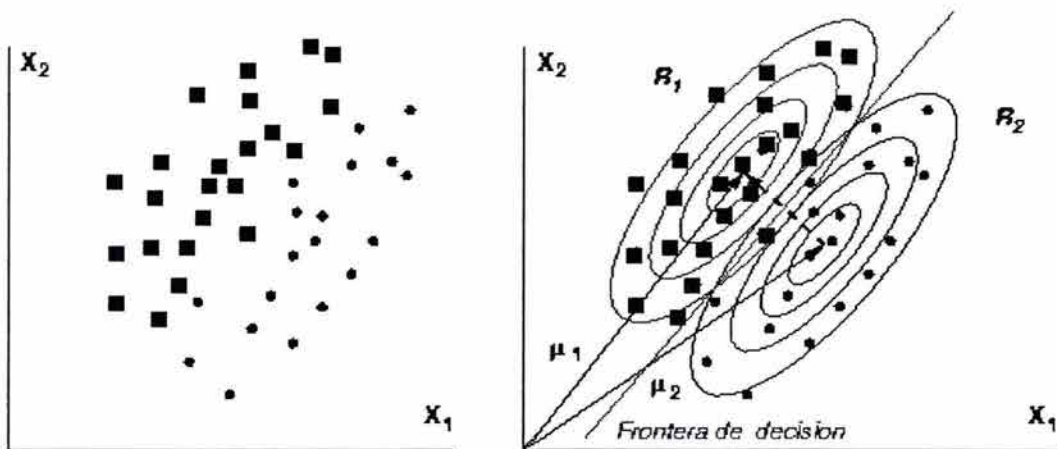


fig. 5.9

En este caso, las matrices de covarianza de todas las clases son iguales, las variables no son estadísticamente independientes (existe correlación).

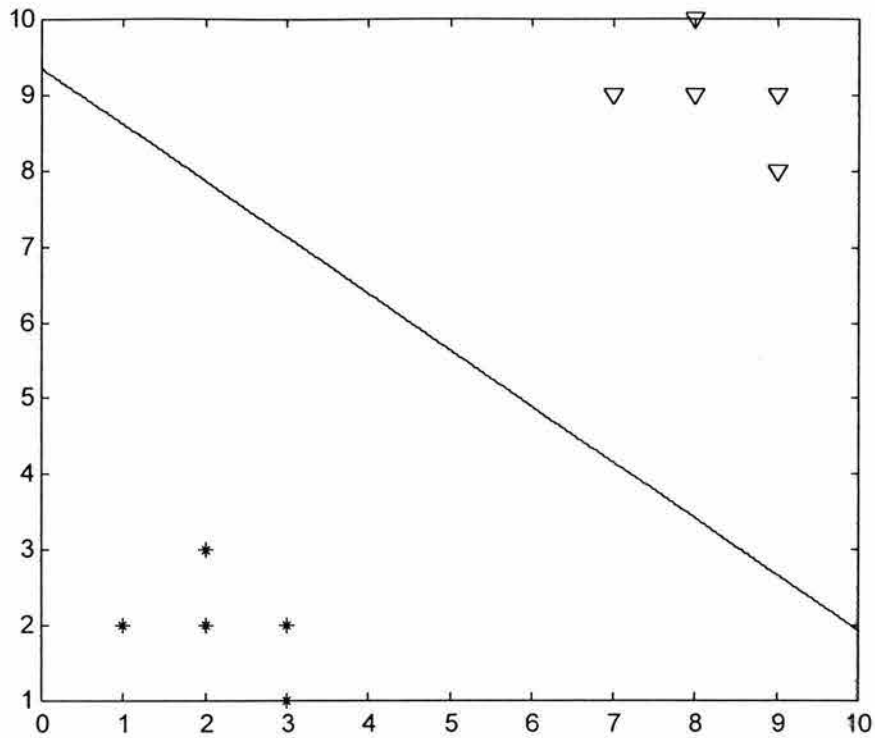


fig. 5.10

$$\begin{array}{l}
 \text{clase 1} \left\{ \begin{array}{l}
 \varpi_1 = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\} \\
 \mathbf{m}_1 = \begin{pmatrix} 2.2 \\ 2.0 \end{pmatrix} \\
 \mathbf{C}_1 = \begin{pmatrix} 0.56 & -0.20 \\ -0.20 & 0.40 \end{pmatrix}
 \end{array} \right.
 \end{array}
 \qquad
 \begin{array}{l}
 \text{clase 2} \left\{ \begin{array}{l}
 \varpi_2 = \left\{ \begin{pmatrix} 7 \\ 9 \end{pmatrix}, \begin{pmatrix} 8 \\ 9 \end{pmatrix}, \begin{pmatrix} 9 \\ 8 \end{pmatrix}, \begin{pmatrix} 9 \\ 9 \end{pmatrix}, \begin{pmatrix} 8 \\ 10 \end{pmatrix} \right\} \\
 \mathbf{m}_2 = \begin{pmatrix} 8.2 \\ 9 \end{pmatrix} \\
 \mathbf{C}_2 = \begin{pmatrix} 0.56 & -0.20 \\ -0.20 & 0.40 \end{pmatrix}
 \end{array} \right.
 \end{array}$$

$$d(\mathbf{x}) = -0.826x_1 - 1.11x_2 + 10.35$$

ec. 5.3.40

Caso 2: Cuando la matriz de covarianza \mathbf{C}_k es diagonal de la forma $\sigma_k^2 \mathbf{I}$, donde $\sigma_k^2 = |\mathbf{C}_k|$. El significado físico de este caso es representado por un cluster cuyos componentes son iguales a lo largo de los ejes principales y su distribución es de forma esférica.

Sustituyendo $\mathbf{C}_k = \sigma_k^2 \mathbf{I}$, con la misma varianza entre todas las clases y desarrollando los términos en los paréntesis de la ec. 5.3.32

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

ec. 5.3.41

eliminando los términos que no cambian para diferentes valores de k , así como considerando que las $p(\varpi_k)$ es igual para toda k , reescribiendo la ec. anterior:

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

ec. 5.3.42

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

ec. 5.3.43

La cual es una ecuación lineal.

Para el caso de dos clases, $M=2$:

$$d(\mathbf{x}) = d_1(\mathbf{x}) - d_2(\mathbf{x})$$

ec. 5.3.44

$$d(\mathbf{x}) = \mathbf{x}^T (\mathbf{m}_1 - \mathbf{m}_2) - \frac{1}{2} (|\mathbf{m}_1|^2 - |\mathbf{m}_2|^2)$$

ec. 5.3.45

Si retomamos la ec. 5.3.32 y bajo las consideraciones antes mencionadas el resultado sería:

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

ec. 5.3.46

o simplemente:

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

ec. 5.3.47

la cual es clasificador de distancia mínima.

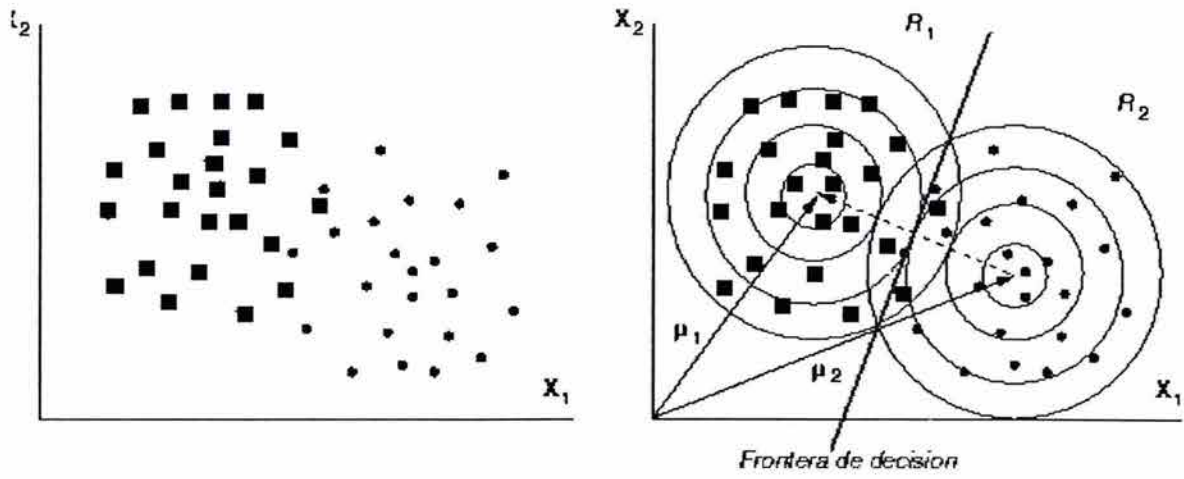


fig. 5.11

Este es el caso más simple, en el que las variables son estadísticamente independientes

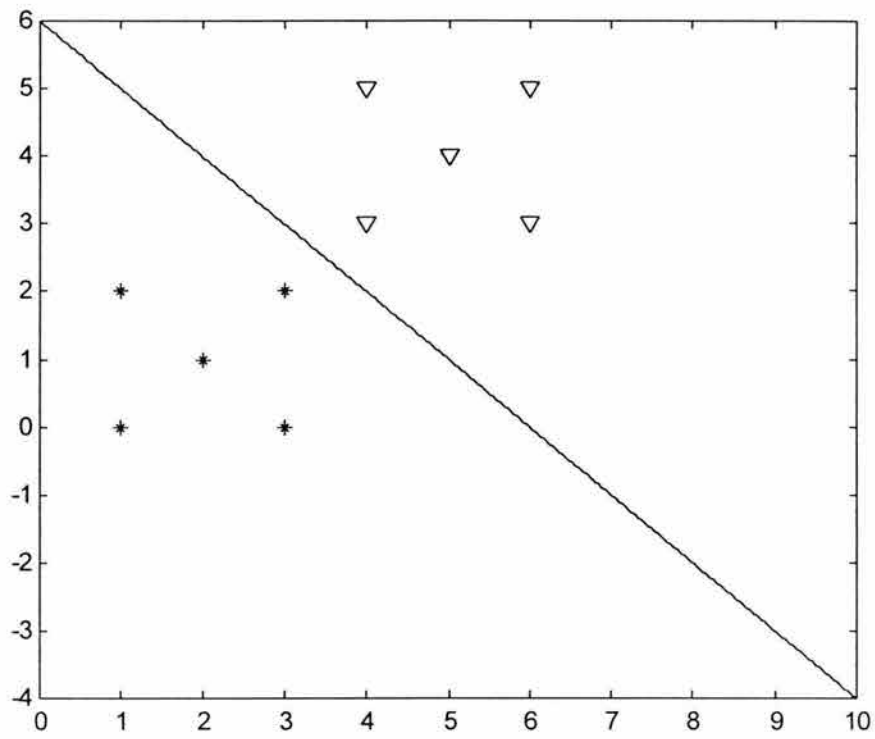


fig. 5.12

$$\begin{array}{l}
 \text{clase 1} \left\{ \begin{array}{l}
 \varpi_1 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\} \\
 \mathbf{m}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\
 \mathbf{C}_1 = \begin{pmatrix} 0.80 & 0 \\ 0 & 0.80 \end{pmatrix}
 \end{array} \right.
 \end{array}
 \qquad
 \begin{array}{l}
 \text{clase 2} \left\{ \begin{array}{l}
 \varpi_2 = \left\{ \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 6 \\ 3 \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \end{pmatrix} \right\} \\
 \mathbf{m}_2 = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \\
 \mathbf{C}_2 = \begin{pmatrix} 0.80 & 0 \\ 0 & 0.80 \end{pmatrix}
 \end{array} \right.
 \end{array}$$

$$d(\mathbf{x}) = -3.75x_1 - 3.75x_2 + 22.5$$

ec. 5.3.48

5.4 CASO BINARIO DE SEGMENTACIÓN

Para el caso binario de segmentación, encontramos que el vector \mathbf{x} tiene componentes binarios independientes como a continuación se muestra:

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$$

$$x_i = 1 \text{ ó } 0, i=1, 2, \dots, n$$

ec. 5.4.1

Para el caso particular de dos clases, la función discriminante $d(\mathbf{x})$ es:

$$d(\mathbf{x}) = d_1(\mathbf{x}) - d_2(\mathbf{x})$$

ec. 5.4.2

donde $d_1(\mathbf{x}) = \log[p(\mathbf{x}|\varpi_1)p(\varpi_1)]$ y $d_2(\mathbf{x}) = \log[p(\mathbf{x}|\varpi_2)p(\varpi_2)]$, entonces:

$$d(\mathbf{x}) = \boxed{\mathbf{x}}$$

ec. 5.4.3

Para el caso donde existen dos clases:

$$p(\varpi_1) + p(\varpi_2) = 1$$

ec. 5.4.4

por lo que la ecuación 5.4.2. queda como sigue:

$$d(\mathbf{x}) = \log \frac{p(\mathbf{x}|\varpi_1)}{p(\mathbf{x}|\varpi_2)} + \log \frac{p(\varpi_1)}{1 - p(\varpi_1)}$$

ec. 5.4.5

Los componentes x_i , son independientes, por lo que:

$$p(\mathbf{x} | \varpi_j) = p(x_1 | \varpi_j) p(x_2 | \varpi_j) \dots p(x_n | \varpi_j) = \prod_{i=1}^n p(x_i | \varpi_j)$$

ec. 5.4.6

de la ec. 5.4.4 y la ec. 5.4.5.

$$d(\mathbf{x}) = \sum_{i=1}^n \log \frac{p(\mathbf{x} | \varpi_1)}{p(\mathbf{x} | \varpi_2)} + \log \frac{p(\varpi_1)}{1 - p(\varpi_1)}$$

ec. 5.4.7

Por otra parte:

$$p(x_i=1 | \varpi_1) = p_i$$

ec. 5.4.8

$$p(x_i=0 | \varpi_1) = 1 - p_i$$

ec. 5.4.9

De manera similar:

$$p(x_i=1 | \varpi_2) = q_i$$

ec. 5.4.10

$$p(x_i=0 | \varpi_2) = 1 - q_i$$

ec. 5.4.11

Lo cual permite:

$$\log \frac{p(x_i | \varpi_1)}{p(x_i | \varpi_2)} = x_i \log \frac{p_i}{q_i} + (1 - x_i) \log \frac{1 - p_i}{1 - q_i} = \gamma$$

ec. 5.4.12

La validez de la ec. 5.4.12 se puede observar al sustituir x_i por 1 o 0.

Rescribiendo la ecuación 5.4.12:

$$\gamma = \log \frac{p(x_i | \varpi_1)}{p(x_i | \varpi_2)} = x_i \log \frac{p_i(1 - q_i)}{q_i(1 - p_i)} + \log \frac{1 - p_i}{1 - q_i}$$

ec. 5.4.13

Sustituyendo en la ec. 5.4.6:

$$d(\mathbf{x}) = \sum_{i=1}^n x_i \log \frac{p_i(1-q_i)}{q_i(1-p_i)} + \sum_{i=1}^n \log \frac{1-p_i}{1-q_i} + \log \frac{p(\varpi_1)}{1-p(\varpi_1)}$$

ec. 5.4.14

donde:

$$w_i = x_i \log \frac{p_i(1-q_i)}{q_i(1-p_i)}$$

ec 5.4.15

y

$$w_{n+1} = \sum_{i=1}^n \log \frac{1-p_i}{1-q_i} + \log \frac{p(\varpi_1)}{1-p(\varpi_1)}$$

ec 5.4.16

Finalmente, la ec.5.4.14 puede ser representada como:

$$d(\mathbf{x}) = \sum_{i=1}^n w_i x_i + w_{n+1}$$

ec. 5.4.17

la cual es una función lineal discriminante óptima en \mathbf{x} .

5.5 APROXIMACIÓN MÁXIMO A POSTERIORI-MAP

La *aproximación máximo a posteriori* (MAP) consiste en buscar la clase h (de entre todas las clases H posibles) más probable si hemos observado una serie de datos D .

$$h_{MAP} = \arg \max_{h \in H} P(h | D)$$

ec. 5.5.1

de la ec. 5.2.8 del teorema de Bayes:

$$P(h | D) = \frac{P(D | h)P(h)}{P(D)}$$

ec. 5.5.2

$$h_{MAP} = \arg \max_{h \in H} P(h | D)$$

ec. 5.5.3

ESTA TESIS NO SALE
DE LA BIBLIOTECA

$$h_{MAP} = \arg \max_{h \in H} \frac{P(D | h)P(h)}{P(D)}$$

ec. 5.5.4

ya que $P(D)=cte.$ por ser independiente de h y no contener información discriminante, se puede considerar lo siguiente:

$$h_{MAP} = \arg \max_{h \in H} P(D | h)P(h)$$

ec. 5.5.5

De manera que para obtener la correspondiente h_{MAP} es necesario el cálculo para cada clase involucrada de su respectivo $P(D|h)$ conocido a priori y $P(h)$.

De forma práctica, se considera de manera inicial que la probabilidad *a priori* de cada clase es uniforme, una vez terminado este proceso, se calcula nuevamente la probabilidad *a priori* de cada clase.

La fig. 5.13 muestra un ejemplo de segmentación MAP para dos clases.

5.6 CASO APARTE: ALGORITMO DE K-MEDIAS

El nombre de este algoritmo está ya consagrado en la literatura especializada y hace referencia a que existen K clases, siendo necesario, por tanto, conocer *a priori* el número de clases existentes. Es un algoritmo sencillo, pero muy eficiente, siempre que el número de clases se conozca *a priori* con exactitud.

Partiendo de un conjunto de objetos a clasificar X_1, X_2, \dots, X_p , el algoritmo de las K-medias realiza las siguientes operaciones.

Paso 1: Estableciendo previamente el número exacto de clases existentes, digamos K , se escogen al azar entre los elementos h a agrupar K vectores, de forma que van a constituir los centroides (al ser los únicos elementos) de las K clases. Es decir $\alpha_1 : Z_1(1); \alpha_2 : Z_2(1) \dots \alpha_k : Z_k(1)$, en donde se ha introducido entre paréntesis el índice iterativo de este algoritmo.

Paso 2: Como se trata de un proceso recursivo con un contador n , en la iteración genérica n se distribuye todas las muestras $\{X\}_{1 \leq j \leq p}$ entre las K clases, de acuerdo a la siguiente regla:

$$X \in \alpha_j(n) \text{ si } \|X - Z_j(n)\| < \|X - Z_i(n)\|$$

$$i = 1, 2, \dots, K$$

$$i \neq j$$

ec. 5.6.1

en donde se han indexado las clases (que son dinámicas) y sus correspondientes centroides.

Paso 3: Una vez redistribuidos los elementos a agrupar entre las diferentes clases, es preciso recalcular o actualizar los centroides de las clases. El objetivo en el cálculo de los nuevos centroides de las clases es minimizar el índice de rendimiento siguiente:

$$J_i = \sum_{X \in \alpha_i(n)} \|X - Z_i(n)\|^2$$

$$i = 1, 2, \dots, K$$

ec. 5.6.2

Este índice se minimiza utilizando la media muestral o aritmética de $\alpha_i(n)$:

$$Z_i(n+1) = \frac{1}{N_i(n)} \sum_{X \in \alpha_i(n)} X$$

$$i = 1, 2, \dots, K$$

ec. 5.6.3

Siendo $N_i(n)$ el número de elementos de la base α_i en la iteración n .

Paso 4: Se comprueba si el algoritmo ha alcanzado una posición estable. Es decir, si se cumple:

$$Z_i(n+1) = Z_i(n)$$

$$i = 1, 2, \dots, K$$

ec. 5.6.4

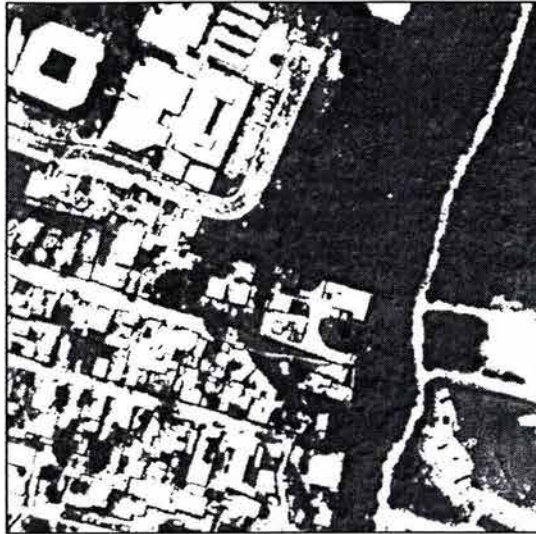
Si se cumple, el algoritmo finaliza. En caso contrario, salta al paso 2.

El algoritmo de las K-medias, tal como se dijo anteriormente, es simple y extraordinariamente eficiente si el número de clases se conoce *a priori* con exactitud, es decir, es muy sensible al parámetro K. Un valor de K superior al número real de clases dará lugar a clases ficticias, mientras que un K inferior producirá menos clases de las reales.

Una forma de detectar una posible mala elección del parámetro K es analizar las dispersiones estadísticas de las clases formadas. Cuando estas dispersiones sean sensiblemente diferentes entre sí, siendo algunas de ellas muy elevadas con respecto a las demás, se puede sospechar que se ha manejado un valor de K bajo. Análogamente, si algunas distancias interclases (separación entre los centroides de dos clases) son muy pequeñas respecto a las demás distancias interclases, entonces es válido plantearse que el parámetro K introducido es alto. Desgraciadamente, las dos situaciones anteriores pueden deberse exclusivamente a la naturaleza de las variables características escogidas.

En la fig. 5.14 se aplicó el algoritmo de K-medias de la fig. 2.4, para dos clases (urbano, no-urbano), el valor del píxel original perteneciente a cada clase fue sustituido, como se puede

observar las sombras formadas entre las casas se confunde con la clase no-urbana, también se puede observar las diferencias con respecto a la fig. 5.13.



segmentación MAP de la fig. 2.4 para dos clases: urbano, no-urbano
fig. 5.13



segmentación K-medias
fig 5.14

6

CAMPOS ALEATORIOS DE MARKOV (CAM)

6.1 GENERALIDADES

La segmentación de imágenes es un proceso que divide una imagen en un conjunto de regiones, basándose en características similares tales como intensidad o textura. La segmentación es uno de los elementos más importantes de cualquier sistema de visión; es el primer nivel para el entendimiento de la imagen y afecta severamente el procesamiento posterior de ésta, proporcionando estructuras útiles tales como regiones y bordes.

Un algoritmo de segmentación bajo el esquema de umbrales suele fallar a menudo, al juntar regiones que deben estar separadas y viceversa, porque la información sobre la uniformidad de una región y la discontinuidad entre regiones diferentes no está bien incorporada a los algoritmos. Por ello interesa incorporar características tolerantes a cierto nivel de distorsión que proporcione robustez a los resultados.

El modelo de Campo Aleatorio de Markov (CAM) tiene una potente capacidad, siendo utilizado en problemas de visión tales como restauración y segmentación de imágenes. El uso práctico del Campo Aleatorio de Markov se atribuye a la equivalencia entre las distribuciones de probabilidad de Campo Aleatorio de Markov y de Campo de Aleatorio de Gibbs establecidas por el teorema de Hammersley-Clifford. Esto permite modelar matemáticamente problemas de visión manejables para el análisis de imagen en el marco bayesiano.

La segmentación de una imagen se considera como un problema de etiquetado, de tal manera que la solución al problema es el conjunto de etiquetas asignado a los *sites* (pixels) de la imagen, así se considera un conjunto de *sites* a los cuales se le pueden asignar un conjunto de etiquetas. Sea S un conjunto discreto con m *sites*:

$$S = \{1, 2, \dots, m\}$$

Para una imagen de $n \times n$ pixels se puede denotar al conjunto de *sites* de la siguiente manera:

$$S = \{(i, j) \mid 1 \leq i, j \leq n\}$$

La relación que guardan entre sí diferentes *sites* es tratada en las vecindades.

L es el conjunto de etiquetas y puede ser un conjunto discreto o continuo. En el caso continuo este conjunto de etiquetas puede corresponder a \mathbb{R} (conjunto de números reales) o a un intervalo determinado de éste:

$$L_c = [x_1, x_h] \subset \mathbb{R}$$

Para el caso discreto, las etiquetas solo podrán tomar un valor determinado dentro de un conjunto de M valores:

$$L_d = \{l_1, \dots, l_M\} = \{1, \dots, M\}$$

En un conjunto de etiquetas ordenado, podemos definir una medida numérica (cuantitativa) de semejanza entre dos 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. En el problema de etiquetado se tiene que asignar a cada *site* de S una etiqueta de L , ya sea continuo o discreto. Si se quiere detectar un borde en una imagen, S contendrá a todos los pixels pertenecientes a la imagen y el conjunto de etiquetas $L = \{\text{borde}, \text{no borde}\}$ para cada estado $i \in S$, les serán asignados.

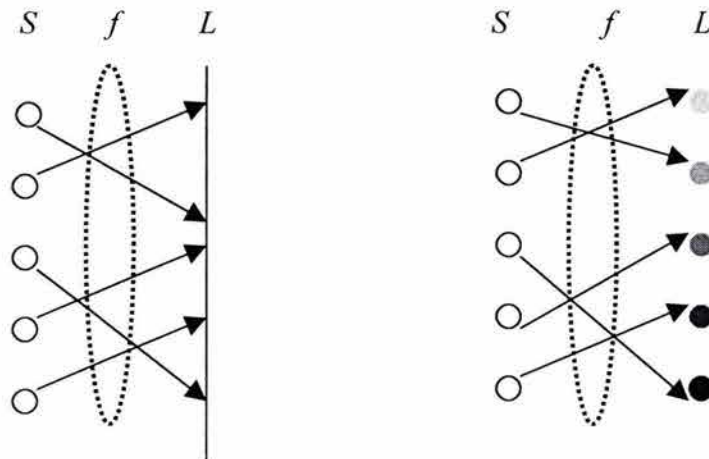
El etiquetado se representaría:

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

donde a cada *site* de S se le ha asignado una de las etiquetas de L con la función mapping o de mapeo:

$$f : S \rightarrow L$$

donde a cada *site* de S le asigna una única etiqueta de f_i de L .



etiquetado continuo y etiquetado discreto

fig. 6.1

En la terminología de campos aleatorios un determinado conjunto de etiquetas recibe el nombre de configuración, se considera un caso específico de configuración a una imagen, ya que ésta responde a un conjunto determinado de etiquetas.

Cuando todos los *sites* tiene el mismo conjunto de etiquetas de L , el conjunto de todas las configuraciones posibles, el espacio de configuración es el producto cartesiano:

$$\text{Espacio de Configuración} \quad \mathfrak{S} = L \times L \times L \times \dots \times L = L^m$$

m es el número total de *sites*.

Por ejemplo, si se retoma la fig. 6.2, se presentan tres posibles configuraciones (de 2^{16} posibles) en una *lattice* de 16 *sites* con $L=\{0, 1\}$.

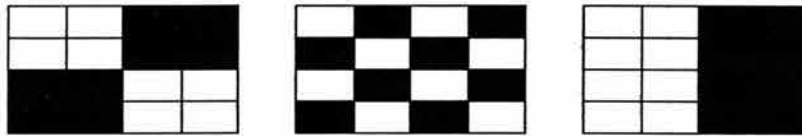


fig. 6.2

En términos de homogeneidad y de continuidad, se clasifica un problema de etiquetado en una de las cuatro categorías siguientes:

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

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

La restauración de imágenes en niveles de gris, o el suavizado de imágenes, son ejemplos de PE1. El conjunto de *sites* S corresponde a los pixels de la imagen y el conjunto de etiquetas L a un intervalo real. La restauración consiste en estimar la señal de la imagen real a partir de una imagen degradada o con ruido.

La restauración de imágenes binarias, o imágenes multinivel, son ejemplos de PE2. Del mismo modo que la restauración continua, se intenta estimar la señal de la imagen real. La diferencia es que cada pixel de la imagen resultante aquí asume un valor discreto y, por tanto, L es conjunto discreto de etiquetas.

Para una *lattice* (rejilla) regular S , el conjunto de vecinos del *site* a está definido como el conjunto de *sites* próximos dentro de un radio r

$$N_a = \{(a, s) \in S \mid \text{dist}(a, s)^2 \leq r, a \neq s\}$$

ec. 6.1

donde $\text{dist}(a, s)$ es la distancia de Euclides entre el *site* a y el *site* s , r es un número entero que puede tomar los valores de $\{1, 2, 4, 5, 8, 9, \dots\}$.

Por ejemplo, para la vecindad de orden 1 con respecto al *site* a :

$$N_a^1 = \{b, c, d, e\}$$

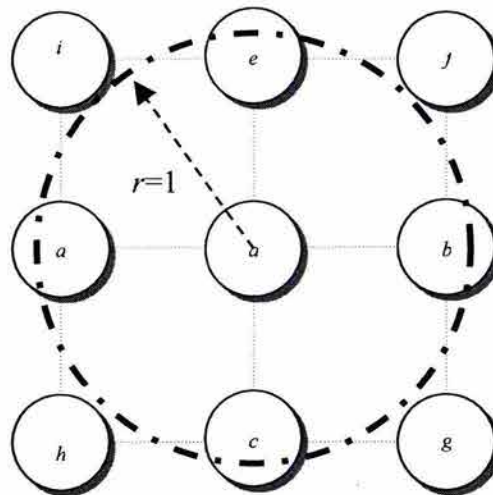


fig. 6.3

para la vecindad de orden 2 del site a :

$$N_a^2 = N_a^1 \cup \{f, g, h, i\}$$

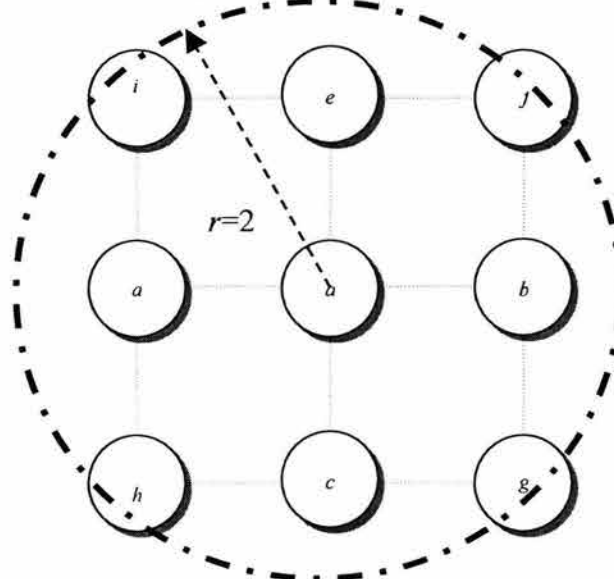


fig. 6.4

vecindad de orden 3:

$$N_a^3 = N_a^2 \cup \{j, k, l, m\}$$

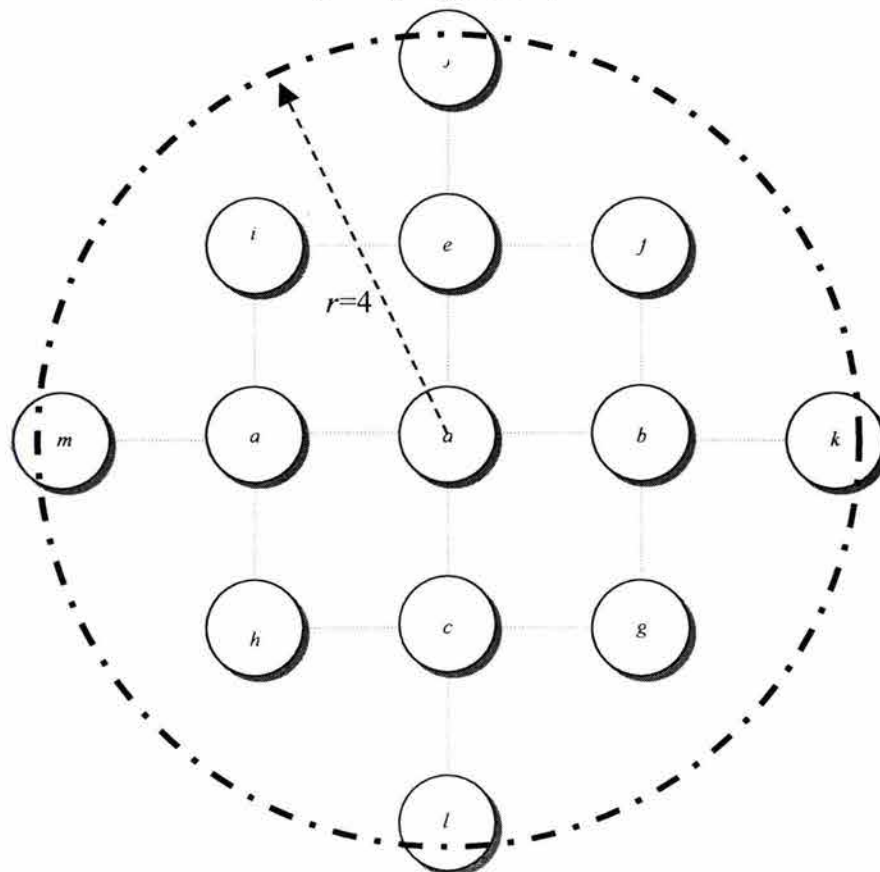


fig. 6.5

Para el sistema de vecinos de primer orden, llamado sistema de 4-vecinos, cada *site* (interior) tiene cuatro *sites* vecinos, según lo mostrado en la fig. 6.3. En el sistema de vecinos de segundo orden, también llamado sistema de 8-vecinos, hay ocho vecinos para cada *site* (interior), según lo mostrado en la fig. 6.4.

Con base en estas definiciones se introduce el concepto de *conexidad*, que constituye una relación de vecindad espacial entre un píxel y sus próximos vecinos.

Dos *sites* s y s' tienen 4-*conexidad* si s' pertenece al conjunto de N_s^1 .

Dos *sites* s y s' tienen 8-*conexidad* si s' pertenece al conjunto de N_s^2 .

Los números $n=1, \dots, 5$ mostrados en la fig. 6.4 indica el sistema de vecinos de orden n -ésimo para el *site* x .

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

fig. 6.6

Cuando se ha especificado un orden entre los elementos de S , el conjunto de vecinos se puede determinar más explícitamente. Por ejemplo, sea $S = \{1, 2, \dots, m\}$ un conjunto ordenado de *sites* cuyos elementos son los índices de los píxels de una imagen 1D, entonces un estado interior $i \in \{2, \dots, m-1\}$ tiene dos vecinos $N_i = \{i-1, i+1\}$ y los *sites* de la frontera (los dos extremos) tienen un vecino cada uno, $N_1 = \{2\}$ y $N_m = \{m-1\}$. El conjunto de *sites* de una rejilla regular rectangular $S = \{(i, j) | 1 \leq i, j \leq n\}$ corresponde a los píxels de una imagen en dos dimensiones de tamaño $n \times n$. Cada *site* interior, representado por (i, j) , tiene 4 vecinos $N_{i,j} = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}$, los estados de la frontera tienen tres y los estados de las esquinas tienen dos.

6.2.2 Cliques

El par (S, N) constituye un grafo, donde N establece la relación de los nodos en S . Un *clique* c de (S, N) está definido como un subconjunto de *sites* en S . Un *clique* puede formarse de un solo estado, dos estados, tres estados o más estados y está representado por $C_1, C_2, y C_3, \dots$, respectivamente, donde:

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

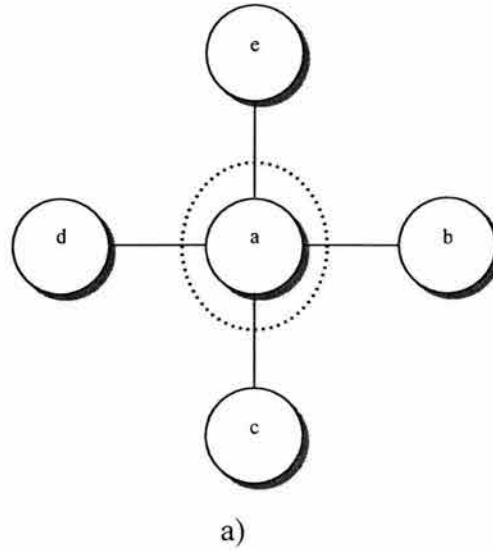
$$C_2 = \{(i, i') | i' \in N_i, i \in S\} \text{ siendo vecinos entre sí.}$$

$$C_3 = \{(i, i', i'') | i, i', i'' \in S\}$$

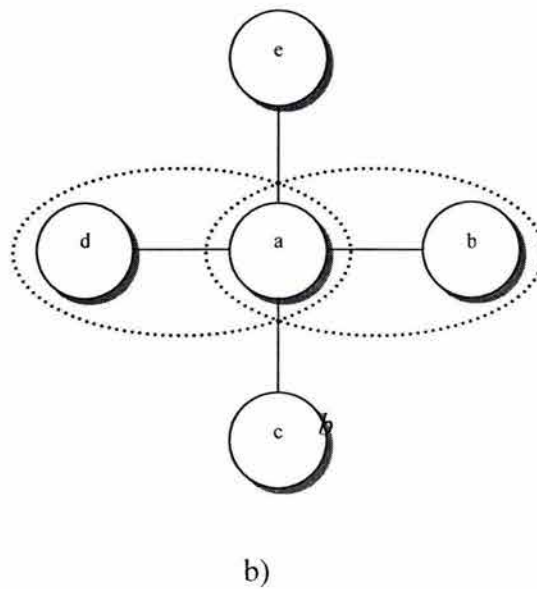
El conjunto de todas las *cliques* para $\{S, N\}$ será $C = C_1 \cup C_2 \cup \dots$

En una vecindad de primer orden, los *cliques* serían:

para C_1



para C_2



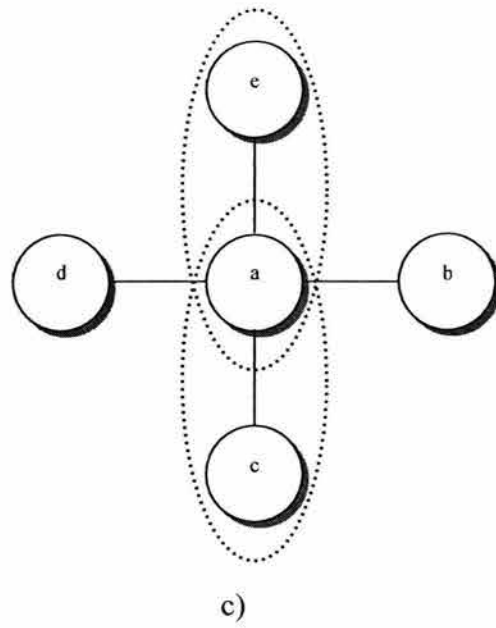
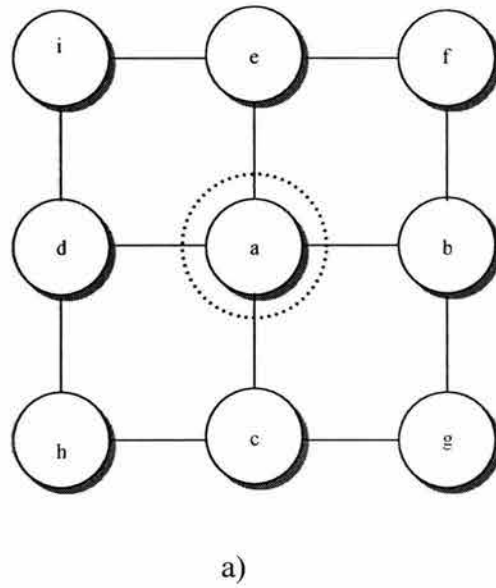


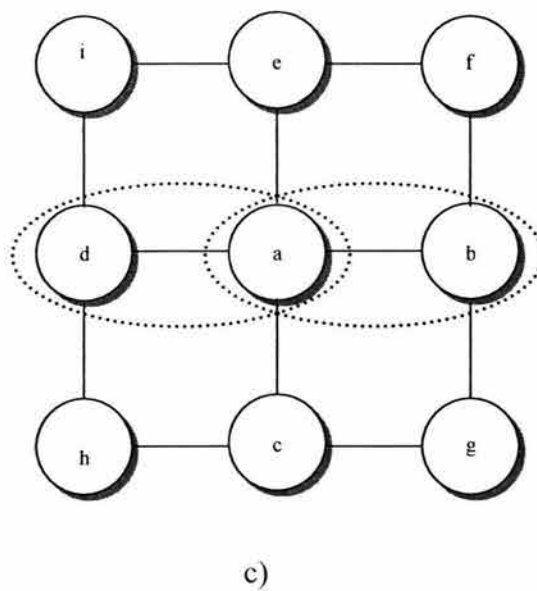
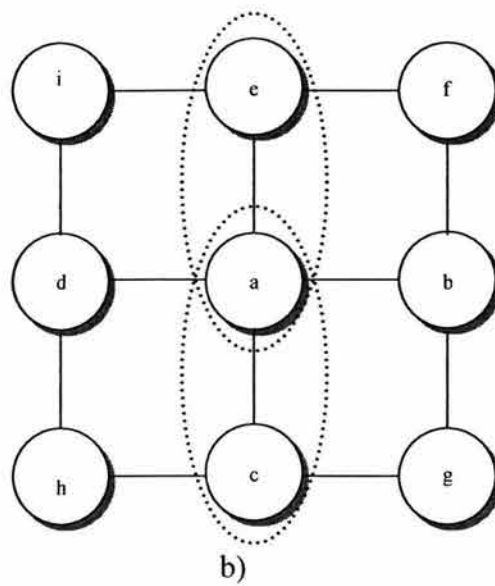
fig. 6.7

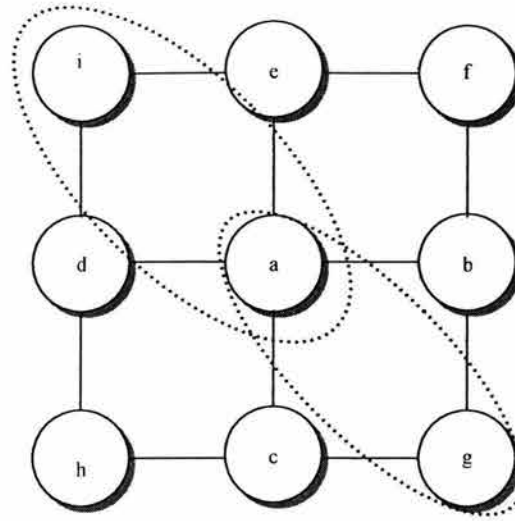
En una vecindad de segundo orden, los *cliques* serían:

para C_1

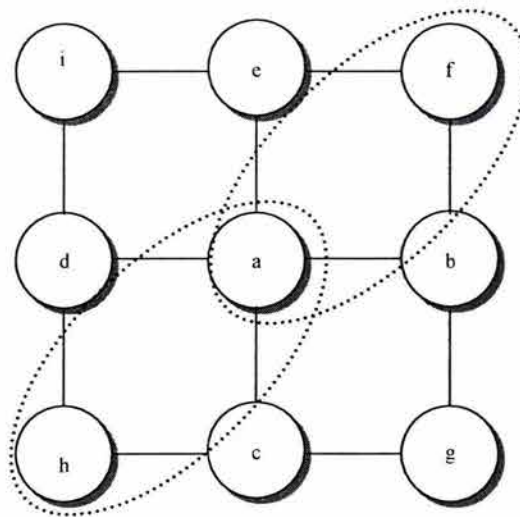


a)

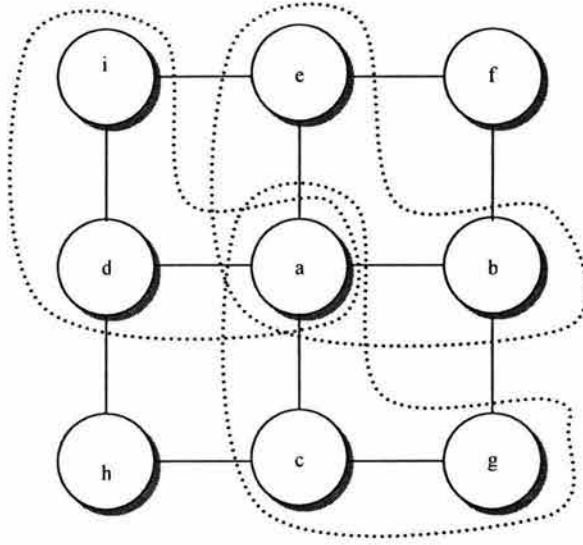
para C_2 



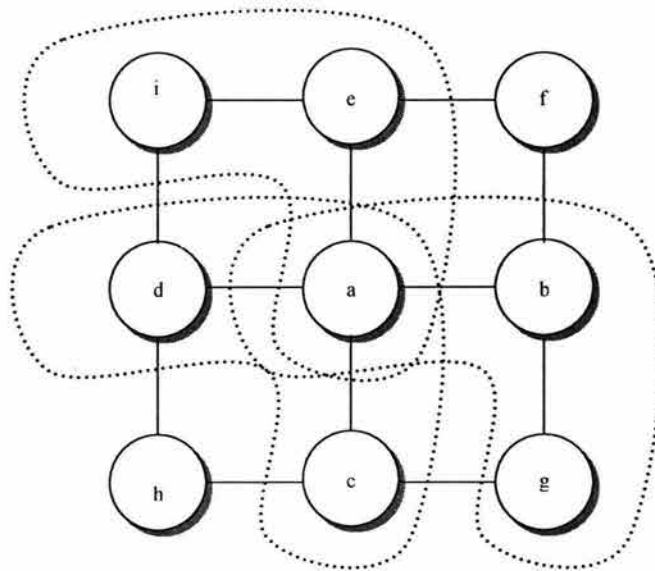
d)



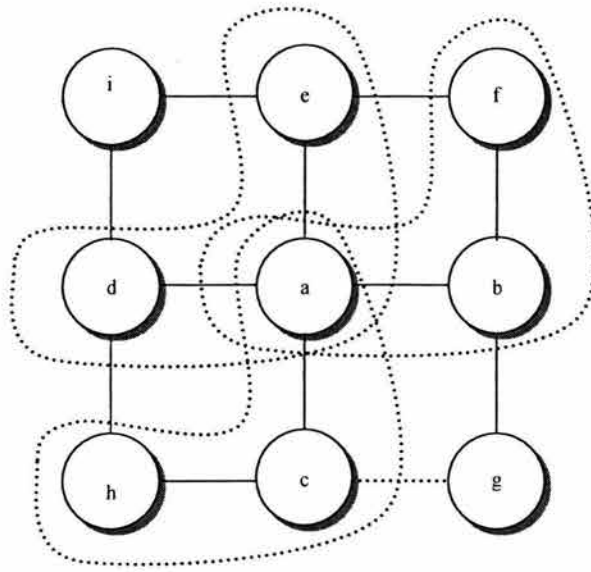
e)

para C_3 

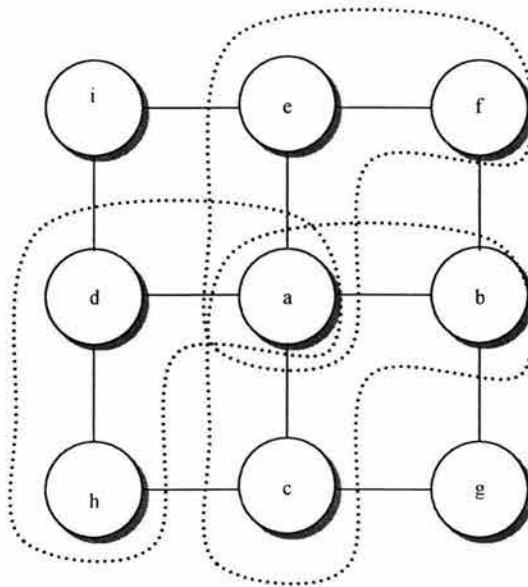
f)



g)



h)



i)

para C_4

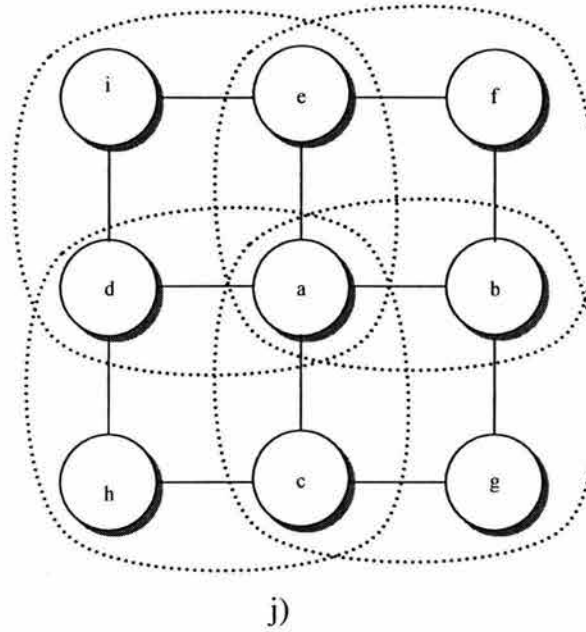


fig. 6.8

6.3 PROPIEDADES DE MARKOV

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

Sea $F = \{F_1, \dots, F_m\}$ una familia de variables aleatorias definidas en S , donde cada variable aleatoria F_i toma un valor de f_i en L . Llamaremos Campo Aleatorio a la familia de variables aleatorias F , f es una configuración de \mathfrak{S} .

Una familia de variables aleatorias F se dice que es un Campo Aleatorio de Markov (CAM) sobre S con respecto al sistema de vecindad N si y solo si las siguientes condiciones se satisfacen:

Positividad

$$P(f) > 0$$

ec. 6.2

$$\forall f \in \mathfrak{S}$$

Markovianidad

$$P(f_i | f_{S-\{i\}}) = P(f_i | f_{N_i})$$

ec. 6.3

donde $f_{S-\{i\}}$ denota al conjunto de etiquetas en todos *sites* exceptuando a $\{i\}$ y $f_{N_i} = \{f_i \mid i' \in N_i\}$ es el conjunto de etiquetas que pertenecen a los *sites* de la vecindad de i .

Un Campo Aleatorio de Markov es caracterizado por su propiedad local (Markovianidad) mientras que un Campo Aleatorio de Gibbs esta caracterizado por su propiedad global (distribución de probabilidad Gibbs). *El teorema de Hammersley- Clifford establece la equivalencia entre estas dos propiedades: 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.* La demostración se puede encontrar en [LI01].

Una distribución de Gibbs tiene la siguiente forma:

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

ec. 6.4

donde

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

ec. 6.5

T: es la temperatura , $T=1$
 $U(f)$: función de energía
 Z: función de partición

La función de energía $U(f)$ viene definida por

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

ec. 6.5

y $V_c(f)$ es el potencial de *clique*.

$U(f)$ es la suma de todas las funciones *cliques* potenciales, $V_c(f)$, para todos las *cliques* posibles C , donde el valor de $V_c(f)$ depende de la configuración local de la *clique* c .

Se dice que un Campo Aleatorio de Gibbs es homogéneo si $V_c(f)$ es independiente de la posición relativa de la *clique* c en S . Se dice que es isotrópico si la posición de V_c es independiente de la orientación de c . Es más simple especificar una distribución de Campo Aleatorio de Gibbs si es homogéneo e isotrópico que si no lo es. La homogeneidad se asume en la mayoría de los modelos de Campos Aleatorios de Markov por conveniencia matemática y computacional.

$P(f)$ mide la probabilidad de la ocurrencia de una configuración determinada. Las configuraciones más probables son aquellas con energías más bajas. La temperatura T controla la forma de la distribución. Cuando la temperatura es alta, todas las configuraciones tienden a estar igualmente

distribuidas. Con la temperatura cercana a cero, la distribución se concentra alrededor del mínimo global de energía.

La energía de una distribución de Gibbs se puede expresar como la suma de varios términos correspondientes a *cliques* de diferentes tamaños, es decir:

$$U(f) = \sum_{c \in C} V_c(f) = \sum_{\{i\} \in C_1} V_1(f_i) + \sum_{\{i,j\} \in C_2} V_2(f_i, f_j) + \sum_{\{i,j,k\} \in C_3} V_3(f_i, f_j, f_k) + \dots$$

ec. 6.6

La función anterior implica una distribución homogénea de Gibbs porque V_1, V_2, \dots , son independientes de las localizaciones i, j . En las distribuciones no homogéneas de Gibbs, las funciones *clique* se deben considerar de la forma: $V_1(i, f_i), V_2(i, j, f_i, f_j)$ y así sucesivamente.

Un caso importante es cuando se consideran las *cliques* de tamaño dos como máximo. Por lo que la función anterior quedaría

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

ec.6.7

La probabilidad condicional puede escribirse como:

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

ec. 6.8

6.4 MODELOS DE MARKOV

A continuación se introducen brevemente los modelos más comunes en el estudio de los Campos Aleatorios de Markov.

6.4.1 Modelo de Ising

El modelo de *Ising* surge del problema de modelar materiales ferromagnéticos en lo que se conoce como fenómeno de *Ising*, en el cual se observa una serie de dipolos en una línea que pueden estar orientados hacia arriba (+1) o hacia abajo (-1) y el estado de cada dipolo se ve influenciado por los dipolos cercanos.

Las restricciones de contexto entre dos etiquetas son un primer nivel a tener en cuenta en la información del contexto. Son muy utilizadas debido a su simplicidad y bajo coste computacional. Se codifican en la función de energía de Gibbs como funciones *clique* potenciales para cada par de *sites*. Entonces la función *clique* potencial queda definida por:

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

ec. 6.8

que llamamos energía de segundo orden porque se calcula para cada *clique* formada por dos *sites*.

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

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

ec.6.9

EL modelo anterior recibe el nombre de modelo auto-logístico, si f_i y f_j toman valores en el conjunto de etiquetas discreto $L=\{0,1\}$ o $L=\{+1,-1\}$. La correspondiente función de energía es de la forma

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

ec. 6.10

donde $\beta_{i,j}$ puede ser interpretado como el coeficiente de interacción. Cuando el sistema de vecindad que se considera es el de 4-vecinos en una *lattice* 2D, entonces el modelo auto-logístico recibe el nombre de modelo de *Ising*. La probabilidad condicional para el modelo de Ising es:

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

ec. 6.11

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

6.4.2 Modelo de Potts

El modelo de Potts también es conocido como Multi-Level Logistic model, MLL o modelo generalizado de Ising. Si hay $M > 2$ etiquetas discretas en el conjunto de etiquetas $L = \{1, \dots, M\}$, el potencial del *clique* depende del tipo c (tamaño y orientación posible).

Para los *cliques* que contienen más de un *site* se define las funciones de potencial para cada *clique*, la cual es de la forma:

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

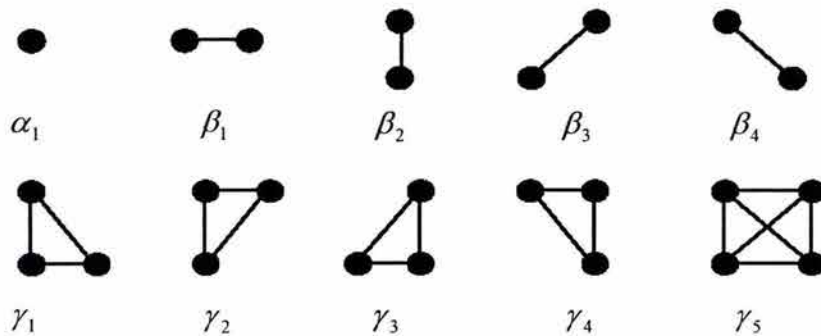
ec. 6.12

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

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

ec. 6.13

donde α_I es el potencial para la etiqueta I . La fig.6.8 muestra los tipos de *cliques* y los parámetros asociados en sistemas de vecinos de segundo orden.



Tipos de *cliques* y parámetros potenciales asociados a una vecindad de segundo orden
fig. 6.9

Se asume que un modelo MLL es de segundo orden, sí y solo sí α (para *cliques* de un estado) y β (para *cliques* de dos estados) son parámetros distintos de cero. La función *clique* potencial para *clique* de dos estados es

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

ec. 6.14

donde β_c es el parámetro β para las *cliques* de tipo c , y C_2 es el conjunto de *cliques* de dos estados. Para una vecindad de primer orden hay cuatro tipos de *cliques* de dos estados fig. 6.9 y cuatro diferentes β_c . Cuando el modelo es isotrópico los cuatro toman el mismo valor. Debido a su simplicidad, el modelo de Potts es usado extensamente para modelar regiones y texturas.

Cuando el modelo es isotrópico, la probabilidad condicional puede ser expresada como sigue:

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

ec. 6.15

donde $n_i(I)$ es el número de estados en N_i que tienen etiqueta I .

El caso particular para el cual las etiquetas son $\{0,1\}$ se reduce al modelo de *Ising*.

6.5 OPTIMACIÓN

En general, para resolver el problema de segmentación bajo el esquema MAP-CAM, después de haber definido las funciones de potencial para los diferentes *cliques*, es necesario minimizar la función de energía *a posteriori*. La solución se encuentra en el espacio de configuraciones \mathcal{S} , siendo un problema combinatorio, se utiliza un algoritmo de optimación como los es el Recocido Simulado que encuentra una solución global. Para minimización local, en [LI02] se describen métodos como ICM y Relaxation Labeling.

6.5.1 Recocido Simulado (*Simulated Annealing*)

El algoritmo de recocido simulado está basado en una analogía entre el recocido de sólidos y la problemática de resolver problemas de optimación combinatoria de gran escala. Por esta razón el algoritmo se conoce como recocido simulado. Recocido denota un proceso de calentamiento de un sólido a una temperatura en la que sus granos deformados cristalizan para producir nuevos granos. La temperatura de recocido o de recristalización, depende del tipo de material, del grado de deformación del mismo. Seguida a la fase de calentamiento, viene un proceso de enfriamiento en donde la temperatura se baja poco a poco. De esta manera, cada vez que se baja la temperatura, las partículas se acomodan en estados de más baja energía hasta que se obtiene un sólido con sus partículas acomodadas conforme a una estructura de cristal.

En general, encontrar el mínimo global de una función objetivo con muchos grados de libertad sujeta a restricciones es un problema completo, ya que la función objetivo tiende a tener varios mínimos locales.

Un procedimiento para resolver problemas de optimación difíciles debe muestrear valores de la función objetivo de forma tal que se tenga una alta probabilidad de encontrar soluciones cerca del óptimo y debe también poder implementarse de manera eficiente. El recocido simulado es una técnica que cumple con estos requerimientos.

```

Inicializar T y f;
Repetir
    muestrear aleatoriamente f de N(f) bajo T;
    decrementar T;
    mientras (T → 0);
regresar f;

```

algoritmo de recocido simulado
fig. 6.10

Inicialmente T es fijada en un valor alto y f es una configuración que se inicializa de manera aleatoria.

En la práctica, el algoritmo de Recocido Simulado, emplea un algoritmo de Metrópolis, con un esquema de temperatura variable, a fin de minimizar la función de energía *a posteriori* $U(X|Y)$. El objetivo del problema de optimación consiste en encontrar un elemento de Ω_0 donde:

$$\Omega_0 = (\varpi \in \Omega, \text{ donde } \varpi \text{ es de energía mínima})$$

U^* es la energía máxima en todas las configuraciones y U_* es la energía mínima, la desviación entre las dos energías es $\Delta U = U^* - U_*$.

Como se verá en 6.5.2, al emplear un muestreador de Gibbs con una temperatura T_n , que varía después de cada barrido, la variación de la temperatura debe seguir el esquema:

$$T_n \rightarrow 0 \text{ cuando } n \rightarrow \infty$$

$$T_n = \frac{N\Delta U}{\log(n)} \text{ para } n \geq 2$$

donde N es el número de *sites* en la imagen.

Para toda configuración inicial η (de la imagen) y para toda la configuración de una interacción cualquiera ϖ , se tiene

$$\lim_{n \rightarrow \infty} P(Y^{(n)} = \varpi | Y^{(0)} = \eta) = \pi(\varpi)$$

donde $\pi(\varpi)$ corresponde a la medida de probabilidad de Gibbs definida por los Campos Aleatorios de Markov. Así, existe una convergencia hacia una realización dada, con una probabilidad que sigue la distribución de Gibbs global.

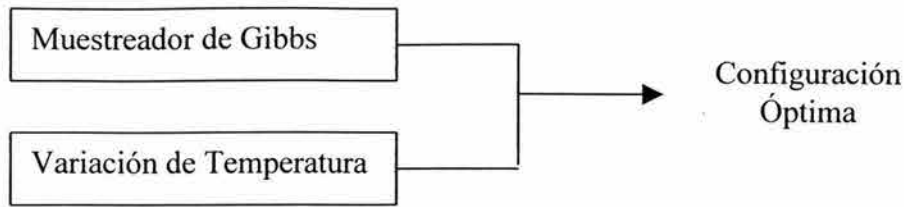


fig. 6.11

El Algoritmo de Recocido Simulado aplicado al CAM quedaría como sigue:

1. Inicialización

- Determinación aleatoria de la solución $Y^{(0)}$
- Escoger una temperatura inicial $T^{(0)}$ suficientemente elevada

2. Barrido secuencial de 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 sorteo:

variable aleatoria con distribución uniforme $\xi \in \Omega$

3. Criterio de Metrópolis

- Calcular variación de la energía local asociada al site r:

$$\Delta U = U(\xi | V_r^{(n)}, X) - U(Y_r^{(n)} | V_r^{(n)}, X)$$

- Pueden surgir dos casos

$$\text{Si } \Delta U \leq 0 \Rightarrow Y_r^{(n+1)} = \xi$$

$$\text{Si } \Delta U > 0 \Rightarrow Y_r^{(n+1)} = \xi \text{ con probabilidad } \xi = e^{-\frac{\Delta U}{T}}$$

4. Continuar con la etapa 2

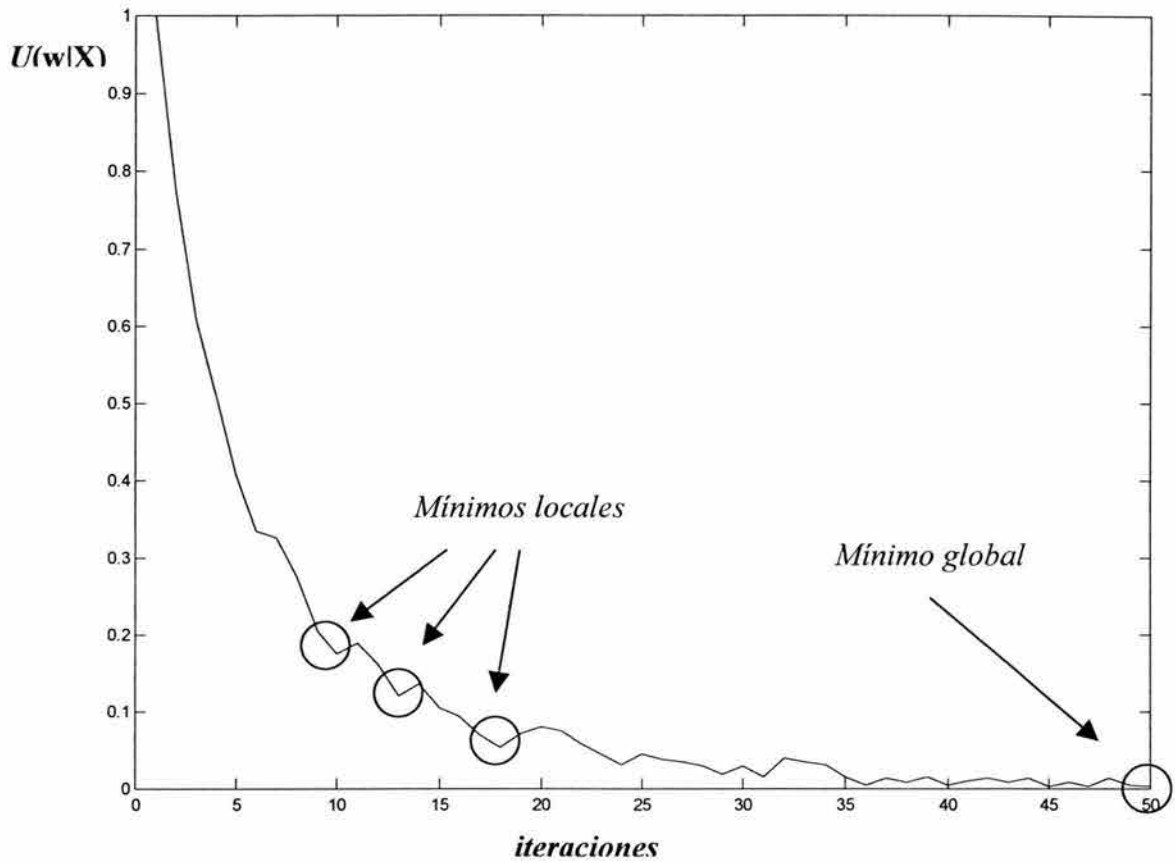
5. Serie de temperaturas decrecientes hacia 0, por lo que $\lim_{n \rightarrow \infty} T_n = 0$

La expresión $\Delta U = U(\xi | V_r^{(n)}, X) - U(Y_r^{(n)} | V_r^{(n)}, X)$ representa tanto la variación local de energía, como la variación global de la energía asociada a la transición $Y_r^{(n+1)} = \xi$. De forma

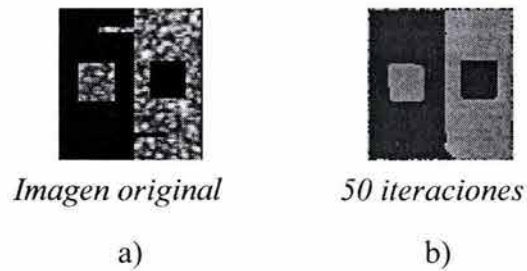
práctica, se simula la probabilidad dada por $p = e^{-\frac{\Delta U^{(n)}}{T^{(n)}}}$, efectuando el sorteo aleatorio de una variable distribuida uniformemente, $\eta \in [0,1]$

$$Y_r^{(n+1)} = \begin{cases} \xi & \text{si } 0 \leq \eta \leq p \\ Y_r^{(n)} & \text{otro caso} \end{cases}$$

La fig. 6.12 es la función de energía *a posteriori* de la imagen segmentada de la fig. 6.13-a, las figuras c, d y e muestran mínimos locales de la función de energía *a posteriori*.



función de energía *a posteriori*
fig. 6.12





10 iteraciones

c)



13 iteraciones

d)



18 iteraciones

e)

fig. 6.13

6.5.2 Esquemas de decremento de temperatura

Para minimizar la función de energía *a posteriori* es necesario decrementar la temperatura para cada configuración, ya que para temperaturas altas, las configuraciones más aleatorias están presentes. Para asegurar la convergencia a un mínimo global, el esquema de decremento de temperatura, deberá seguir las siguientes condiciones:

$$\lim_{n \rightarrow \infty} T^{(n)} = 0$$

ec. 6.18

$$T^{(n)} \geq \frac{m\alpha\Delta}{\ln(1+n)}$$

ec. 6.19

donde $\Delta = \max E(f) - \min E(f)$ y m es el número de *sites*, entonces el sistema converge al mínimo global, la desventaja de la ec. 6.19 es la lentitud, de manera heurística se han adoptado los siguientes esquemas:

1. La primera, establecía una velocidad de enfriamiento de la temperatura de tipo geométrico, $T^{(n+1)} = \alpha \cdot T^{(n)}$, como $0 < \alpha < 1$. Las evidencias empíricas señalaban que valores elevados de α comprendidos entre 0.8 y 0.99 eran los que mejores resultados proporcionaban.
2. Según el segundo esquema de enfriamiento, se ejecuta una sola iteración para cada temperatura, pero por el contrario la temperatura se reduce a una velocidad muy lenta según la fórmula $T \rightarrow \frac{T}{1 + \beta \cdot T}$, siendo β muy pequeño.
3. Entre los esquemas de descenso de temperatura también se encuentran los de la forma $T^{(n)} = \frac{c}{\log(n+1)}$, donde n es el número de iteraciones.

7

FUSIÓN DE DATOS

7.1 GENERALIDADES

La fusión de datos se puede ver como la integración de información proveniente de múltiples fuentes para dar como resultado una unidad específica y comprensiva de datos respecto a una entidad[HALL97].

Por ejemplo, el cerebro humano es el mejor ejemplo de una fusión de datos, integrando información proveniente de diferentes fuentes como lo son: vista, tacto, olfato, gusto y oído, para obtener una información total sobre el tema que centra su atención.

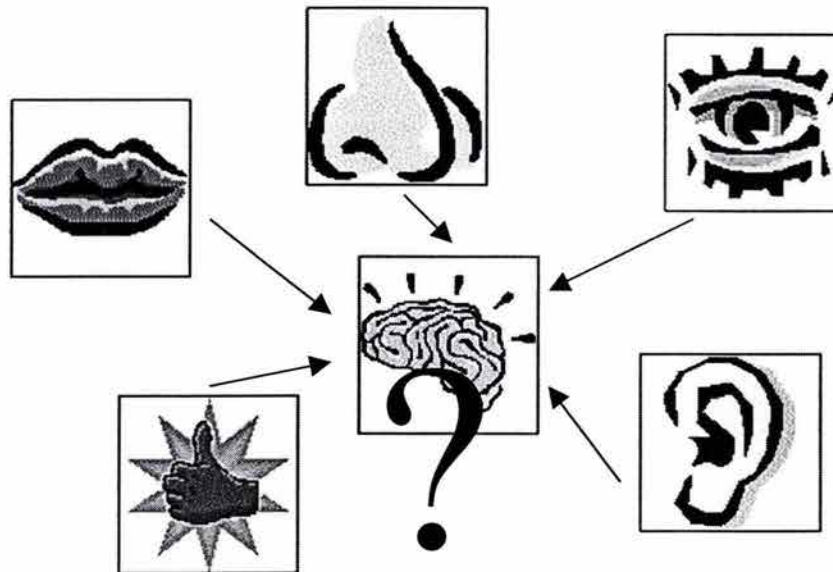


fig. 7.1

Los orígenes de la fusión de datos se encuentran en el control de procesos industriales y específicamente en el monitoreo de líneas de producción. Un buen ejemplo es el monitoreo de la producción de químicos, ya que existe una continua integración de varios parámetros que gobiernan el proceso: temperatura, presión, consumo de combustible, etc.

En el procesamiento digital de imágenes se ha encontrado uno de los campos mas demandantes para fusión de datos, por ejemplo: aplicaciones en imágenes satelitales, imágenes médicas (fig. 7.3 a-c), guía de vehículos y visión artificial (7.3 d-f).

El objetivo de la fusión de imágenes es usar varias de ellas, provenientes de diferentes fuentes, de tal manera que provea un entendimiento completo de la escena, no solo en término de posición y geometría, sino en término de interpretación semántica. Es deseable que la información contenida en un píxel, revele la presencia o ausencia del objeto de interés, en caso afirmativo, sin duda, este píxel será de mayor peso en el proceso de fusión. Dicho en otras palabras, la fusión de imágenes extraerá información útil de las fuentes de información tomando en cuenta información redundante e información complementaria que se presente.

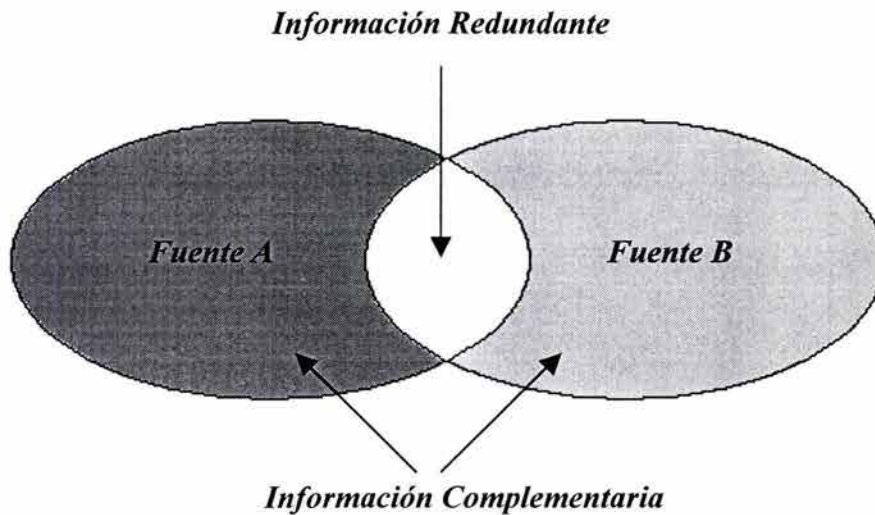
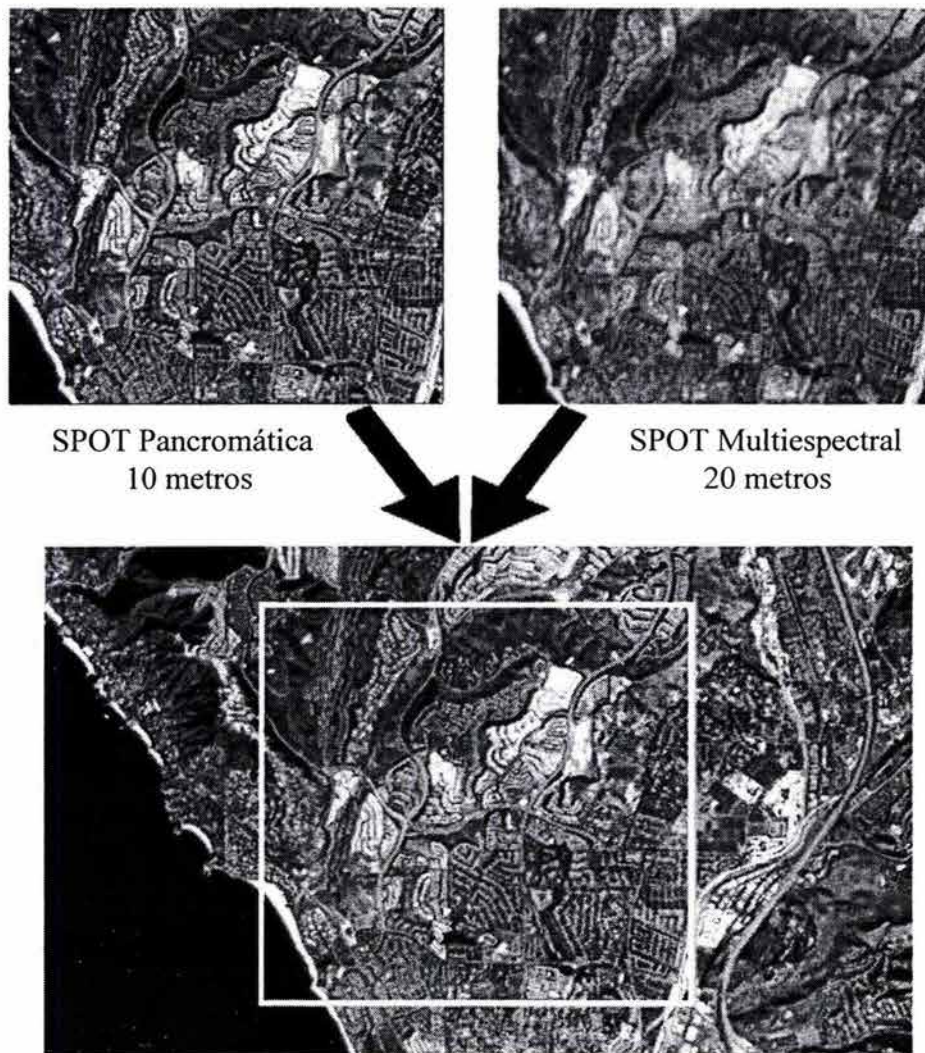
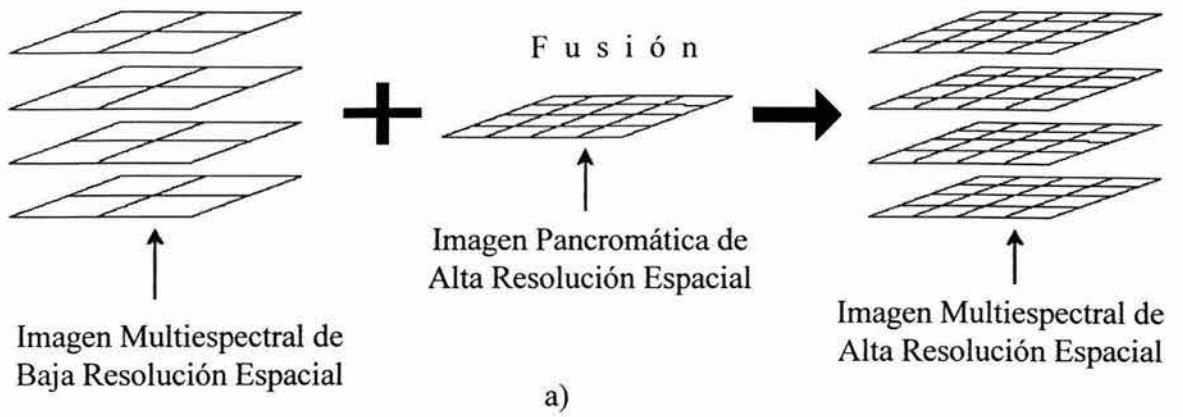
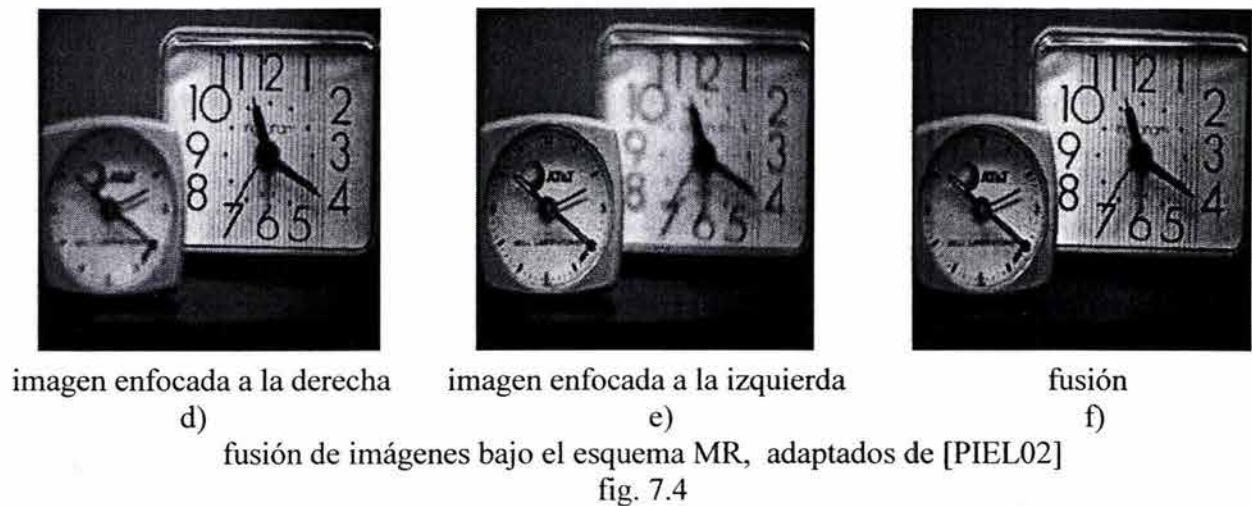
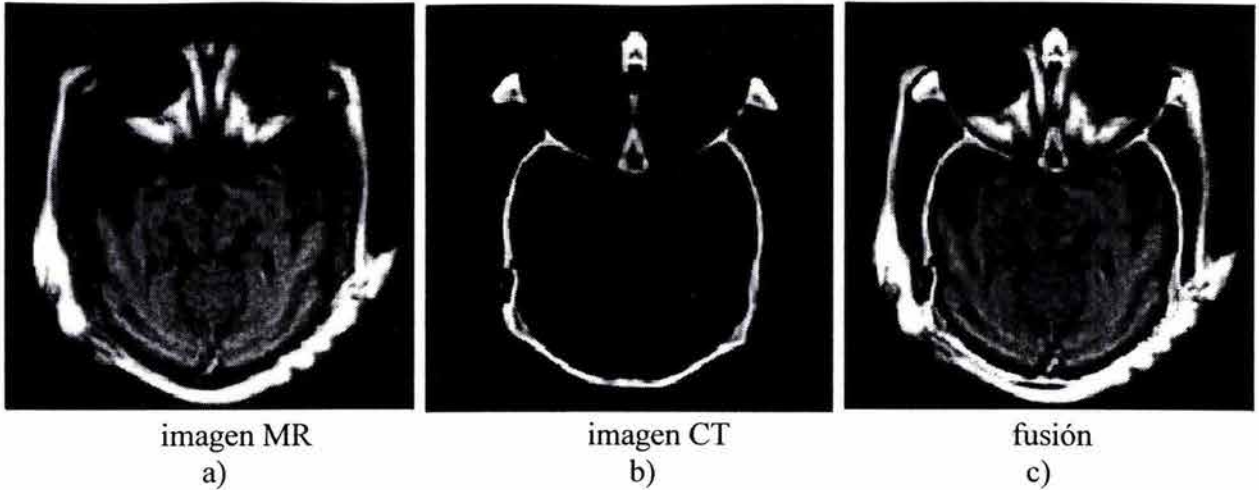


fig. 7.2

En los últimos años se ha realizado un avance en la fusión de datos, entendida como el proceso que permite obtener una unidad de información a partir de imágenes con diferente resolución espacial, espectral o radiométrica. En la fusión de datos se tiene como objetivo mejorar las características de las imágenes de percepción remota, a partir de información que se complementa entre sí. Un ejemplo (fig. 7.3) es la mejora de la resolución espacial de las imágenes pancromáticas SPOT de 10 [m] de resolución espacial de un canal con las multiespectrales SPOT de 3 canales y 20 [m] de resolución espacial para dar como resultado a imágenes multiespectrales de tres canales a una resolución de 10 [m].



fusión
fig. 7.3



Algunas aplicaciones de la fusión de imágenes son:

1. Robótica:

- Guía de vehículos
- Fusión de imágenes estéreo
- Reconocimiento de objetivos

2. Imágenes Médicas

- Fusión de Imágenes de rayos-X, tomografía computarizada (CT) y resonancia magnética (RM)
- Cirugía asistida por computadora

3. Manufactura

- Inspección automática de circuitos electrónicos
- Inspección no destructiva de material
- Sensores en línea de ensamblaje

4. Percepción Remota

- Imágenes adquiridas en diferentes bandas del espectro electromagnético
- Fusión de imágenes pancromáticas y multiespectrales

La fusión de datos es entendida como el conjunto de técnicas diversas cuyo denominador común es la mejora de ciertas características a partir de conjuntos de datos que, por sí solos, carecen de ellas, se divide en general:

Combinación de pesos: una aproximación simple que consiste en que cada píxel tomará el valor respectivo de acuerdo a la contribución de los pixels correspondientes a las demás fuentes de información.

Fusión en Color: la fusión en color por medio de transformaciones en color tienen la ventaja de presentar el resultado como una imagen en color, la técnica más sencilla consiste en considerar a cada imagen de diferente sensor, como un canal de color, resultando una combinación de banda de colores, así como de transformaciones, explotando los detalles que brindan las imágenes en pseudocolor y falso color.

Aproximación Biológica: en este esquema se toma de ejemplo el sistema de visión de las serpientes, donde sus ojos son sensibles a la radiación termal, la cual es interpretada por neuronas, así, la visión termal es sustituida por imágenes de diferentes sensores que serán posteriormente procesadas por redes neuronales.

Por otra parte en los últimos años se han realizado no menos de una docena de métodos para la mejora de la resolución espacial[TAPI01] y que están contenidos en alguna de las técnicas anteriores.

Entre ellos se encuentran:

Método IHS

Método PCS

Método SVR

Contribución Espectral Relativa

ARSIS(Wavelet)

Método de Fusión Mediante Entropía

Normalmente el ojo humano no distingue la proporción de azul, verde y rojo presentes en una imagen, en cambio, sí evalúa la intensidad (I), el color o matiz (H) y la saturación (S). El método IHS permite pasar del espacio de color al espacio de intensidad, color y saturación.

La intensidad o brillo es la medida de energía total en todas las longitudes de onda, siendo por lo tanto la responsable por la sensación de brillo de esa energía incidente sobre el ojo.

El matiz o color de un objeto es la medida de la longitud de onda medio de la luz que se refleja o se emite, definiendo, el color del objeto.

La saturación o pureza expresa el intervalo de longitud de onda alrededor de la longitud de onda media, en el cual la energía es reflejada o transmitida. Un alto valor de saturación resulta en un color espectralmente puro, en tanto que un bajo valor indica una mezcla de longitudes de onda que irá a producir tonos pasteles (apagados).

La transformación de las componentes rojo, verde y azul (RGB) en las componentes intensidad, matiz y saturación (IHS) puede ser utilizada para producir composiciones en color con una reducida correlación interbanda y consecuentemente con mejor utilización del espacio de colores, adicionalmente, permite combinar diferentes tipos de imágenes o imágenes de diferentes sensores. Esta transformación es realizada a través de algoritmos matemáticos que relacionan el espacio RGB al IHS.

Para producir composiciones en color, se seleccionan tres bandas de una imagen y se asocia cada banda a una de las componentes RGB. La transformación IHS se realiza pixel a pixel. Cada pixel en la imagen de salida corresponderá a un punto en el espacio IHS. El resultado es un conjunto de tres imágenes nuevas: una de intensidad, una de matiz y la otra de saturación. Estas imágenes son realizadas, de modo de expandir el intervalo de intensidad y saturación a través de las técnicas convencionales de contraste, y nuevamente transformarlas para el sistema RGB, permitiendo así una mejor separación de los rasgos que se desean observar.

La transformación IHS puede ser utilizada para combinar imágenes de diferentes sensores y de diferente resolución espacial, como en la unión de imágenes SPOT-HRV (pancromático) y TM-Landsat, por ejemplo.

El procedimiento para combinar imágenes de diferentes sensores consiste en calcular las componentes H, I, S a partir de tres bandas seleccionadas, por ejemplo, del TM y aplicar un contraste en las componentes H y S resultantes y en una imagen SPOT. La componente I es substituida por la banda del SPOT y se aplica la transformación inversa (IHS - RGB).

Después de la transformación, la imagen en color tendrá la resolución espacial de la imagen SPOT y la resolución espectral de las tres bandas TM.

El método anterior es uno de los más usados. Su limitación fundamental radica en la distorsión que provoca en las características espectrales de la imagen original. Tiene como ventaja la rapidez en la realización de las operaciones.

Frecuentemente se observa que las bandas individuales de una imagen multiespectral son altamente correlacionadas, o sea, las bandas son similares visual y numéricamente.

Esta correlación proviene del efecto de las sombras resultantes de la topografía, del solapamiento de las ventanas espectrales entre bandas adyacentes y del propio comportamiento espectral de los objetos.

El análisis de las bandas espectrales individuales puede ser ineficiente debido a la información redundante presente en cada una de esas bandas.

La *transformación por Sustitución de Componentes Principales* (PCS) es una técnica de realce que reduce o remueve esta redundancia espectral, o sea, genera un nuevo conjunto de imágenes cuyas bandas individuales presentan informaciones no disponibles en otras bandas.

Esta transformación se deriva de la matriz de covariancia entre las bandas y crea un nuevo conjunto de imágenes donde cada valor de pixel es una combinación lineal de los valores originales. El número de componentes principales es igual al número de bandas espectrales utilizadas y son ordenadas de acuerdo con el decremento de la variancia de nivel de gris. La primera componente principal tiene la mayor variancia (mayor contraste) y la última por consecuencia, la menor variancia.

Los autovalores representan la longitud de los ejes de las componentes principales de una imagen y son medidos en unidad de variancia. Existe un vector de módulo unitario asociado a cada autovalor denominado autovector. Los autovectores representan las direcciones de los ejes de las componentes principales. Son factores de ponderación que definen la contribución de cada banda original para una componente principal, en una combinación aditiva y lineal.

Para facilitar la percepción de esas contribuciones, se deben transformar los autovectores en porcentajes. Conociendo la señal de cada coeficiente de autovector, se pueden comparar los porcentajes con las curvas espectrales de materiales conocidos (por ejemplo, vegetación, agua, diferentes tipos de suelo), determinándose así en cuál(es) componente(s) principal(es) la información espectral deseada irá ser concentrada.

En el caso de la fusión aplicada a este procedimiento, consiste en dos procesos sucesivos, la sustitución de la intensidad en el espacio de color –que puede ser vista como la imagen de más alta resolución espacial-, por la primera componente principal, por una combinación de ésta con alguna banda multiespectral, o por una combinación de la primera y segunda componentes principales. El segundo paso consiste en realizar la transformación inversa para recomponer la imagen dando como resultado la imagen fusionada.

7.2 ESQUEMA DE FUSIÓN SHAN-YU

Se han realizado esfuerzos para el estudio y tratamiento de imágenes de percepción remota, uno de estos es aprovechar la información disponible de la zona en estudio. Una manera de hacerlo, es recopilando las imágenes adquiridas por varios sensores en diferentes épocas. Las imágenes satelitales son bastante complejas y un algoritmo único puede no tener éxito en el análisis de las imágenes con precisión y exactitud. En el esquema de fusión Shan-Yu se evalúa la calidad de cada resultado (error global) con el conocimiento previo de la información concerniente a la disposición geométrica de la zona de estudio (mapa), para después, basado en esta evaluación, se considera la fusión de estos resultados intermedios.

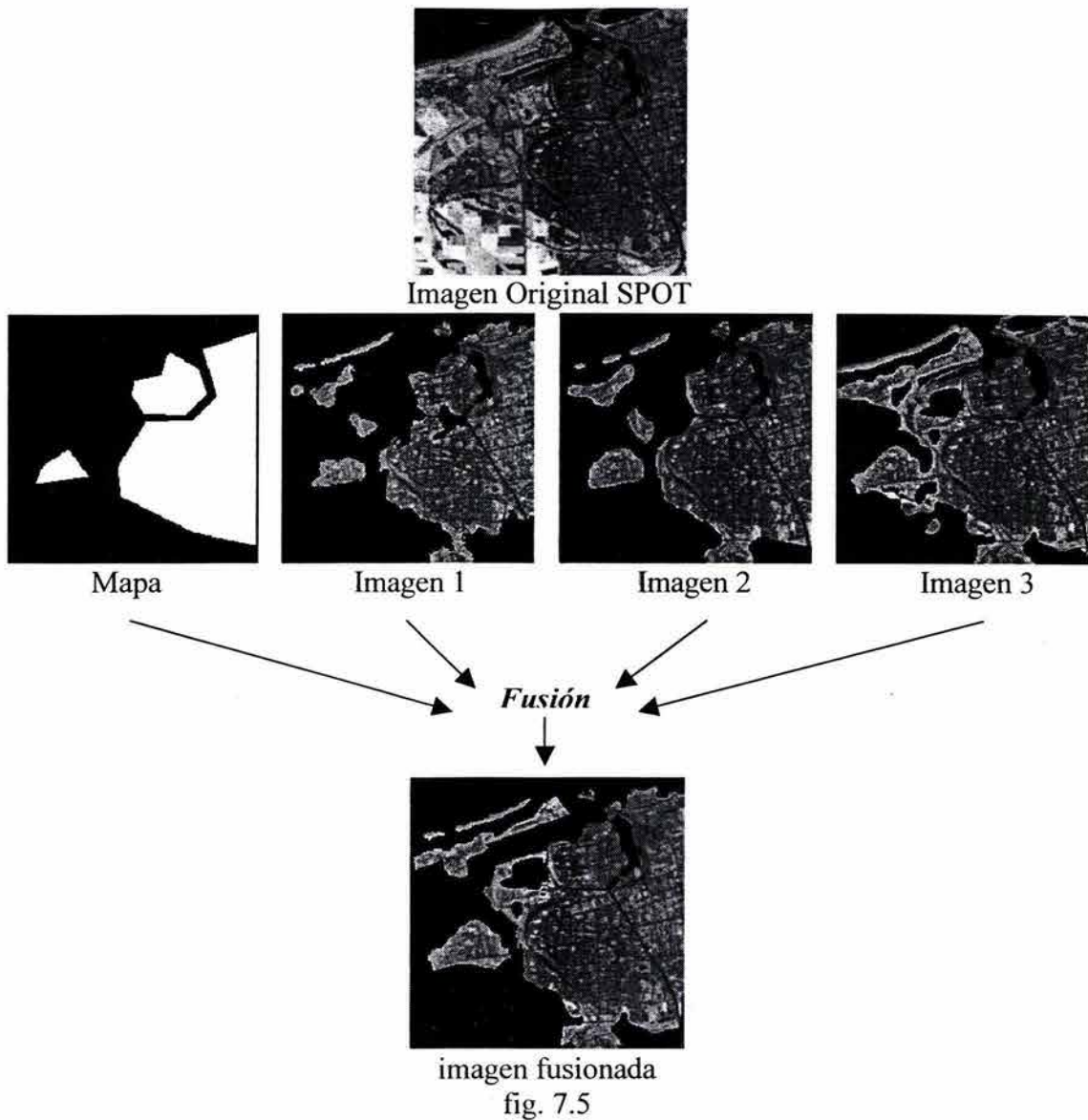


imagen fusionada
fig. 7.5

La información contextual en el análisis de cada *site* perteneciente a la imagen, se incorpora, al tomar en cuenta a los vecinos de éste. Para evaluar la calidad de un resultado al analizar una imagen, primero se calcula la medida de error local $E(s)$ para cada *site* s , para después sumarlos para obtener la medida de error global E_{global} para cada imagen. La medida de error global determina si la imagen es suficientemente buena para usarla en el proceso de fusión. La medida de error local relacionada a cada *site* determinará el grado de importancia para el proceso de fusión que tomará éste pixel. Por otra parte, se debe proporcionar como punto de inicio un esbozo de la zona geográfica a procesar, muchas veces esta zona es burda e imprecisa, lo cual es un elemento importante a tomar en cuenta para considerar el error local de cada *site* y error global en cada imagen.

Para cualquier $s \in S$, $Map(s)$ denota la etiqueta de un *site* s de acuerdo con la información contenida en el mapa, siendo $Img(s)$ la etiqueta del *site* s resultado del análisis. Denotando

$Map(s) = \lambda_i$ e $Img(s) = \lambda_j$, $\lambda \in L$, el cambio de etiqueta de λ_i (en el mapa) por λ_j (en el resultado del análisis de la imagen) incurre en un error que llamaremos riesgo:

$$x_{\lambda_i, \lambda_j} = \begin{cases} = 0 & \text{si } i = j \\ > 0 & \text{si } i \neq j \end{cases}$$

La función indicador $\mathbf{1}_{\{A(i)=x\}}$ de un conjunto A está definida por

$$\mathbf{1}_{\{A(i)=x\}} = \begin{cases} 1 & x \in A \\ 0 & \text{otro caso} \end{cases}$$

Un *site* no puede ser analizado sin considerar los *sites* vecinos. Sea s' un sitio en la vecindad V_s del *site* s , y $d(s, s')$ la distancia euclídeana entre s y s' , se calcula las siguientes dos métricas:

$$a(s) = \sum_{s' \in V_s} \frac{\mathbf{1}_{\{Map(s')=\lambda_j\}}}{d(s, s')} \quad \text{ec. 7.1}$$

$$b(s) = \sum_{s' \in V_s} \frac{\mathbf{1}_{\{Img(s')=\lambda_j\}}}{d(s, s')} \quad \text{ec. 7.2}$$

donde $\mathbf{1}_{\{\bullet\}}$ es la función indicador.

La medida de error local $E(s)$ del sitio s está definido como:

$$E(s) = \frac{x_{\lambda_i, \lambda_j}}{r \cdot a(s) + b(s) + 1} \quad \text{ec. 7.3}$$

donde r es un coeficiente. La medida de error global E_{global} es la suma del error local para todos los *sites* de la imagen.

Una vez obtenido la medida de error global se decide que imágenes se considerarán en el proceso de fusión. Para hacer esto, se toman los K mejores resultados de entre un conjunto de imágenes disponibles. Otra manera es usar un valor umbral h y se utilizan los resultados para la fusión donde $G_{global} < h$, no siendo trivial la elección de h .

Posteriormente al cálculo del error local para cada *site*, se calcula un coeficiente de confianza para cada uno, basándose en el resultado del error local:

$$C(s) = 1 - \frac{E(s)}{\sum_{\lambda_i, \lambda_j \in L} x_{\lambda_i, \lambda_j}} \quad 0 \leq C(s) \leq 1$$

ec. 7.4

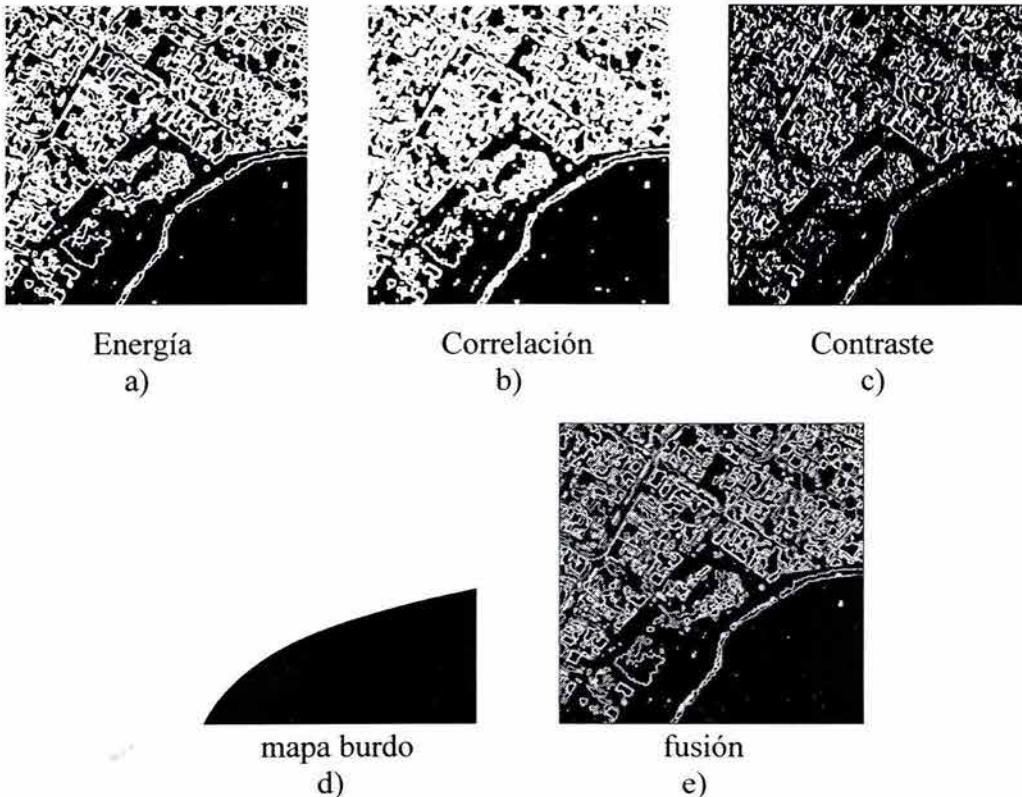
La decisión de la etiqueta final para cada *site* depende en la acumulación -suma- del coeficiente de confianza del mismo *site* a través de los resultados obtenidos:

$$\lambda(s) = \arg \max_{\lambda_i \in L} \sum_{k=1}^K C_k(s) \mathbf{1}_{\{Img^k(s) = \lambda_i\}}$$

ec. 7.5

donde $Img^k(s)$ es la k-ésima imagen entrada para el proceso de fusión y $C_k(s)$ es el coeficiente de confianza del sitio s en el resultado.

Visualmente, los resultados de la fusión se pueden observar en el siguiente ejemplo aplicado a descriptores de textura:



clases: {urbano, noubano}, adaptado de [MORA03]

fig. 7.6

8

ESQUEMA DE PROCESAMIENTO

8.1 ESQUEMA GENERAL

El esquema general de procesamiento para la detección remota de derrames de petróleo, se ilustra a continuación:

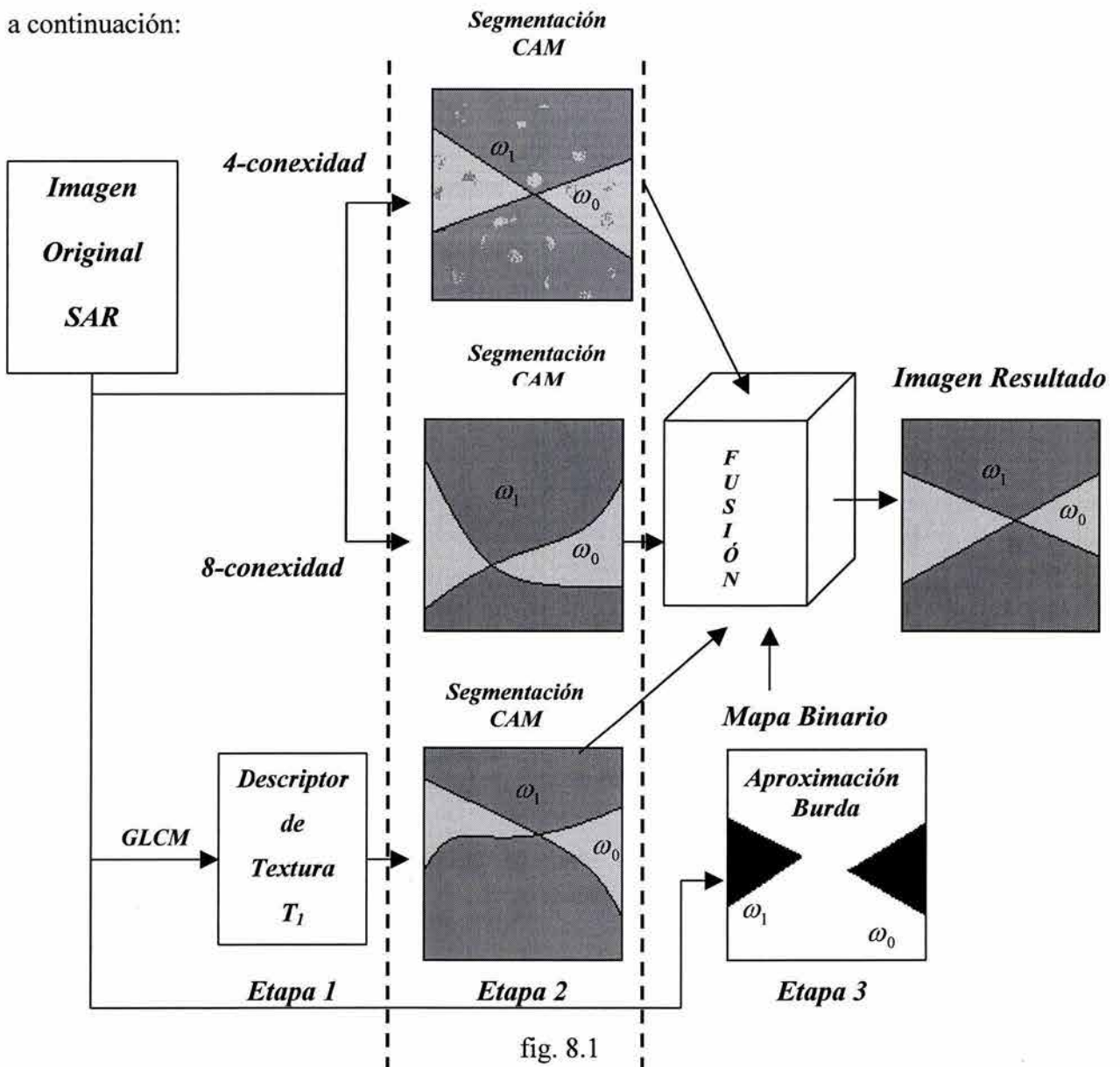


fig. 8.1

Debido a la naturaleza de las imágenes adquiridas por satélite, los esquemas tradicionales de procesamiento no resultan del todo exitosos, por lo que son necesarios esquemas más complejos para el análisis de las imágenes de percepción remota. El esquema general que se propone, está basado en la hipótesis de que diferentes imágenes segmentadas contienen información redundante e información complementaria entre sí.

El esquema general de procesamiento se divide en tres etapas, las cuales son:

1. *Cálculo del descriptor de textura para Imagen SAR*
2. *Segmentación por Campos Aleatorios de Markov:*

para la imagen SAR original:

- *en 4-conexidad*
- *en 8-conexidad*

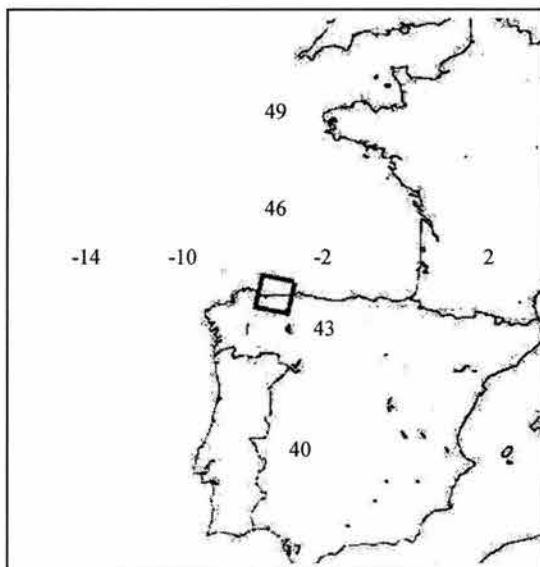
para la imagen del descriptor de textura:

- *en 8-conexidad*

3. *Fusión de las imágenes*

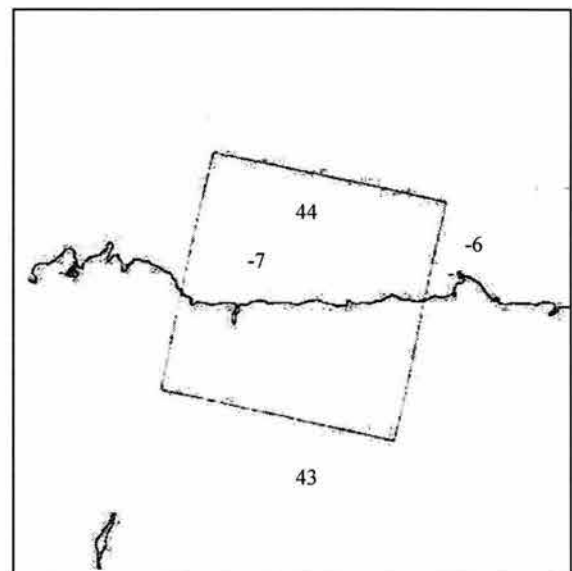
8.1.1 Selección de sitios

La imagen original fue adquirida utilizando el Radar de Apertura Sintética (SAR), perteneciente al satélite ERS-2, el día 22 de diciembre de 2002. La órbita de adquisición fue la 40114 y corresponde a la zona de Galicia y Asturias en España (fig. 8.2-a y 8.2-b). En la imagen de estudio se observan manchas de petróleo (fig. 8.2-c) como consecuencia del derrame del buque "Prestige".



zona de Galicia y Asturias en España

a)



vista en detalle de la zona del derrame

b)

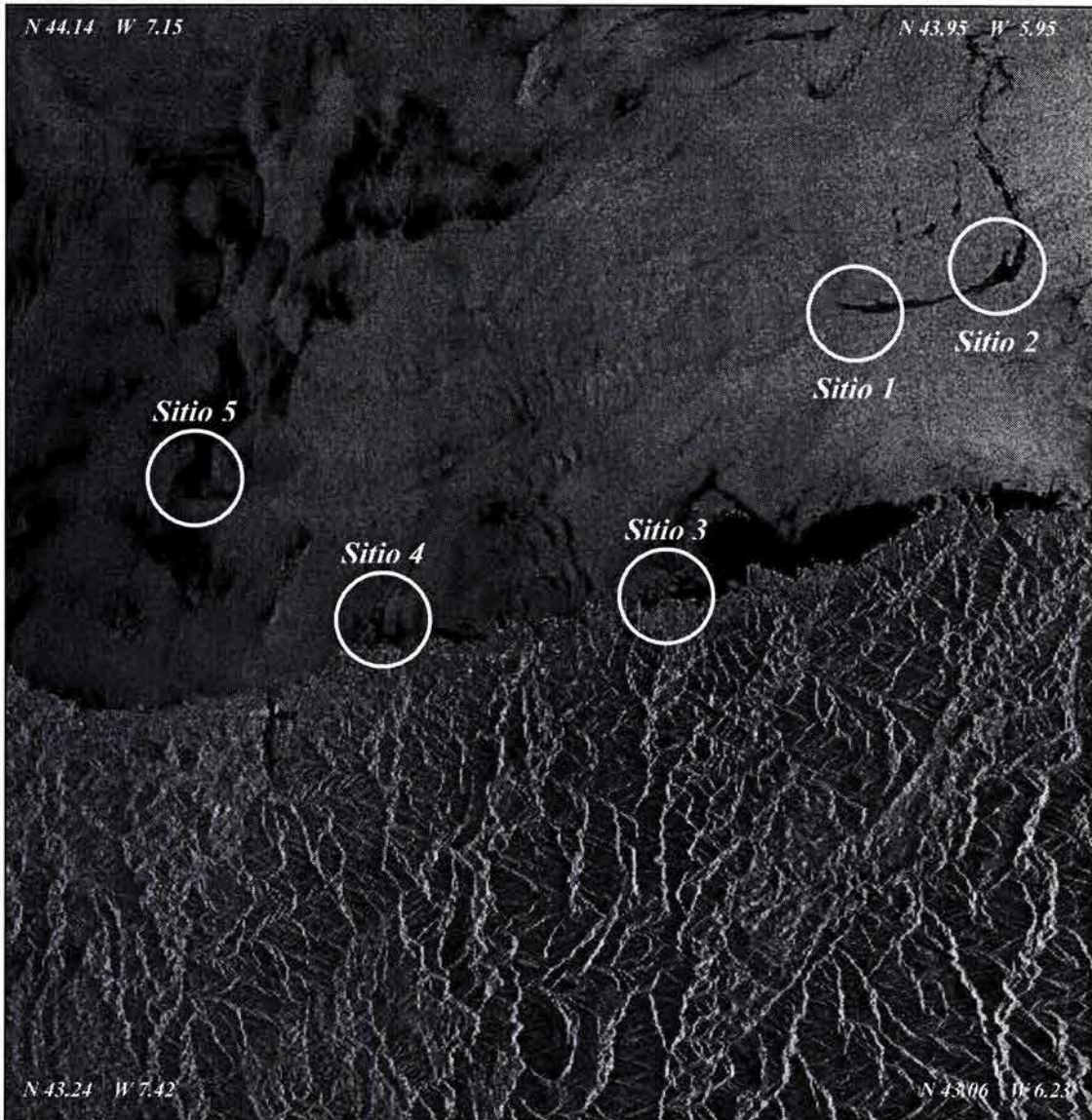


imagen SAR de la zona de estudio

c)

fig. 8.2

La figura original tiene una dimensión de 8131 pixels x 8000 pixels @ 8 bits/pixels, la imagen ha sido procesada para una la resolución espacial de 12.5 [m] y una resolución radiométrica de 256 niveles de gris.

Para la aplicación del esquema markoviano propuesto, se realizó la selección de cinco sitios representativos de la imagen. Las dimensiones de cada imagen es de 512 pixels x 512 pixels @ 8 bits/pixels. Esta selección se realizó considerando la geometría de la mancha y que la zona seleccionada fuera representativa de las diferentes escenas de la imagen. La ubicación de cada uno de los sitios se ilustra en la fig. 8.2-c.

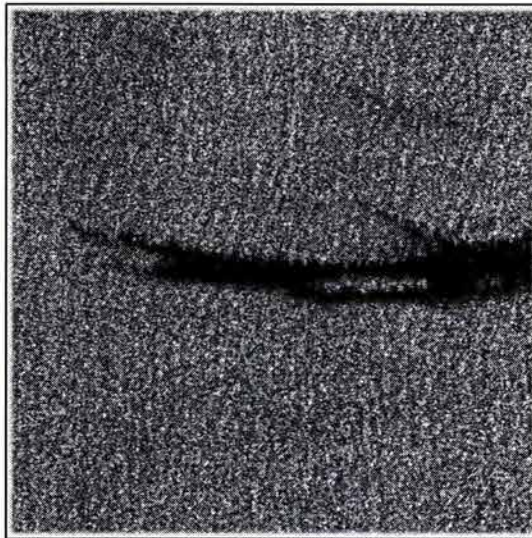
Sitio 1

fig. 8.3

Sitio 1: Este sitio presenta un buen contraste visual entre el mar y la mancha de petróleo. La forma geométrica es sencilla.

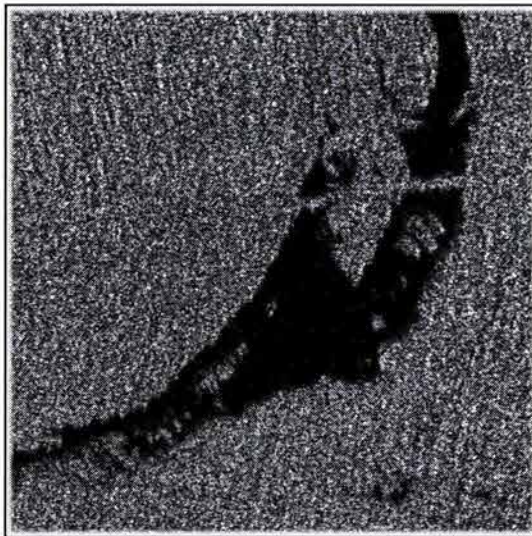
Sitio 2

fig. 8.4

Sitio 2: El sitio dos forma parte de la misma mancha de petróleo que el sitio 1. La geometría es más compleja que el sitio anterior, especialmente en la zona interna de la mancha.

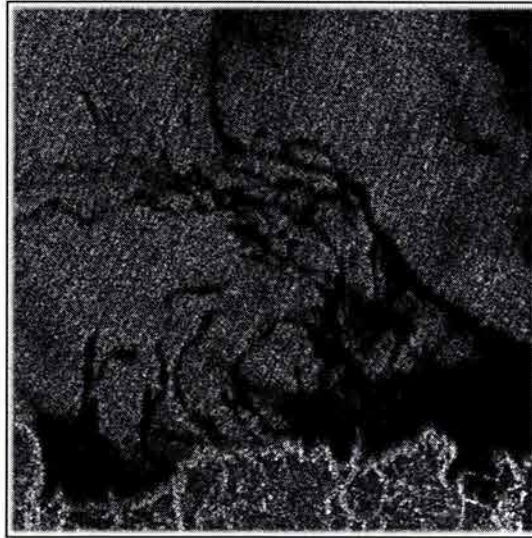
Sitio 3

fig. 8.5

Sitio 3: El sitio 3 contiene la geometría mas caprichosa de todos los sitios seleccionados, presenta zonas difusas entre las clases presentes.

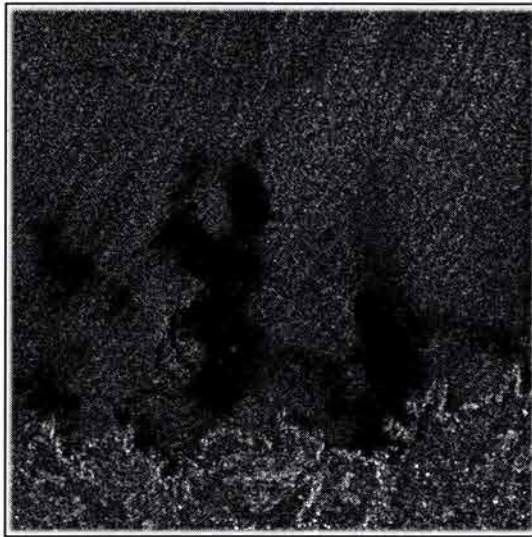
Sitio 4

fig. 8.6

Sitio 4: En esta imagen se muestra una geometría simple con zonas difusas.

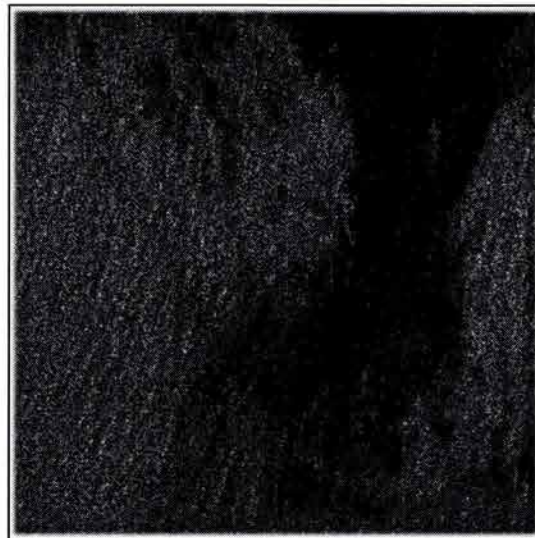
Sitio 5

fig. 8.7

Sitio 5: Se presenta un caso donde las clases llegan a confundirse. La discriminación visual entre clases es la más complicada de todos los sitios.

De acuerdo con información de la Agencia Espacial Europea*, el sitio 1 y el sitio 2, pertenecen con mucha seguridad a una mancha de petróleo, no siendo este el caso para los demás sitios.

8.2 ANÁLISIS DE TEXTURA

Existen diferentes aproximaciones para el análisis de texturas, en este caso, se seleccionó el esquema estadístico de la Matriz de Co-ocurrencia de Niveles de Gris (GLCM) propuesto por [HARA73].

La primera etapa del esquema markoviano propuesto consiste en el cálculo del descriptor de textura para la imagen SAR, que permita la discriminación entre clases: {*petróleo*, *nopetróleo*}.

Es importante considerar que la imagen resultado de esta parte del procesamiento, debería conservar los rasgos característicos de la geometría de la escena, ya que por ejemplo, existen estructuras de 5 pixels de grosor (fig. 8.4 y 8.5), que sería de interés preservar para la segmentación.

En el cálculo de la GLCM, son necesarios tres parámetros que intervienen en el procesamiento: *cuantización*, *desplazamiento* y *tamaño de la ventana*. No se considera la orientación como un parámetro de importancia.

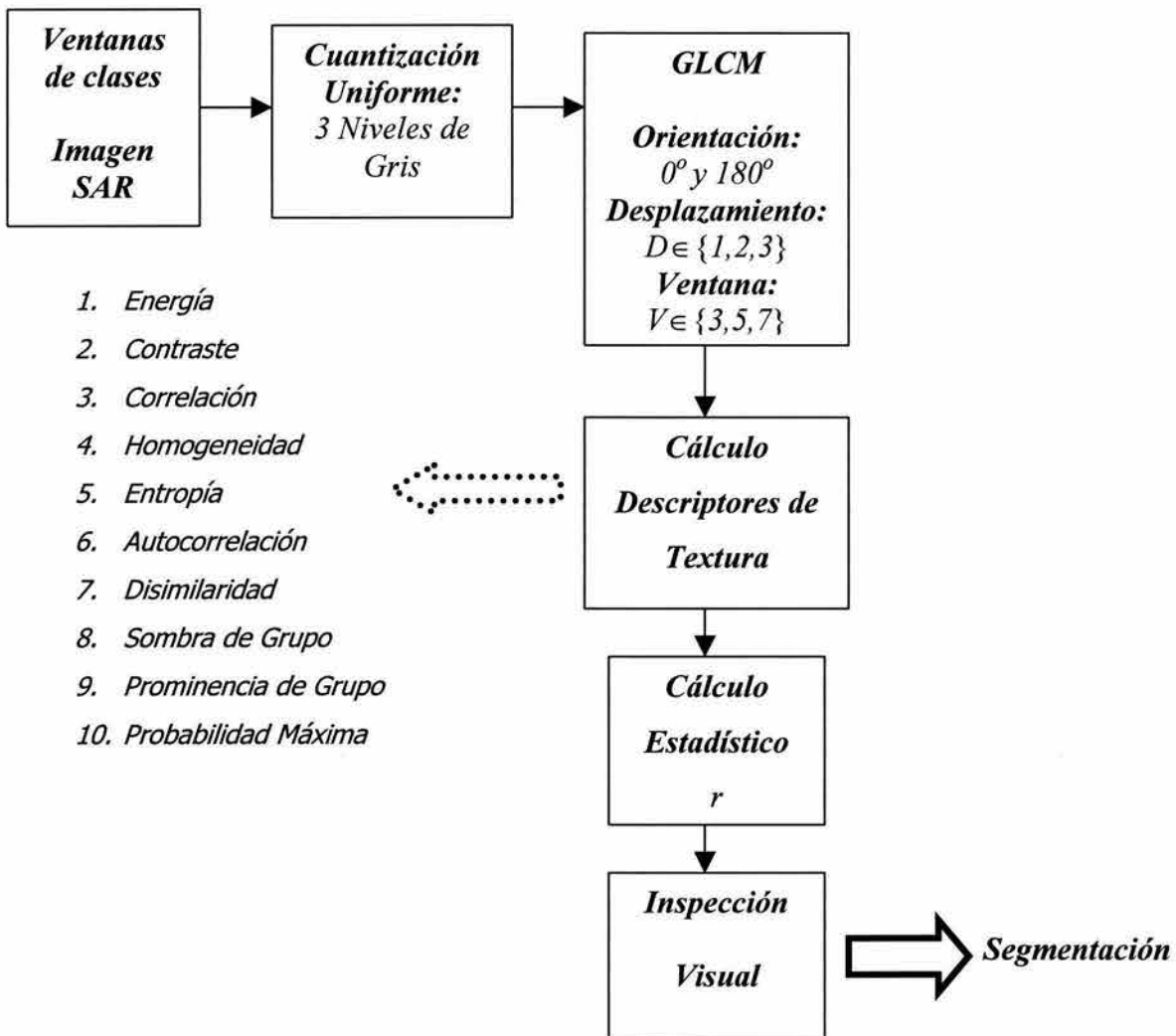
El esquema de *cuantización uniforme* fue seleccionado para la reducción de los niveles de gris de la imagen original. De forma heurística se propuso de 3 niveles de gris.

* http://earth.esa.int/ew/oil_slicks/galicia_sp_02/os_galicia_nov_02.html

El *tamaño de ventana* sobre la cual es calculada la GLCM es del orden de las estructuras que se desean detectar.

El valor de *desplazamiento* está limitado por el tamaño de la ventana. La elección del desplazamiento está restringido al mejor resultado visual que se obtenga dado el *tamaño de ventana* fijo.

De forma práctica, para la elección de los parámetros de *tamaño de ventana* y *desplazamiento* se realizó el siguiente procedimiento:



procedimiento para la elección de los parámetros de *tamaño de ventana* y *desplazamiento*
fig. 8.8

1.- En cada una de las imágenes de los sitios se realizó la selección de ventanas de 64 pixels x 64 pixels, cuyas medias y desviaciones estándar fueran representativas de las clases: {petróleo, nopetróleo}.

<i>Sitio</i>	$\mu_{\text{petróleo}}$	$\sigma_{\text{petróleo}}$	$\mu_{\text{nopetróleo}}$	$\sigma_{\text{nopetróleo}}$
1	18.5	5.6	101.7	34.8
2	19.2	5.7	100.6	34.3
3	17.1	5.0	62.5	22.5
4	18.6	6.0	61.4	20.4
5	18.5	5.8	53.1	18.0

tabla 8.1

Como se puede observar en la tabla 8.1, en el sitio 1 y el sitio 2 los valores de $\mu_{\text{petróleo}}$, $\sigma_{\text{petróleo}}$ y $\mu_{\text{nopetróleo}}$, $\sigma_{\text{nopetróleo}}$ son prácticamente iguales entre sí, el mismo caso se presenta en el sitio 3 y el sitio 4.

2.- A estas ventanas se les aplica el esquema de *cuantización uniforme*, para reducir de 256 niveles de gris a 3 niveles de gris.

3.- Para el par de ventanas de clases: {petróleo, nopetróleo} correspondientes a cada sitio, se calculó la GLCM para dos casos:

a) Modificando el *Tamaño de Ventana* y *Desplazamiento* fijo

Desplazamiento: 1

Tamaño de Ventana: 3, 5, 7

b) Modificando el *Desplazamiento* y *Tamaño de Ventana* fijo

Tamaño de Ventana: 7

Desplazamiento: 1, 2, 3

4.- Posteriormente se calculó el descriptor de textura correspondiente para cada ventana de entrenamiento y se escaló el resultado a 256 niveles de gris.

5.- Se calculó el estadístico r^{**} asociado a cada descriptor de textura para medir la separación entre clases:

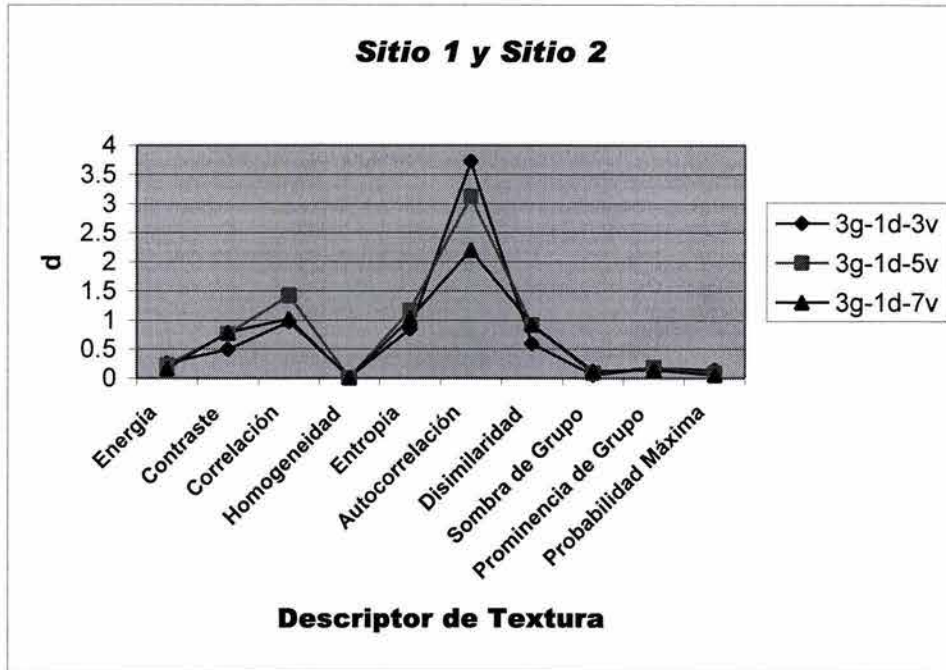
$$r = \frac{(\mu_{\text{petróleo}} - \mu_{\text{nopetróleo}})^2}{\sigma_{\text{petróleo}}^2 + \sigma_{\text{nopetróleo}}^2}$$

ec. 8.1

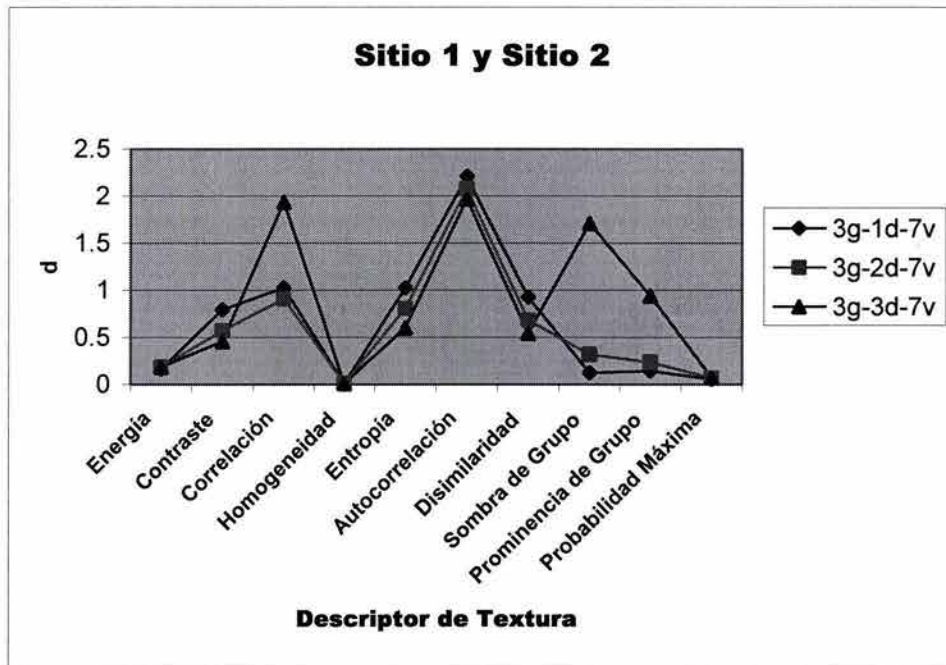
6.- Se seleccionaron los parámetros de los descriptores de textura que presentaran valores máximos.

7.- Se aplican los descriptores de textura correspondientes y se realiza la selección visualmente.

** el estadístico r es un caso particular de la distancia de Mahalanobis



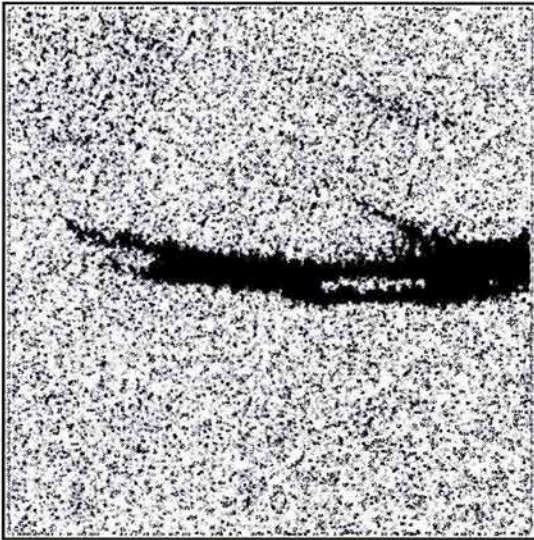
variación parámetro *tamaño de ventana*
a)



variación parámetro *desplazamiento*
b)
fig. 8.9

Los descriptores de Autocorrelación y Correlación, muestran tanto visual (fig. 8.10 y fig. 8.11) como estadísticamente una mejor discriminación entre clases.

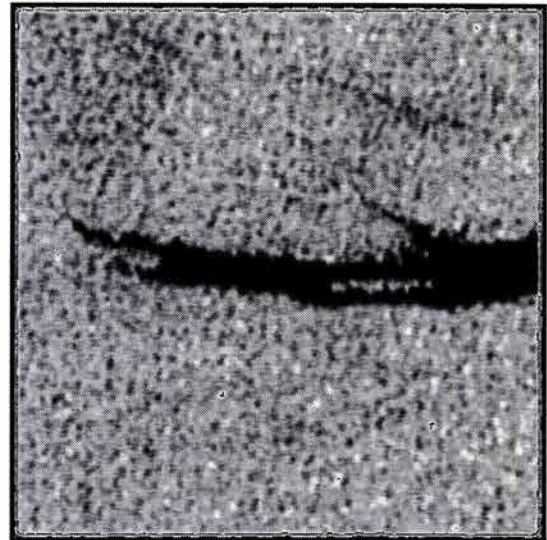
Sitio 1



Autocorrelación

a)

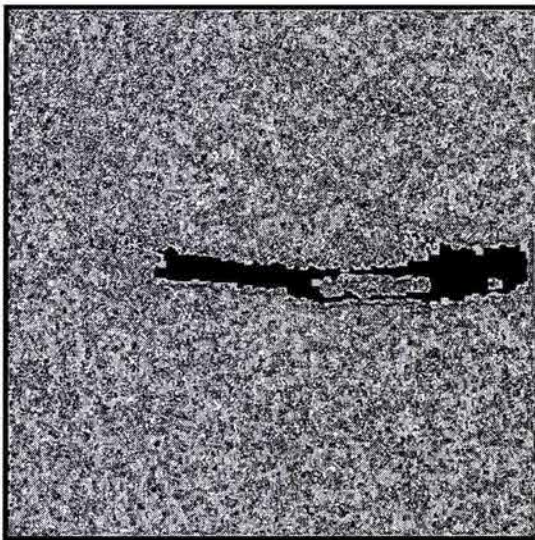
Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:3



Autocorrelación

b)

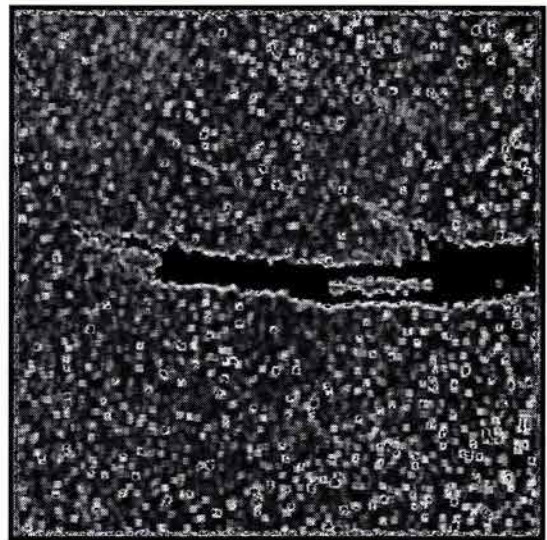
Niveles de Gris:3
Desplazamiento:3
Tamaño de Ventana:7



Correlación

c)

Niveles de Gris:3
Desplazamiento:3
Tamaño de Ventana:7



Sombra de Grupo

d)

Niveles de Gris:3
Desplazamiento:3
Tamaño de Ventana:7

fig. 8.10

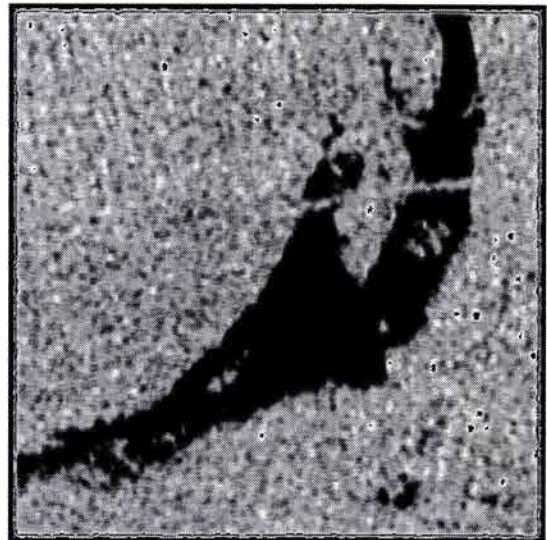
Sitio 2



Autocorrelación

a)

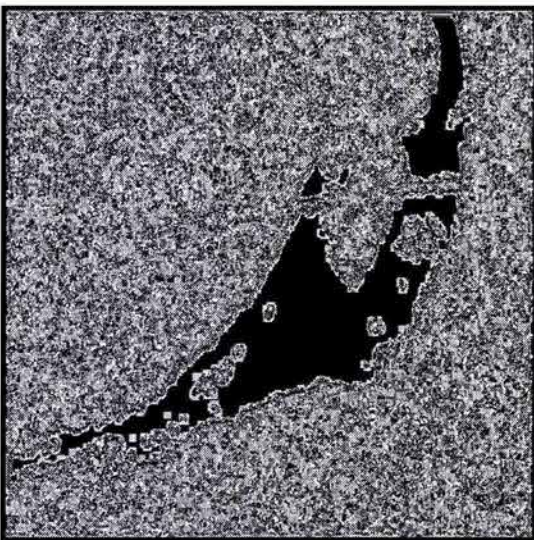
Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:3



Autocorrelación

b)

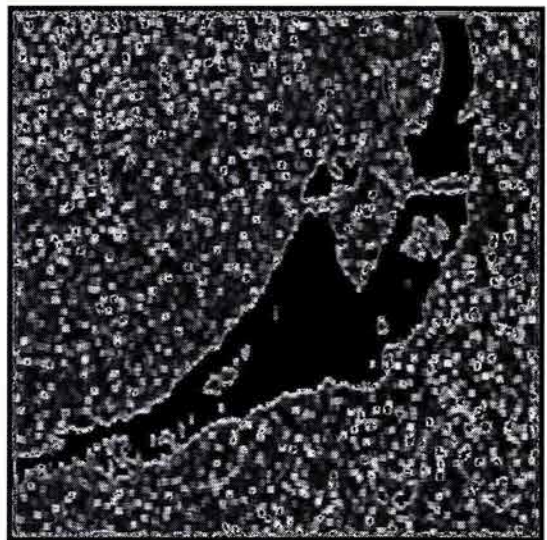
Niveles de Gris:3
Desplazamiento:3
Tamaño de Ventana:7



Correlación

c)

Niveles de Gris:3
Desplazamiento:3
Tamaño de Ventana:7

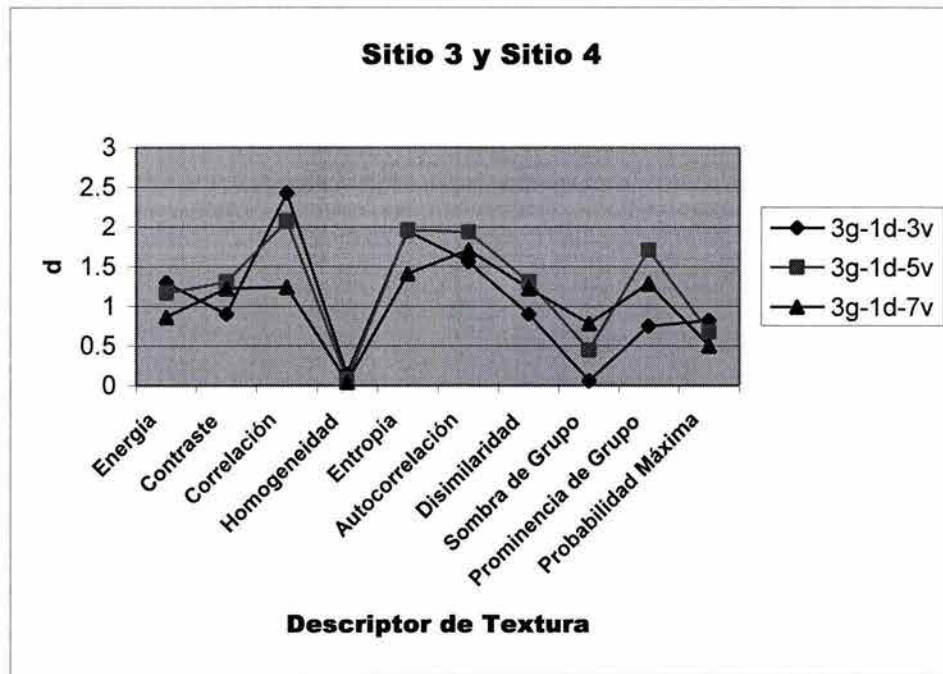


Sombra de Grupo

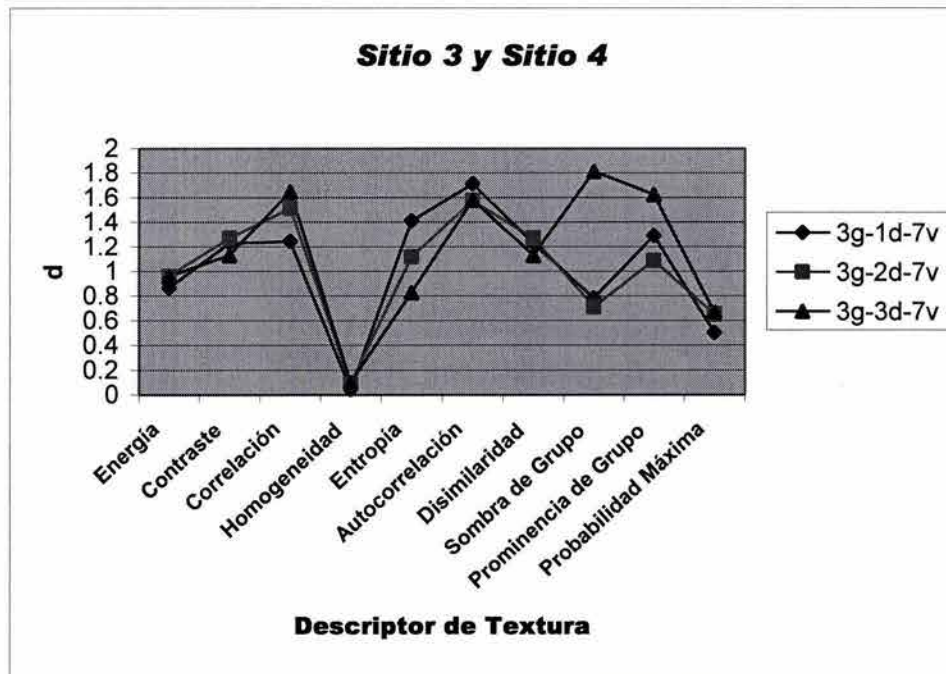
d)

Niveles de Gris:3
Desplazamiento:3
Tamaño de Ventana:7

fig. 8.11



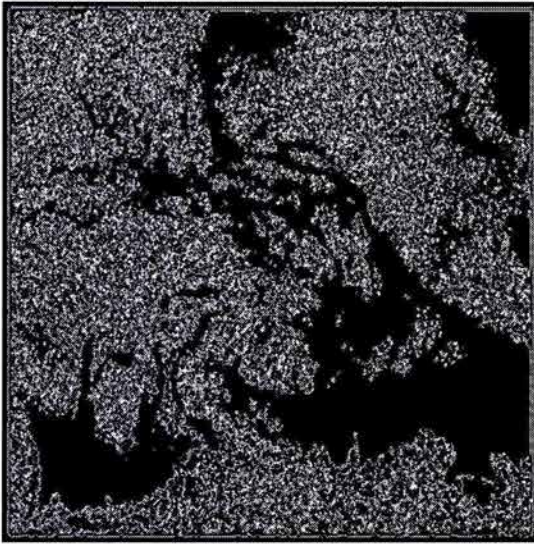
variación parámetro *tamaño de ventana*
a)



variación parámetro *desplazamiento*
b)
fig. 8.12

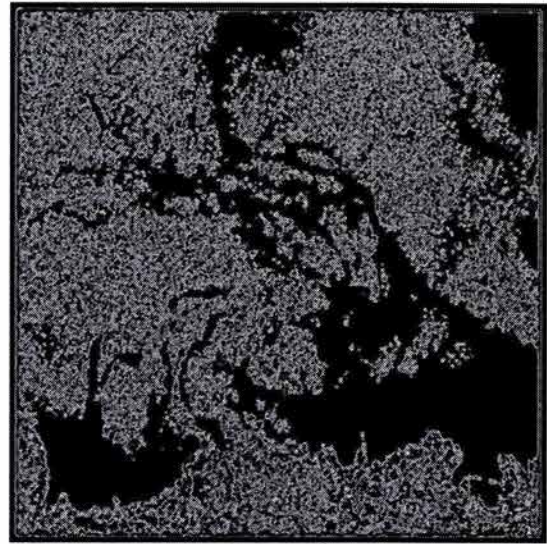
La forma general de la gráfica se modifica con respecto a la anterior. Los descriptores de Contraste y Prominencia de Grupo aumentan a razón de que la imagen es menos contrastada.

Sitio 3



Correlación

a)

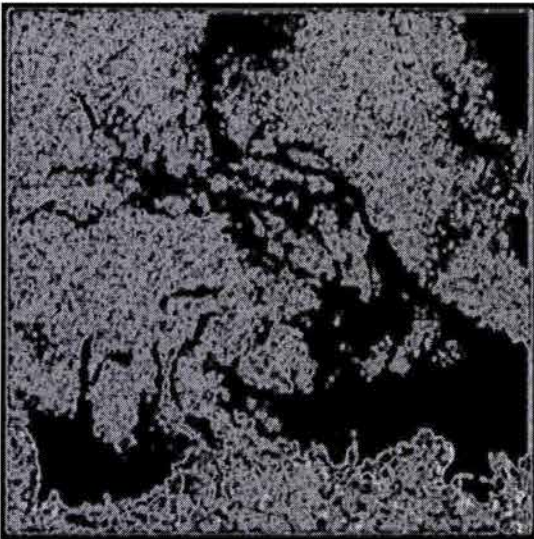


Entropía

b)

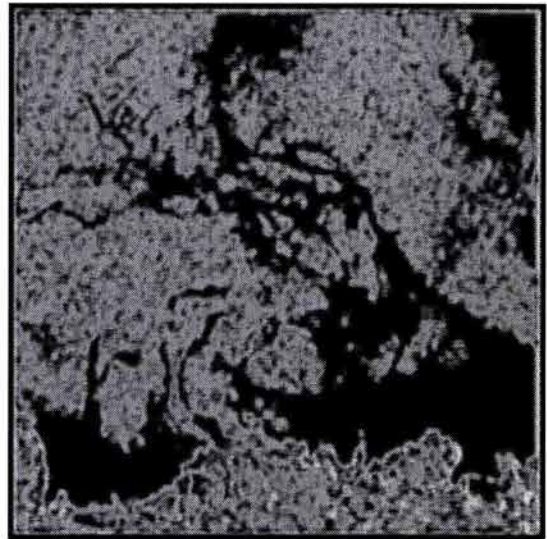
Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:3

Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:3



Entropía

c)



Entropía

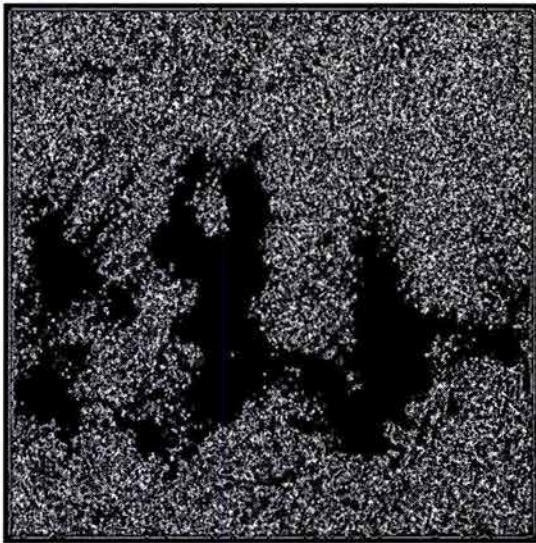
d)

Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:5

Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:7

fig. 8.13

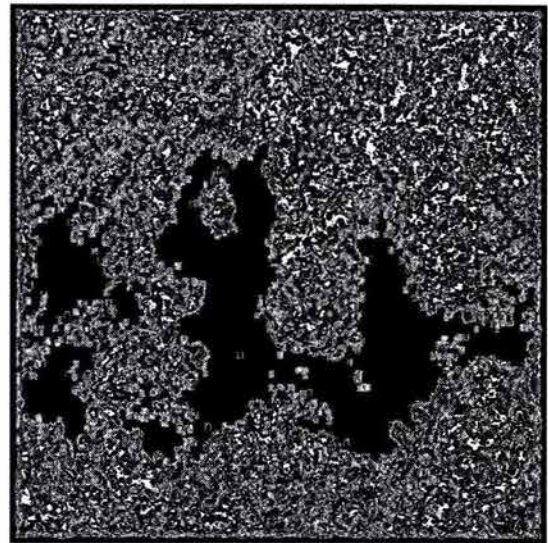
Sitio 4



Correlación

a)

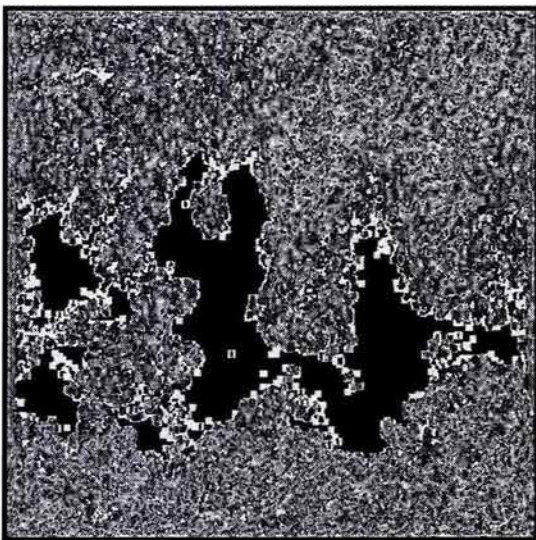
Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:3



Correlación

b)

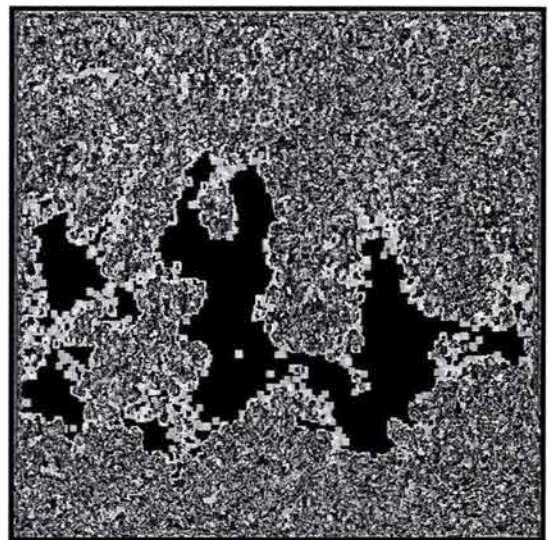
Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:7



Correlación

c)

Niveles de Gris:3
Desplazamiento:2
Tamaño de Ventana:7

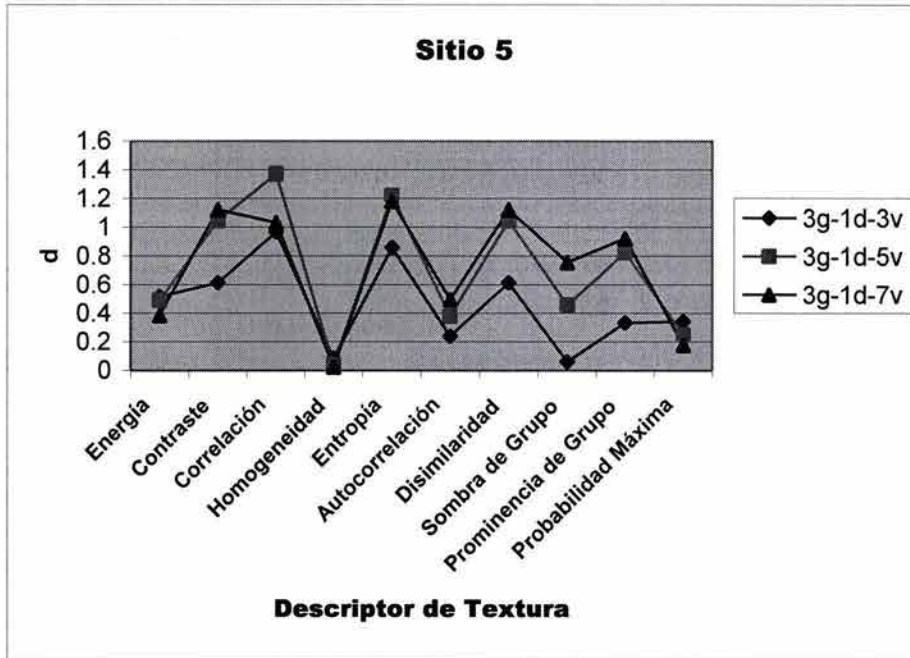


Correlación

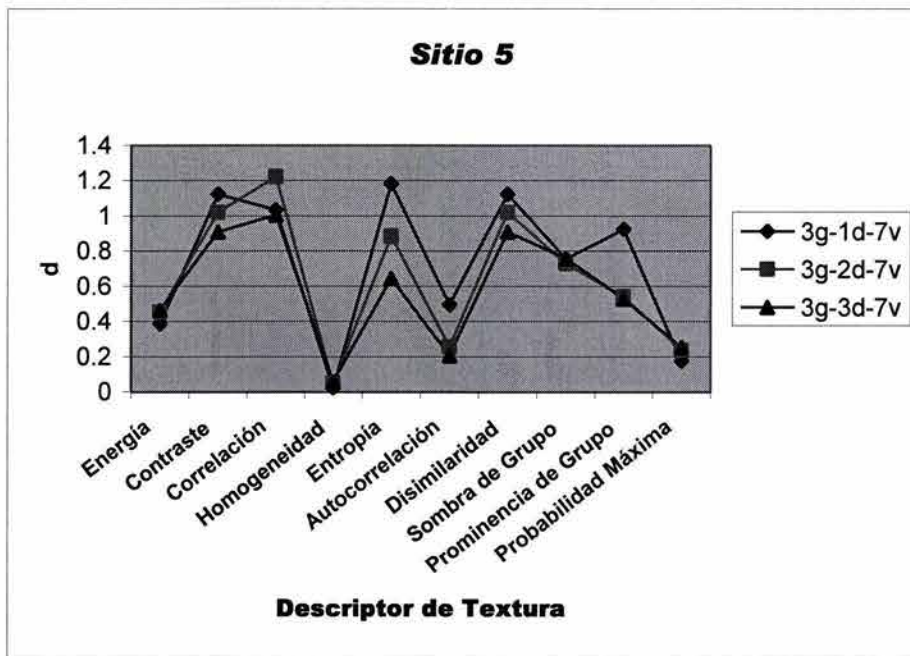
b)

Niveles de Gris:3
Desplazamiento:3
Tamaño de Ventana:7

fig. 8.14



variación parámetro *tamaño de ventana*
a)

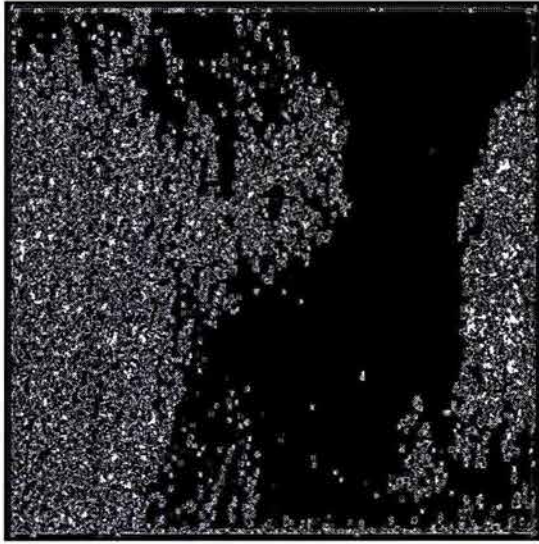


variación parámetro *desplazamiento*
b)

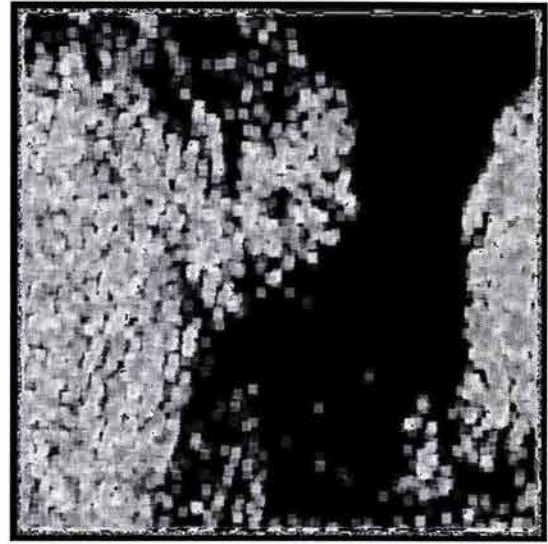
fig. 8.15

Este sitio es la imagen con menor contraste de todas. Los descriptores de Correlación y Prominencia de Grupo muestran visual y estadísticamente su capacidad para discriminar entre las clases.

Sitio 5



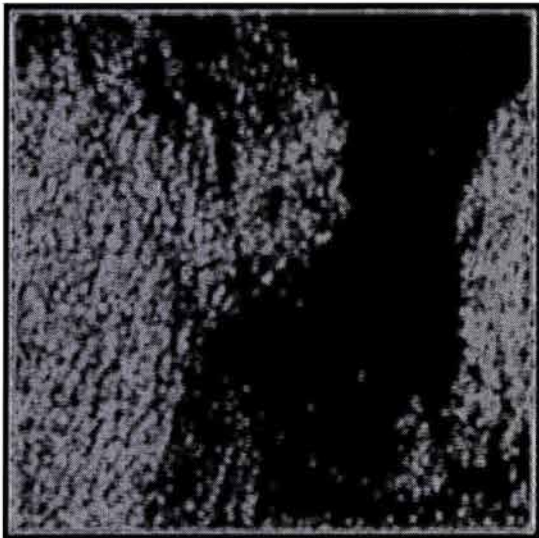
Correlación
a)



Prominencia de Grupo
b)

Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:5

Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:7



Entropía
c)



Disimilaridad
d)

Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:7

Niveles de Gris:3
Desplazamiento:1
Tamaño de Ventana:7

fig. 8.16

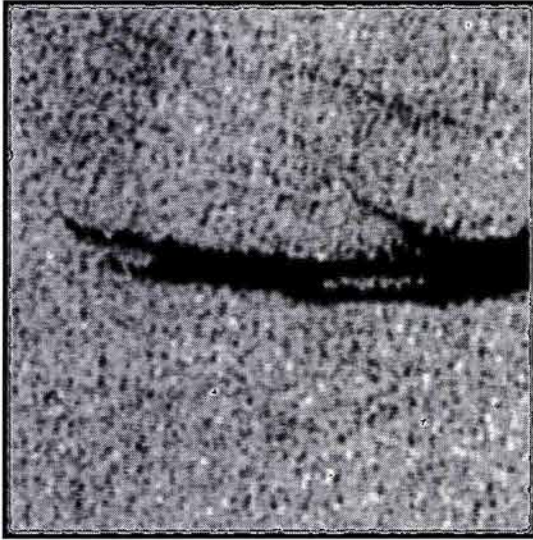
*Selección de Descriptores de Textura**Sitio 1*

fig. 8.17

Descriptor de Textura: *Autocorrelación*
 Niveles de Gris: 3
 Desplazamiento: 3
 Tamaño de la Ventana: 7

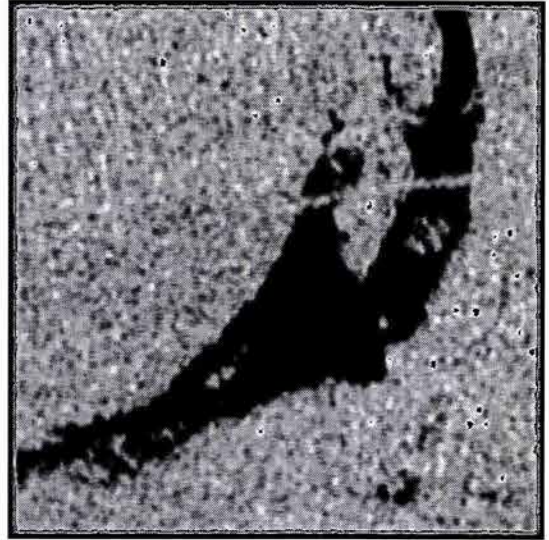
Sitio 2

fig. 8.18

Descriptor de Textura: *Autocorrelación*
 Niveles de Gris: 3
 Desplazamiento: 3
 Tamaño de la Ventana: 7

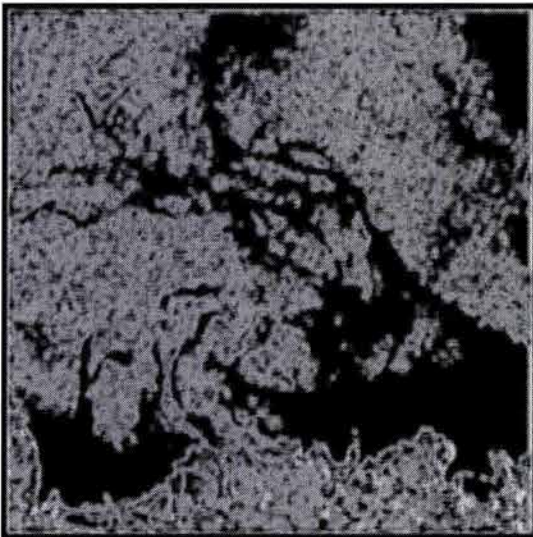
Sitio 3

fig. 8.19

Descriptor de Textura: *Prominencia de Grupo*
 Niveles de Gris: 3
 Desplazamiento: 1
 Tamaño de la Ventana: 7

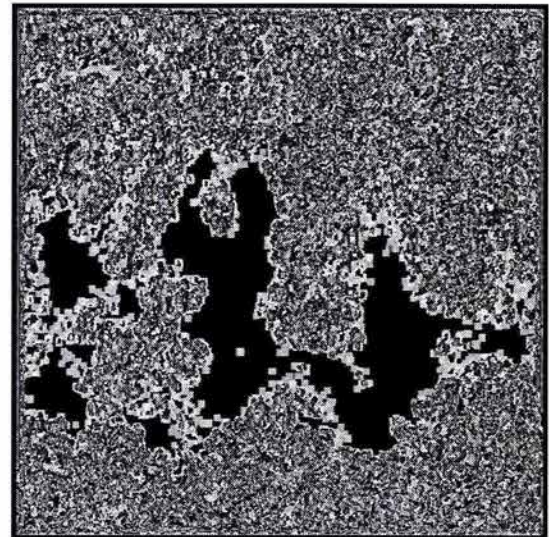
Sitio 4

fig. 8.20

Descriptor de Textura: *Correlación*
 Niveles de Gris: 3
 Desplazamiento: 3
 Tamaño de la Ventana: 7

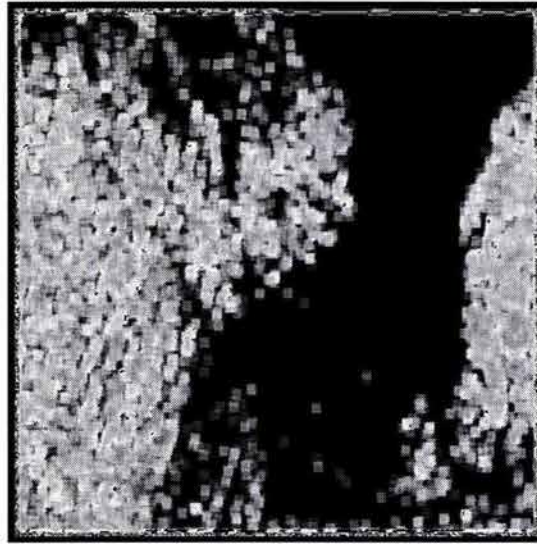
Sitio 5

fig. 8.21

Descriptor de Textura: *Prominencia de Grupo*
 Niveles de Gris: 3
 Desplazamiento: 1
 Tamaño de la Ventana: 7

La selección final de los sitios para la segmentación posterior se realizó de manera visual.

8.3 SEGMENTACIÓN MARKOVIANA

La segmentación se considera como un problema de etiquetado en el marco de estudio de los Campos Aleatorios de Markov. A cada uno de los *sites* $s \in S$ en una *lattice* regular (fig. 8.22-a), le corresponderá una etiqueta en L , dando lugar a una configuración (fig. 8.22-b).

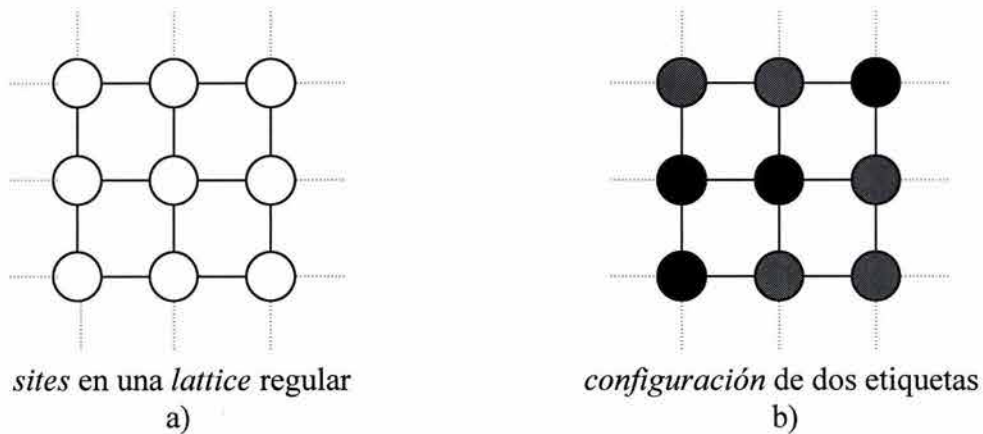
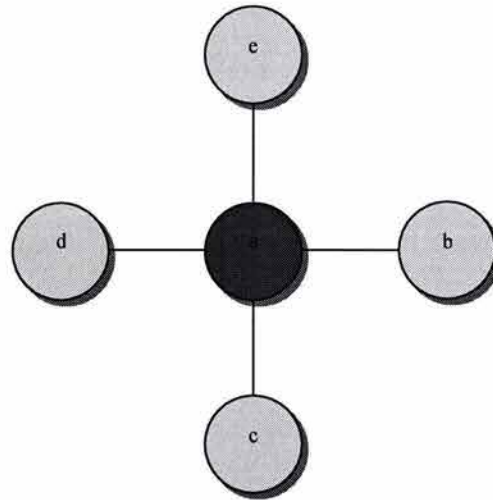


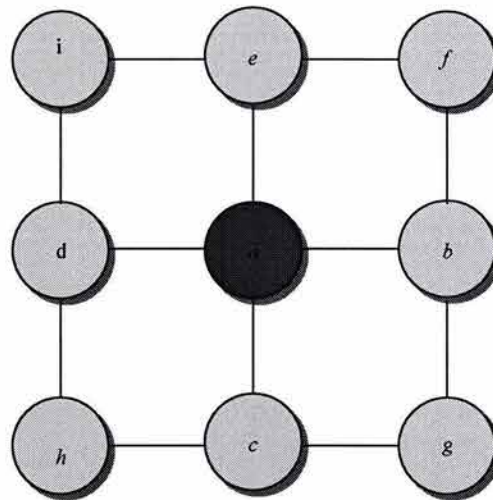
fig. 8.22

Cada *site* tiene un conjunto de posibles etiquetas $L = \{\lambda_1, \lambda_2\} = \{\text{petróleo}, \text{nopetróleo}\}$. Cada una de estas etiquetas corresponde a las clases en las que se desea asignar a los *sites*.

Los *sites* están relacionados entre sí por el sistema de vecindades. Las vecindades más comunes en segmentación de imágenes son las de primer y segundo orden.



vecindad de primer orden
a)

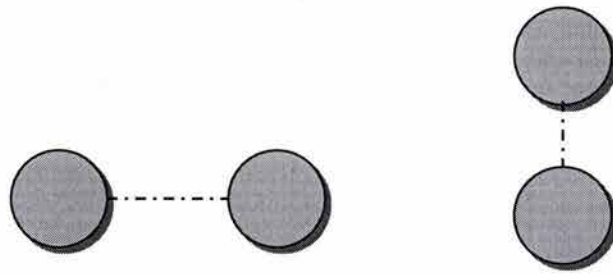


vecindad de segundo orden
b)

fig. 8.23

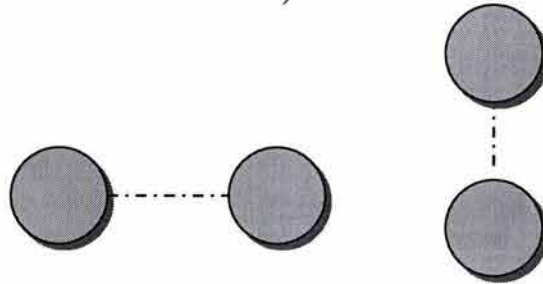
El grafo $G=(S,N)$ es llamado *clique*. Un *clique* c de G es un subconjunto de *sites* en S donde todos los *sites* son vecinos entre sí. C es el conjunto de todos los *cliques* de G y $C_a = \{c \mid s_a \in c\}$ el conjunto de *cliques* conteniendo el *site* a . Como se vió en 6.2.2, conforme el tamaño de la vecindad aumenta, el número de *cliques* crece considerablemente.

Los *cliques* utilizados para análisis en 4-*conexidad* y 8-*conexidad* son los siguientes:



cliques de dos *sites* asociados a una vecindad de primer orden

a)



cliques de dos *sites* asociados a una vecindad de segundo orden

b)

fig. 8.24

Como el análisis que se realizará es del *site* central con respecto a los *sites* vecinos y no entre vecinos solamente, uno de los *sites* perteneciente al *clique* deberá ser forzosamente el *site* central de estudio.

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

$$P(f_i | f_{S-(i)}) = P(f_i | f_{N_i})$$

ec. 8.2

Markovianidad

es decir, la probabilidad de que el *pixel* f tome el nivel de gris f_i , dado todos los niveles de gris f_{S-i} de todos los *pixels* restantes de la imagen, es igual a la probabilidad de que el *pixel* i tome el nivel de gris f_i dados los valores de los niveles de gris f_{N_i} de la vecindad de *pixels* N_i .

Un Campo Aleatorio de Markov es caracterizado por su propiedad local (Markovianidad) mientras que un Campo Aleatorio de Gibbs esta caracterizado por su propiedad global (distribución de probabilidad Gibbs). El teorema de *Hammersley-Clifford* establece la equivalencia entre estas dos propiedades. La importancia de esta equivalencia es que de forma práctica, un CAM queda definido con la distribución de Gibbs.

Una distribución de Gibbs es de la siguiente forma:

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

ec. 8.3

donde

$$Z = \sum_{f \in \Omega} e^{-\frac{1}{T} U(f)}$$

ec. 8.4

T: es la temperatura , $T=1$

$U(f)$: función de energía

Z: función de partición (normalización)

f : es una configuración ***

Como se puede observar de la ecuación anterior, las configuraciones con menor energía *a priori* tiene una mayor probabilidad de ocurrencia.

La fórmula de Bayes implica la probabilidad condicional de tener la clase Y dada la observación X :

$$P(Y | X) = \frac{P(X | Y)P(Y)}{P(X)}$$

ec. 8.5

$P(Y | X)$: probabilidad a posteriori

$P(Y)$: probabilidad a priori de la clase Y

$P(X | Y)$: probabilidad conjunta de X dado Y

$P(X)$: distribución de probabilidad de X

*** Se entiende por configuración a un arreglo particular de etiquetas

$P(Y)$ define un Campo Aleatorio de Markov, de acuerdo con el teorema de *Hammersley-Clifford* y por lo tanto $P(Y)$ puede ser expresada como una distribución de Gibbs:

$$P(Y) = \frac{1}{Z} \exp(-U(Y))$$

ec. 8.6

Z no proporciona información discriminante, por lo que no se considera.

La función de energía *a priori* es la función que permite codificar las restricciones contextuales.

La fórmula que la describe es la siguiente:

$$U(Y) = \sum_{c \in C} V_c$$

ec. 8.7

Para el cálculo de esta función solo se considerarán los *cliques* de segundo orden (dos *sites* vecinos entre sí). Este tipo de convención es ampliamente usada por su forma simple y su bajo costo computacional. De acuerdo con la fig. 8.23-a), si se presenta una vecindad de 4-vecinos para su análisis le corresponden 4 *cliques* de segundo orden (fig. 8.24-a). De forma análoga, a la vecindad de segundo orden o sistema 8-vecinos le corresponden 8 *cliques* de segundo orden.

Otra consideración del modelo de Campo Aleatorio de Markov es la independendencia con respecto a la posición del *clique* en S (homogeneidad) y la orientación del *clique* (isotropía).

El modelo de *Ising* se presenta para el caso donde el análisis es solamente para dos etiquetas (clases), en este caso $\{\text{petróleo}, \text{nopetróleo}\}$. La función *potencial de clique* queda definida por:

$$V_c = -\beta \delta(y_i - y_j)$$

ec. 8.8

el *clique* c contiene a los *sites* i y j , donde y_i e y_j son niveles de gris.

Del teorema de Bayes se considera que el término $P(X)$ no presenta información discriminante, realizando la siguiente aproximación vista en 5.5:

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

ec. 8.7

El método de MAP consiste en encontrar la configuración y que maximiza la distribución *a posteriori* $P(Y|X)$, o bien minimiza una función equivalente de energía:

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

ec. 8.8

por lo que la ec. 8.7 toma la forma:

$$U(Y | X) \approx U(X | Y) + U(Y)$$

ec. 8.8

De forma práctica, en las imágenes SAR se definieron ventanas para caracterizar las medias de las dos clases $\{\mu_{\text{petróleo}}, \mu_{\text{nopetróleo}}\}$. Las funciones de energía conjunta $U(X | Y)$ están dadas por la aproximación de la fig. 8.25:

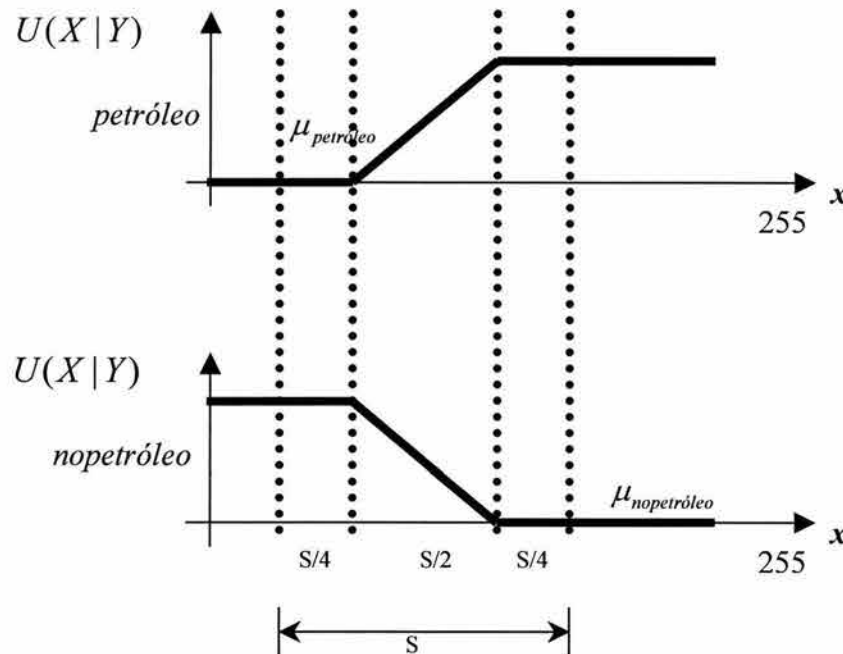


fig. 8.25

La energía *a priori* $U(Y)$ se calcula dependiendo si el análisis es en 8-conexidad ó 4-conexidad, la función de energía *a priori* $U(Y)$ es de la siguiente forma:

$$U(Y) = -\beta \times n_i(Y)$$

ec. 8.9

donde $n_i(I)$ es el número de *sites* en la vecindad N_i que tienen etiqueta Y . β es el coeficiente de interacción.

Los Campos Aleatorios de Markov están ligados a un método específico de optimación, el *recocido simulado*, el cual permite resolver el problema de estimación en el sentido MAP, buscando la convergencia hacia un *mínimo global* de la función de energía a posteriori.

La simulación del algoritmo de recocido simulado es usada para minimizar la energía *a posteriori* y encontrar la configuración $Y^{(n)}$. Esta configuración es la imagen segmentada y es la solución al problema de etiquetado.

El algoritmo de recocido simulado queda descrito a continuación:

Algoritmo de Recocido Simulado

1. Inicialización

- Determinación aleatoria de la solución $Y^{(0)}$
- Escoger una temperatura inicial $T^{(0)}$ suficientemente elevada

2. Barrido secuencial de 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 sorteo:

variable aleatoria con distribución uniforme $\xi \in \Omega$

3. Criterio de Metrópolis

- Calcular variación de la energía local asociada al site r:

$$\Delta U = U(\xi | V_r^{(n)}, X) - U(Y_r^{(n)} | V_r^{(n)}, X)$$

- Pueden surgir dos casos

$$\text{Si } \Delta U \leq 0 \Rightarrow Y_r^{(n+1)} = \xi$$

$$\text{Si } \Delta U > 0 \Rightarrow Y_r^{(n+1)} = \xi \text{ con probabilidad } \xi = e^{-\frac{\Delta U}{T}}$$

4. Continuar con la etapa 2

5. Serie de temperaturas decrecientes hacia 0, por lo que $\lim_{n \rightarrow \infty} T_n = 0$

El método bayesiano asociado a la modelización markoviana proporcionan una función de energía *no convexa* que debe ser minimizada. La aplicación del recocido simulado en la optimización de la función de energía *a posteriori* $U(Y|X)$ asegura la convergencia hacia el mínimo global y evitando la convergencia a los mínimos locales. El cambio de la clase del pixel analizado a la clase propuesta, es aceptado incondicionalmente sólo cuando $\Delta U < 0$.

La simulación del algoritmo de recocido simulado es el siguiente:

Simulación Recocido Simulado

1. Inicialización

- Determinación aleatoria de la solución $Y^{(0)}$.
La clase de cada píxel debe determinarse mediante un generador de números aleatorios, con distribución uniforme, en el rango del número de clases, $\text{clases}=2$.
- Escoger un temperatura inicial $T^{(0)}$ suficientemente elevada, $T^{(0)}=2$.

2. Hacer desde $N_{iter}=1$ hasta $N_{iter}=n$, con $n=60$

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 $\xi \in \Omega=[1,2]$.

2.2 Análisis en 8-conexidad/4-conexidad (según sea el caso)

- $\beta = 0.35$
- Si “actual”=“x”: $\text{prob_actual} = \text{prob_actual} - \beta [\text{actual}]$
- Si “elección”=“x”: $\text{prob_elección} = \text{prob_elección} - \beta [\text{elección}]$

2.3 Calcular la variación de la energía local asociada al site r :

- $\text{prob_actual} = \text{prob_actual} + \text{potencial}[\text{actual}]_{ij}$
- $\text{prob_elección} = \text{prob_elección} + \text{potencial}[\text{elección}]_{ij}$

donde $\text{potencial}[]_{ij}$ representa la energía condicional $U(X|Y)$ e i, j son las coordenadas del píxel analizado.

Si $\text{prob_elección} \leq \text{prob_actual}$:

$Y_{ij} = \text{elección}$

En caso contrario:

- $p = \exp[-(\text{prob_elección} - \text{prob_actual})/T^{(n)}]$
- sorteo de una variable aleatoria ξ_2 , uniformemente distribuida en el rango $[0,1]$
- si $\xi_2 < p$, entonces $Y_{ij} = \text{“elección”}$

3 Regresar al punto 2.1 hasta terminar con el barrido

4 Decrementar la temperatura:

$$T^{(n)} = 0.95 \cdot T^{(n-1)}$$

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

Es conocido que el recocido simulado converge con probabilidad unitaria hacia el óptimo global si la ley de temperatura T_k decrece menos rápidamente que $C/\ln(k)$, para una cierta constante C independiente de k . Debido a la lentitud en convergencia del esquema $C/\ln(k)$, en la práctica, la ley de temperatura se aproxima por una ley exponencial. A causa de esta aproximación, la convergencia hacia el óptimo global está fuera de garantía.

La temperatura inicial T_0 debe ser escogida tal que todas las transiciones pueden ser aceptadas con una probabilidad diferente a cero. Es muy difícil determinar tal valor ya que se encuentra ligado a los valores máximos y mínimos de la función de energía. En la práctica se escoge un valor T_0 relativamente bajo para asegurar una convergencia rápida. Así por ejemplo, un valor de 2 resulta generalmente adecuado.

8.3.1 Segmentación imágenes SAR

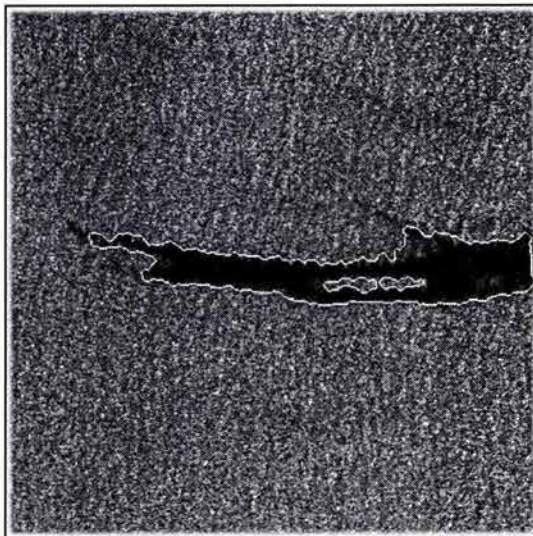
Se realiza la segmentación de cada una de las imágenes originales SAR bajo el esquema estudiado anteriormente. Las imágenes utilizadas en la segmentación son las seleccionadas en 8.1.1. El análisis para la función de energía *a priori* se realiza en 4-conexidad y 8-conexidad.

El resultado de la segmentación presenta solamente dos clases: $\{\text{petróleo}, \text{nopetróleo}\}$. Para mostrar los resultados de la segmentación sobre la imagen original, se delimitó la frontera entre las clases por medio de una línea blanca en la imagen original.

Los parámetros para el algoritmo de segmentación fueron:

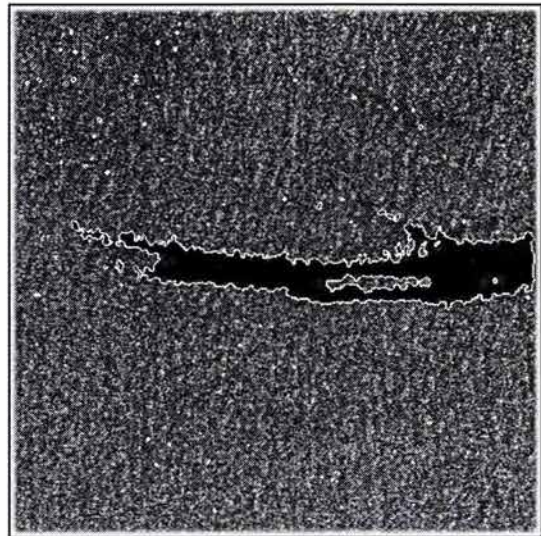
$$\begin{aligned} T &= 2 \\ \beta &= 0.35 \\ \text{Iteraciones} &= 60 \end{aligned}$$

Sitio 1

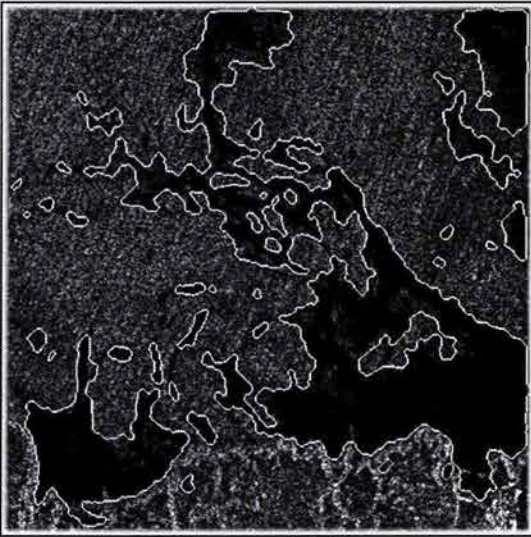
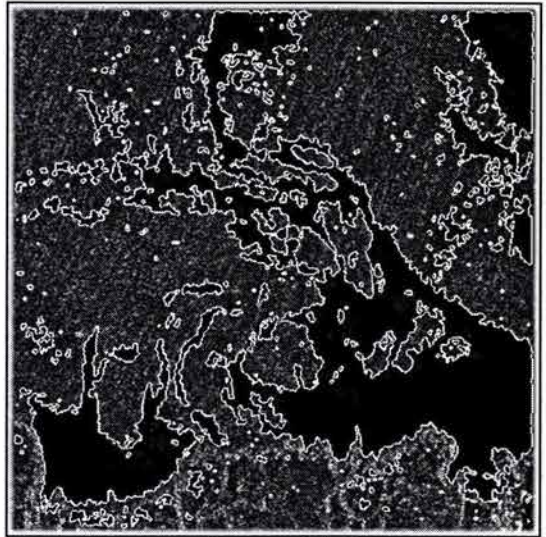


8-conexidad
fig. 8.28

Sitio 1

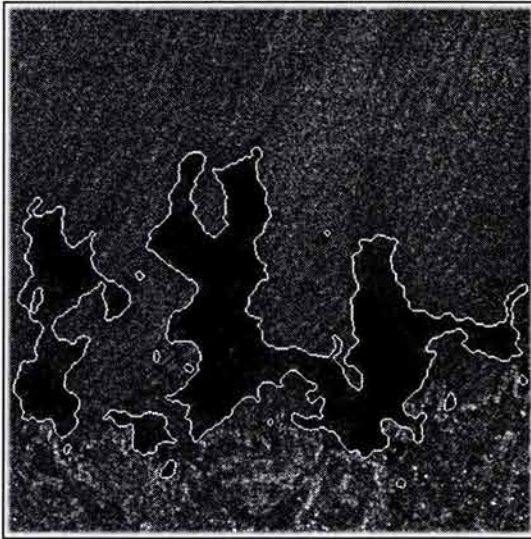
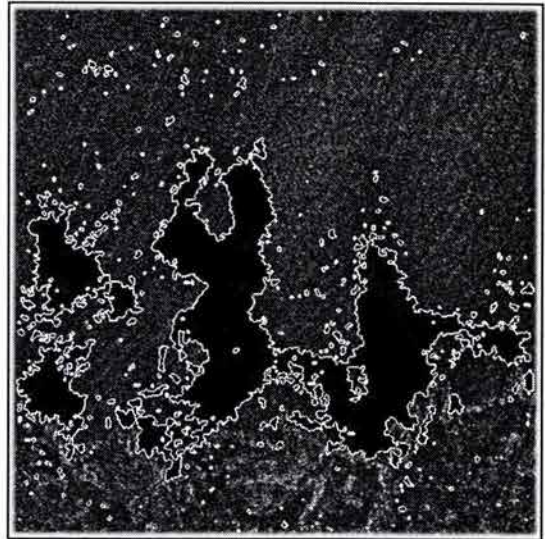
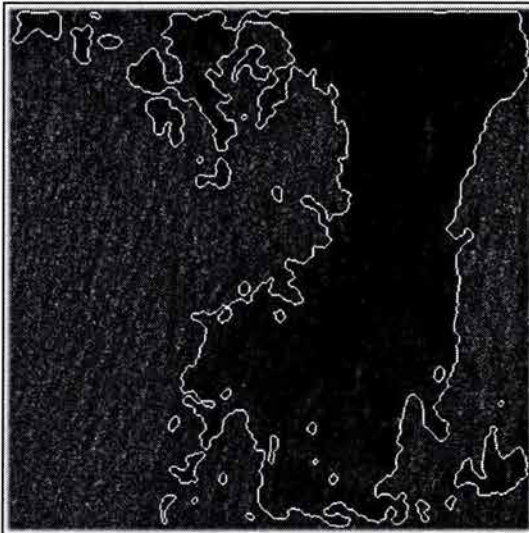
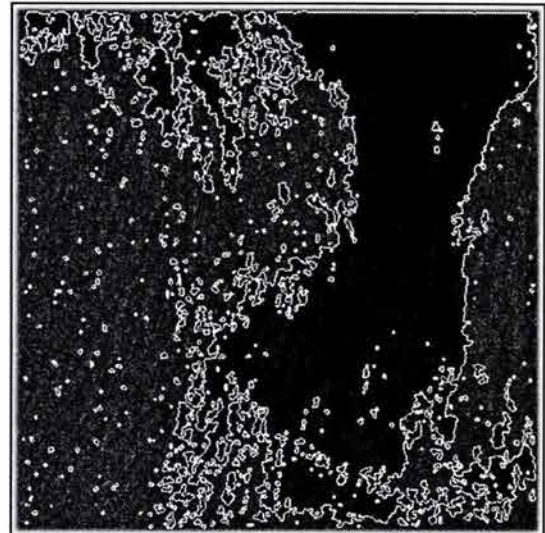


4-conexidad
fig. 8.29

Sitio 28-conexidad
fig. 8.30*Sitio 2*4-conexidad
fig. 8.31*Sitio 3*8-conexidad
fig. 8.32*Sitio 3*4-conexidad
fig. 8.33

El análisis en *8-conexidad* arroja diferentes resultados que el análisis en *4-conexidad*. La mancha de petróleo queda definida de forma general por el primer esquema de análisis, sin embargo, el segundo esquema de análisis brinda mayores detalles de la imagen.

Para imágenes sin muchos detalles, el esquema clásico de *8-conexidad* es una buena aproximación.

Sitio 48-conexidad
fig. 8.34*Sitio 4*4-conexidad
fig. 8.35*Sitio 5*8-conexidad
fig. 8.36*Sitio 5*4-conexidad
fig. 8.37

Al realizar el análisis de *conexidad* y considerar el tamaño de vecindad, el algoritmo tiende a unir zonas aisladas del orden de la vecindad y desaparecer las “islas” de tamaño menor a la vecindad.

Esta característica en el análisis de *8-conexidad* reduce la capacidad para brindar detalles sobre la mancha pero brinda zonas mas homogéneas.

8.3.2 Segmentación descriptores de textura

Se realiza la segmentación de cada una de las imágenes de los descriptores de textura bajo el esquema de segmentación markoviana. Las imágenes utilizadas en la segmentación son comprendidas de la fig. 8.17 hasta la fig. 8.21. El análisis para la función de energía *a priori* se realiza en 8-conexidad.

El resultado de la segmentación presenta solamente dos clases: {*petróleo*, *nopetróleo*}. Para mostrar los resultados de la segmentación sobre la imagen original, se delimitó la frontera entre las clases por medio de una línea blanca en la imagen original.

Los parámetros para el algoritmo de segmentación fueron:

$$\begin{aligned} T &= 2 \\ \beta &= 0.35 \\ \text{Iteraciones} &= 60 \end{aligned}$$

Sitio 1

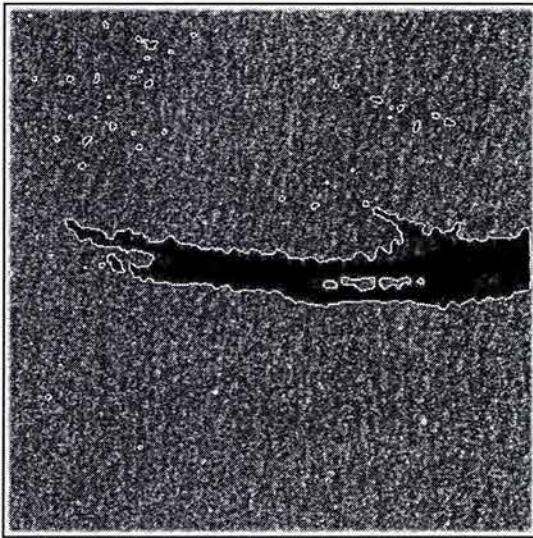


fig. 8.38

Sitio 2



fig. 8.39

La fig. 8.38 presenta “islas” falsas, aún así, esta imagen constituyen una buena aproximación comparada con respecto al resultado de las fig. 8.28 y 8.29. Un caso similar se presenta en la fig. 8.39 y las fig. 8.30 y 8.31.

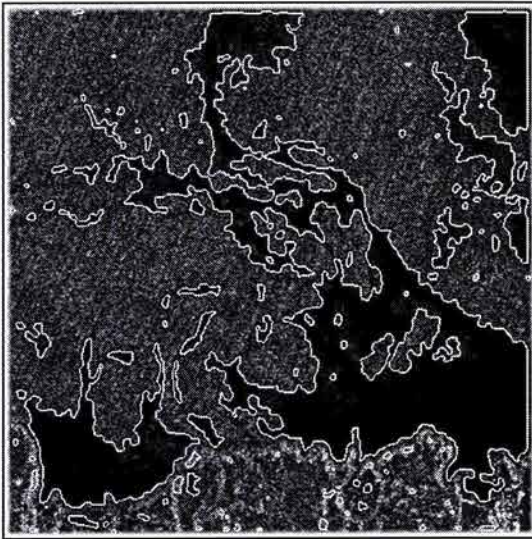
Sitio 3

fig. 8.40

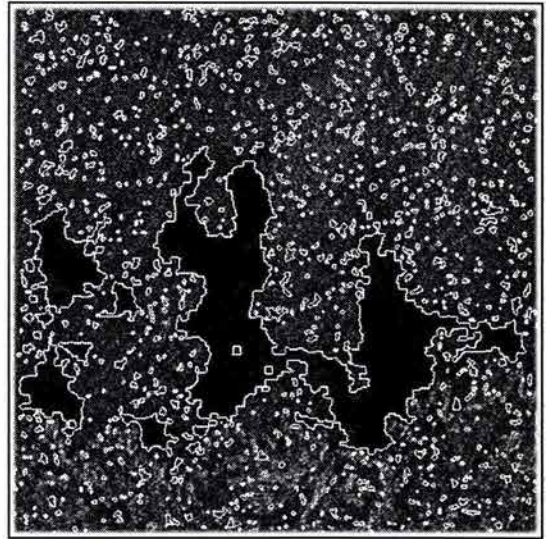
Sitio 4

fig. 8.41

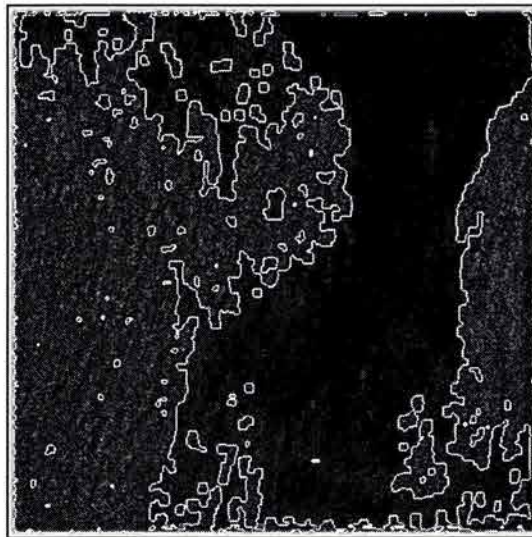
Sitio 5

fig. 8.42

La segmentación de los descriptores de textura bajo el análisis en *8-conexidad* tiene como resultado una descripción general de la mancha.

La calidad del resultado de la segmentación de la imagen del descriptor de textura determina la aparición de falsas “islas”, el ejemplo más claro se encuentra en el sitio 4 (fig. 8.41).

8.4 FUSIÓN DE DATOS

A lo largo de los resultados obtenidos en la segmentación markoviana, se puede observar que las imágenes contienen información complementaria y redundante entre sí. La fusión de datos incorporará esta información en una sola imagen.

Sitio 1

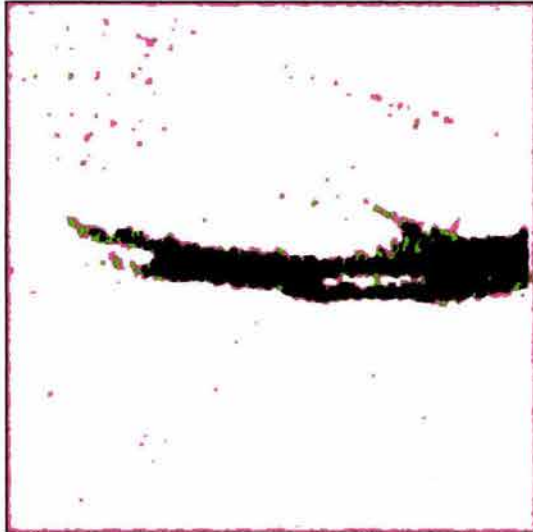


fig. 8.43

Sitio 2



fig. 8.44

Sitio 3

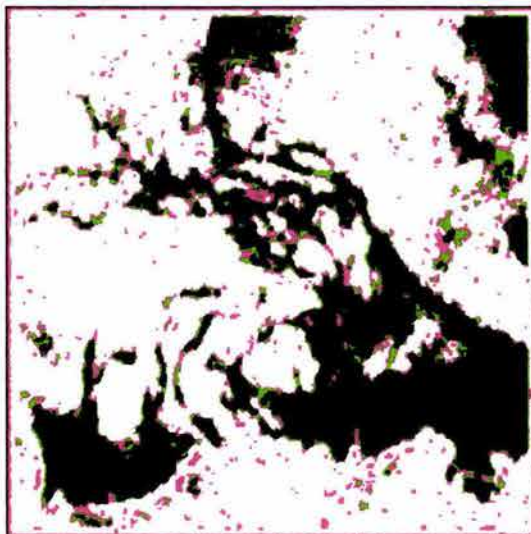


fig. 8.45

Sitio 4

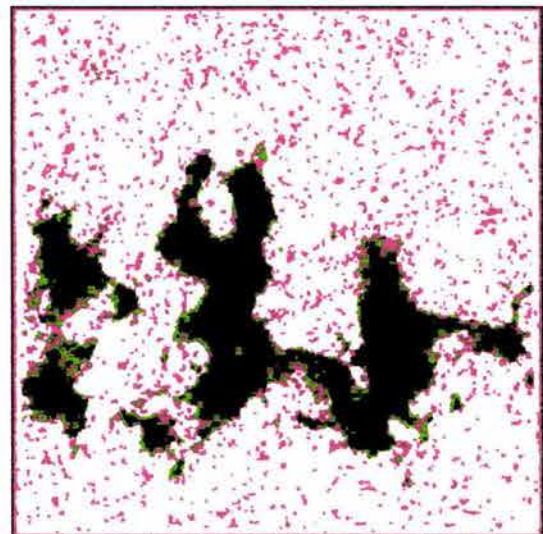


fig. 8.46

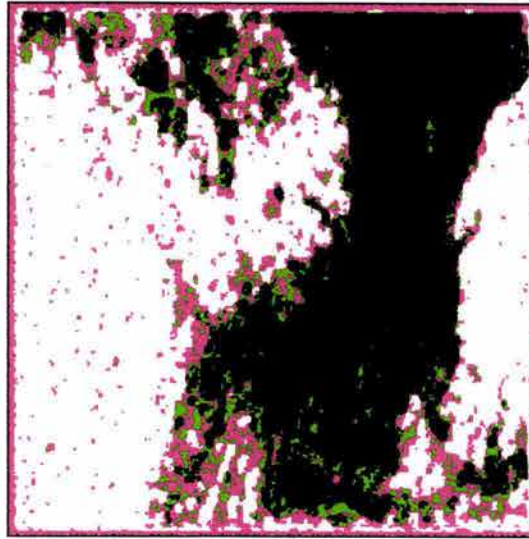
Sitio 5

fig. 8.47

En las figuras anteriores, la zona de color negro es la información redundante presente en las imágenes segmentadas. Esta información redundante comprende a los *sites* que tienen la etiqueta de petróleo en las tres imágenes segmentadas.

Las zonas de verde y rojo son la información complementaria. Las zonas verdes corresponden a los *sites* que tienen la etiqueta de petróleo en dos imágenes segmentadas. Las zonas rojas corresponden a los *sites* que tienen la etiqueta de petróleo en una imagen segmentada.

8.4.1 Fusión de Shan-Yu

El esquema de fusión Shan-Yu [SHAN97] evalúa la calidad de cada *site* con respecto a su vecindad asociada a la imagen a la que pertenece y a la vecindad asociada al mapa. La medida de error local relacionada a cada *site* determinará el grado de importancia para el proceso de fusión que tomará este pixel. Por otra parte, se debe proporcionar como punto de inicio un esbozo de la zona geográfica a procesar (mapa), muchas veces esta zona es burda e imprecisa, lo cual es un elemento importante a tomar en cuenta para considerar el error local de cada *site*.

Para cualquier $s \in S$, $Map(s)$ denota la etiqueta del *site* s en el mapa, siendo $Img(s)$ la etiqueta del *site* s en la imagen. Denotando $Map(s) = \lambda_i$ e $Img(s) = \lambda_j$, $\lambda \in L$, el cambio de etiqueta de λ_i (en el mapa) por λ_j (imagen) esta dada por:

$$x_{\lambda_i \lambda_j} = \begin{cases} 0 & \text{si } i = j \\ 20 & \text{si } i \neq j \end{cases}$$

Un *site* no puede ser analizado sin considerar los *sites* de su vecindad. Sea s' un sitio en la vecindad V_s del *site* s , y $d(s,s')$ la distancia euclidiana entre s y s' , se calcula las siguientes dos métricas:

$$a(s) = \sum_{s' \in V_s} \frac{\mathbf{1}_{\{Map(s')=\lambda_j\}}}{d(s,s')}$$

ec. 8.10

la ecuación anterior relaciona el *site* de la imagen con respecto a su vecindad asociada en el mapa.

$$b(s) = \sum_{s' \in V_s} \frac{\mathbf{1}_{\{Img(s')=\lambda_j\}}}{d(s,s')}$$

ec. 8.11

la ecuación anterior relaciona el *site* de la imagen con respecto a su vecindad asociada en la imagen.

Para los *sites* pertenecientes a una vecindad de primer orden, $d(s,s')=1$. Para los *sites* pertenecientes a una vecindad de segundo orden, $d(s,s')=\sqrt{2}$. La fusión tomó en cuenta una vecindad formada por 8 *sites*.

La medida de error local $E(s)$ del sitio s está definido como:

$$E(s) = \frac{x_{\lambda_i, \lambda_j}}{r \cdot a(s) + b(s) + 1}$$

ec. 8.12

donde $r = 2$. Posteriormente al cálculo del error local para cada *site*, se calcula un coeficiente de confianza para cada uno, basándose en el resultado del error local:

$$C(s) = 1 - \frac{E(s)}{\sum_{\lambda_i, \lambda_j \in L} x_{\lambda_i, \lambda_j}} \quad 0 \leq C(s) \leq 1$$

ec. 8.13

Sustituyendo los valores $x_{\lambda_1, \lambda_2} = 20$ y $x_{\lambda_2, \lambda_1} = 1$.

$$C(s) = 1 - \frac{E(s)}{20 + 1}$$

ec. 8.14

La decisión de la etiqueta final para cada *site* depende de la suma del coeficiente de confianza del mismo *site* a través de los resultados obtenidos:

$$\lambda(s) = \arg \max_{\lambda_i \in L} \sum_{k=1}^K C_k(s) \mathbf{1}_{\{Img_k(s) = \lambda_i\}}$$

ec. 8.15

donde $Img_k(s)$ es la k-ésima imagen entrada para el proceso de fusión y $C_k(s)$ es el coeficiente de confianza del sitio s en el resultado. Para este caso donde se consideren tres imágenes de entrada ($K=3$) para el proceso de fusión con dos clases: $\{\lambda_1, \lambda_2\} = \{\text{petróleo}, \text{nopetróleo}\}$, la ec. 8.15 quedaría como:

$$\lambda_1(s) = C_1(s)\delta(I_1(s) - \lambda_1) + C_2(s)\delta(I_2(s) - \lambda_1) + C_3(s)\delta(I_3(s) - \lambda_1)$$

ec. 8.16

$$\lambda_2(s) = C_1(s)\delta(I_1(s) - \lambda_2) + C_2(s)\delta(I_2(s) - \lambda_2) + C_3(s)\delta(I_3(s) - \lambda_2)$$

ec. 8.17

y la etiqueta correspondiente al *site* s que se analiza:

$$\lambda(s) = \max\{\lambda_1(s), \lambda_2(s)\}$$

ec. 8.18

Para integrar las imágenes procedentes de la segmentación por Campos Aleatorios de Markov, es necesario proponer los mapas que se utiliza para el cálculo del coeficiente $a(s)$ y el coeficiente C .

Estos mapas se muestran en la siguiente página.

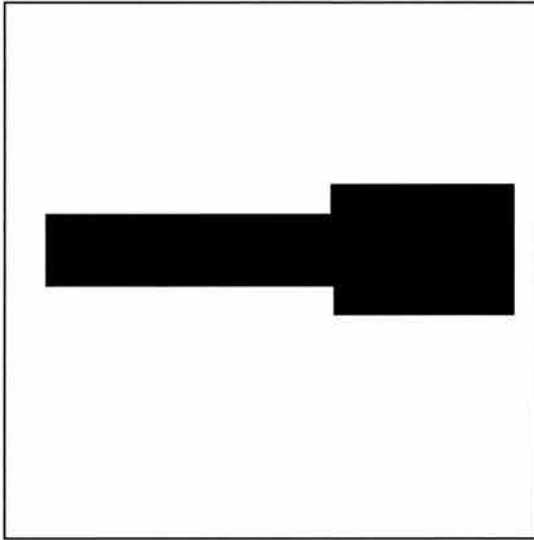
Mapa 1

fig. 8.48

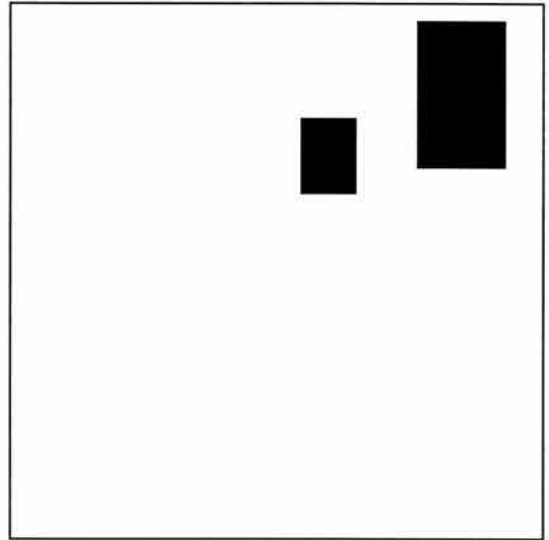
Mapa 2

fig. 8.49

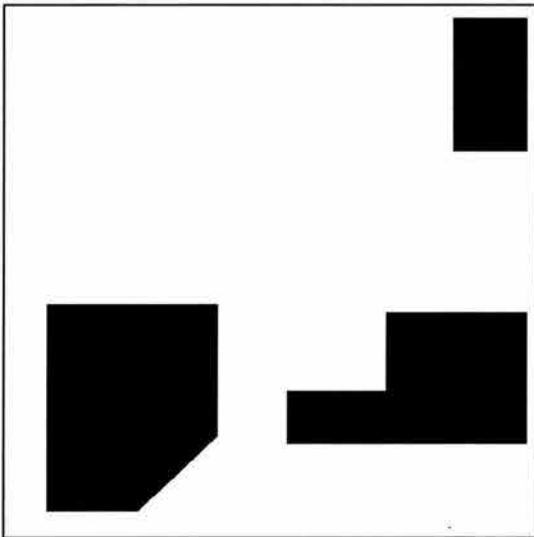
Mapa 3

fig. 8.50

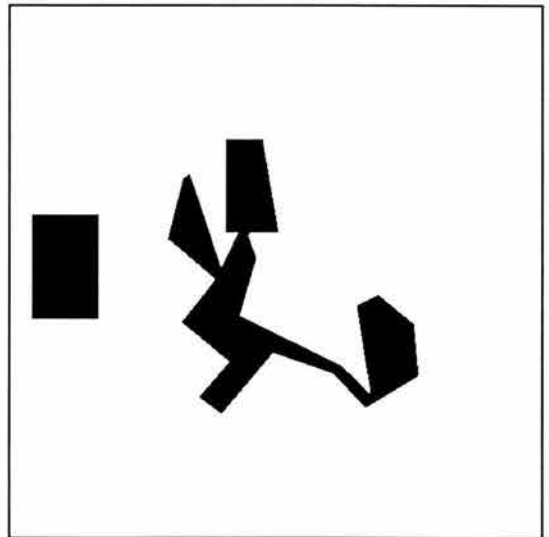
Mapa 4

fig. 8.51

En principio, el esquema de fusión Shan-Yu fue pensado en que existía información cartográfica disponible de la zona que se estudia. Ante la carencia de ésta información, la forma del mapa se propone como una aproximación burda de la zona de interés.

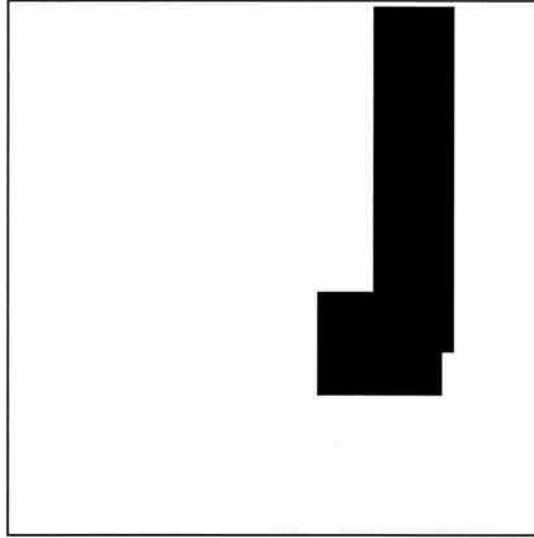
Mapa 5

fig. 8.52

El tercer paso del esquema consiste en la fusión de las tres imágenes segmentadas por Campos Aleatorios de Markov y el mapa propuesto. Los resultados de la fusión bajo el esquema de Shan-Yu se muestran a continuación:

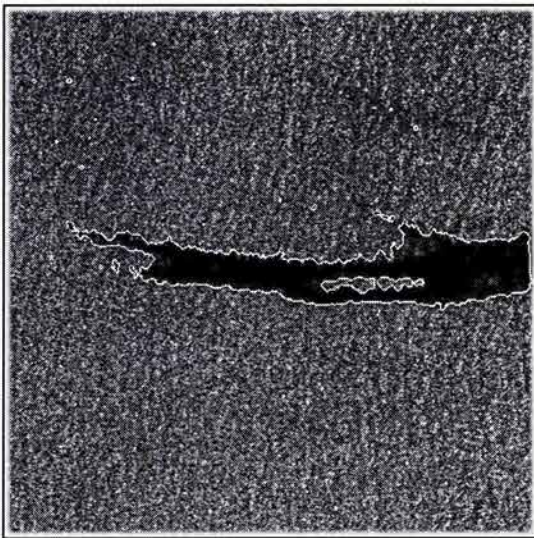
Sitio 1

fig. 8.53

Sitio 2

fig. 8.54

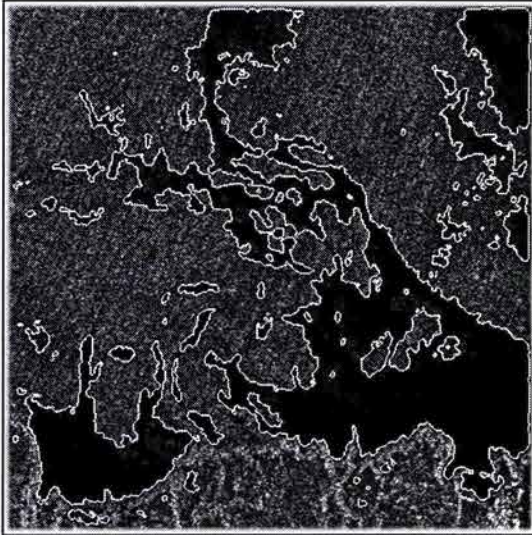
Sitio 3

fig. 8.55

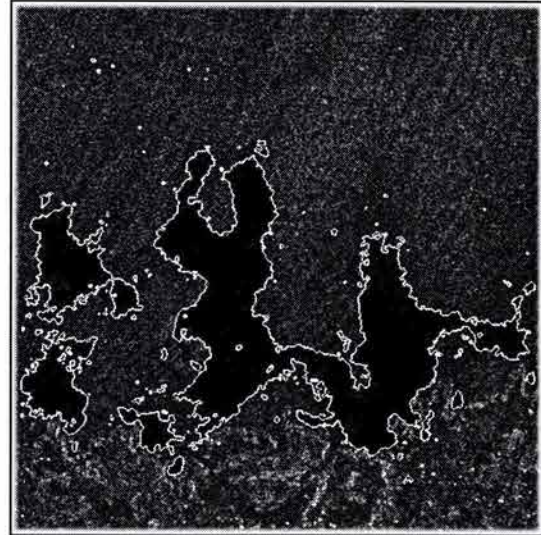
Sitio 4

fig. 8.56

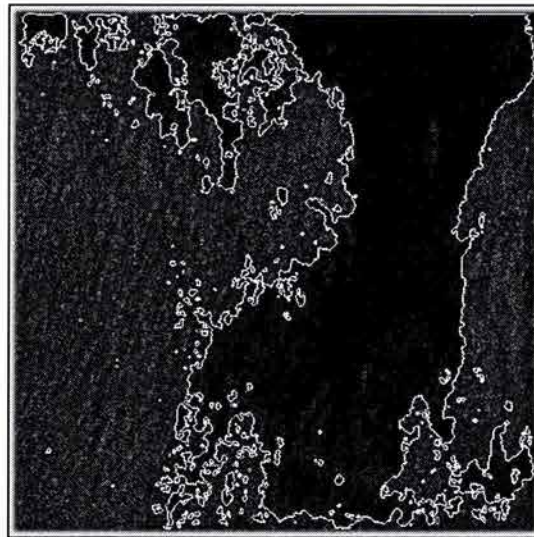
Sitio 5

fig. 8.57

Como se puede observar, el inconveniente en el esquema de fusión Shan-Yu radica en la aparición de “manchas” falsas. La razón consiste en que, a pesar de tomar en cuenta la vecindad de los pixels para determinar la clase a la que pertenece, la posición de la información complementaria no se considera con respecto a la información redundante.

8.4.2 Propuesta de Fusión

En esta propuesta de fusión las imágenes de entrada son binarias: $\{\text{petróleo}, \text{nopetróleo}\}=\{1,0\}$. La consideración principal que se realiza en la imagen suma de las tres imágenes de entrada, consiste en que la imagen suma presenta un nivel de información redundante y dos niveles de información complementaria (fig. 8.43 a 8.47). Esta consideración también sería válida si se tuvieran más imágenes de entrada. El procedimiento de fusión se resume a continuación:

1.- Se suman las imágenes:

$$I_T = I_1 + I_2 + I_3$$

2.- Primer nivel de redundancia

Se barre secuencialmente la imagen

Si $I_T(s)=3$, entonces el *site* s se etiqueta como petróleo

3.- Primer nivel complementario

Se barre secuencialmente la imagen

Si $I_T(s)=2$ se realiza el análisis en 4-*conexidad*

Si dos o más *sites* vecinos a s tienen etiqueta de *petróleo*, entonces el *site* s se etiqueta como *petróleo*.

Se barre secuencialmente la imagen hasta que no haya más cambios de etiquetas.

4.- Segundo nivel complementario

Se barre secuencialmente la imagen

Si $I_T(s)=1$ se realiza el análisis en 4-*conexidad*

Si dos o más *sites* vecinos a $I_T(s)=1$ tienen etiqueta de *petróleo*, entonces el *site* s se etiqueta como petróleo.

Todos los *sites* que no cambiaron durante todo el proceso, se etiquetan como *nopetróleo*.

El esquema de fusión propuesto es más simple que el esquema de fusión de Shan-Yu. La desventaja de este algoritmo es que no toma en cuenta el primer y segundo nivel complementario a menos que existan *sites* pertenecientes al primer nivel de redundancia. Esto provoca la desaparición de estructuras finas de las manchas. Por otra parte, mejora la homogeneidad de las imágenes, eliminando “manchas” falsas.

Las imágenes resultado del esquema propuesto son las siguientes:

Sitio 1

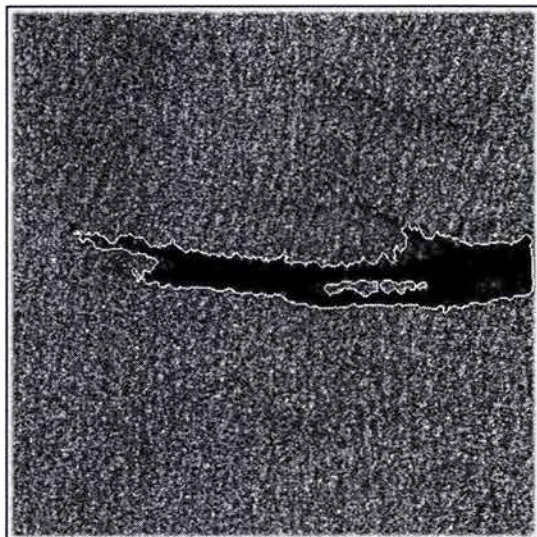


fig. 8.58

Sitio 2



fig. 8.59

Sitio 3



fig. 8.60

Sitio 4

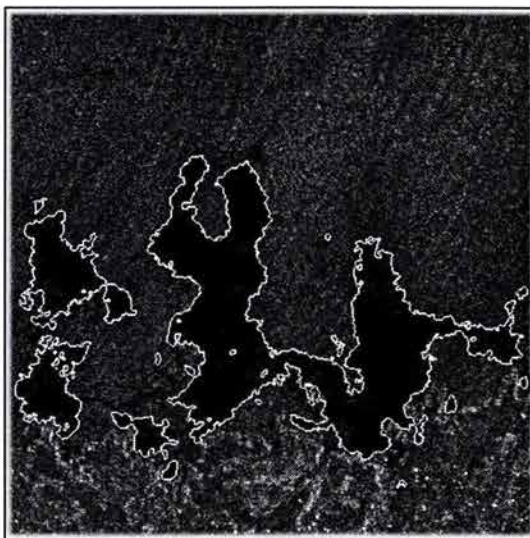


fig. 8.61

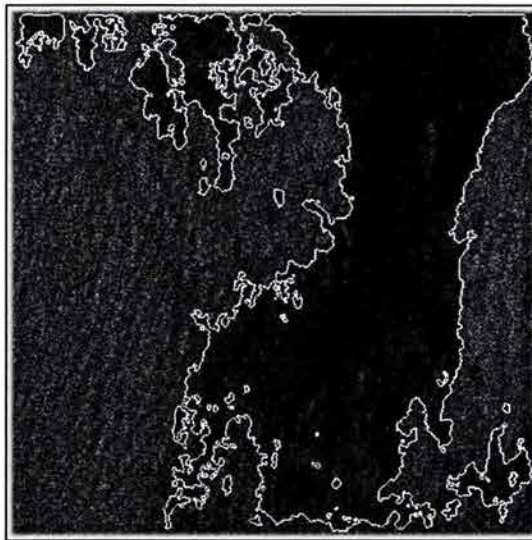
Sitio 5

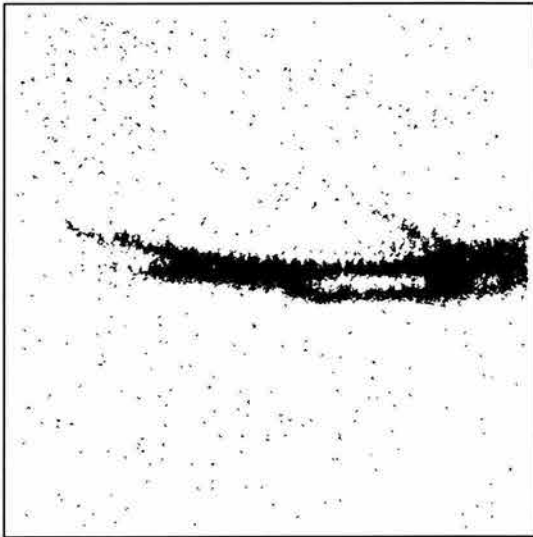
fig. 8.62

Como se puede observar, el resultado de este esquema de segmentación propuesto es comparable con el esquema de fusión de Shan-Yu y mejora la homogeneidad de la imagen.

8.5 EVALUACIÓN DE RESULTADOS

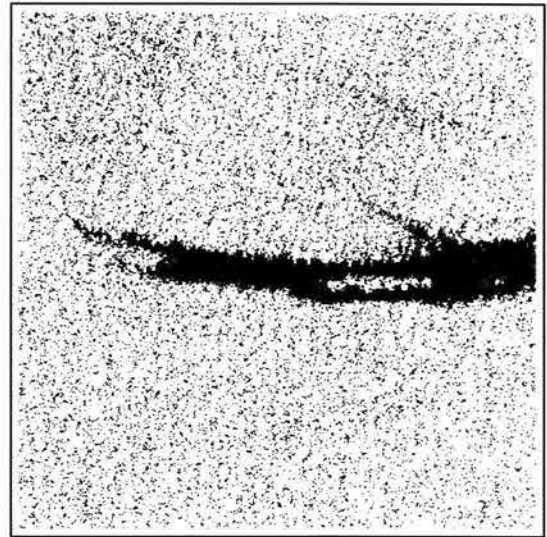
La evaluación de los resultados se realiza de manera visual. El esquema markoviano propuesto se compara con dos esquemas clásicos de segmentación como lo son: segmentación K-medias y segmentación de Bayes. A continuación se presentan los resultados:

Sitio 1



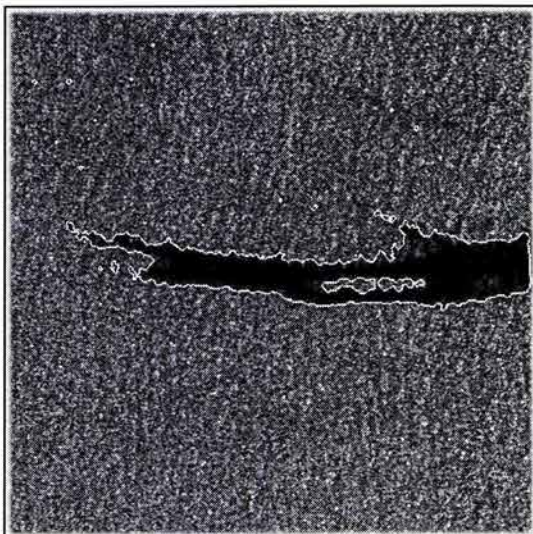
Bayes

fig. 8.63



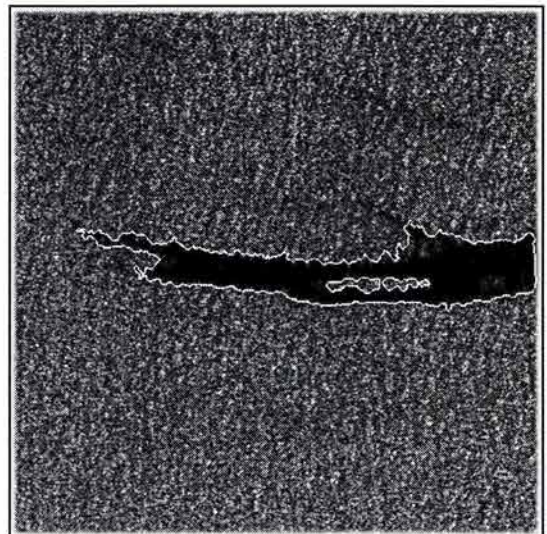
K-medias

fig. 8.64



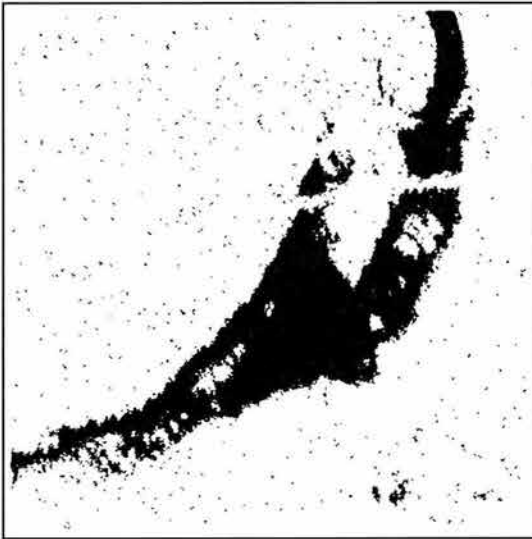
Esquema Markoviano[Shan-Yu]

fig. 8.65



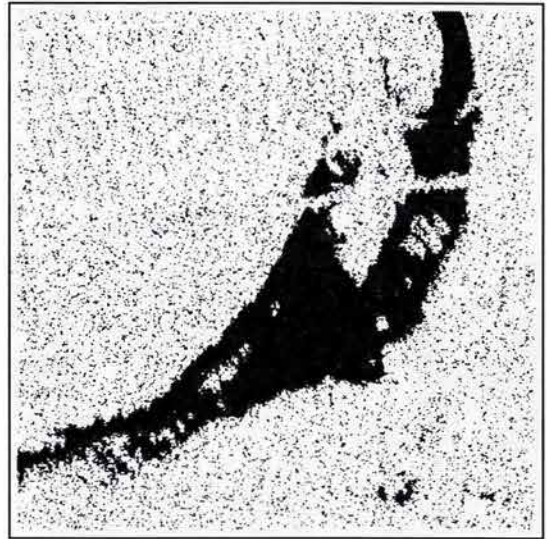
Esquema Markoviano[Propuesta]

fig. 8.66

Sitio 2

Bayes

fig. 8.67



K-medias

fig. 8.68



Esquema Markoviano[Shan-Yu]

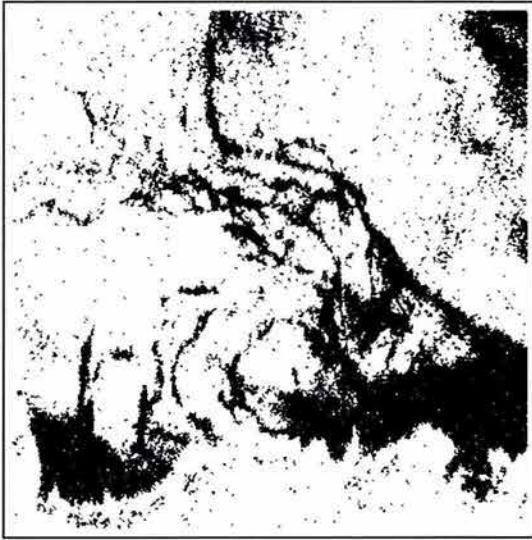
fig. 8.69



Esquema Markoviano[Propuesta]

fig. 8.70

Los resultados del esquema markoviano propuesto, muestran el mejor desempeño de éste con respecto a los resultados de la segmentación de Bayes y K-medias.

Sitio 3

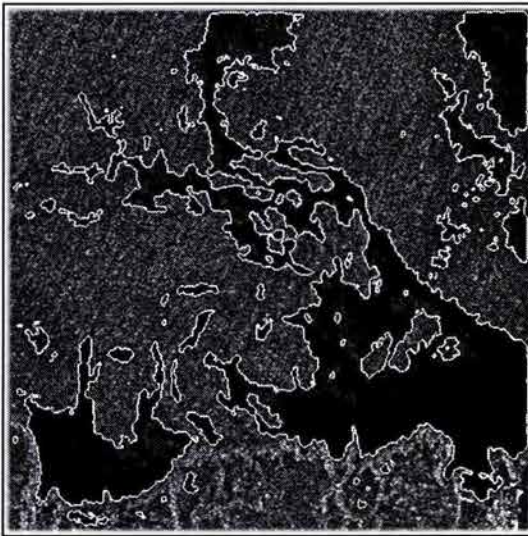
Bayes

fig. 8.71



K-medias

fig. 8.72



Esquema Markoviano[Shan-Yu]

fig. 8.73

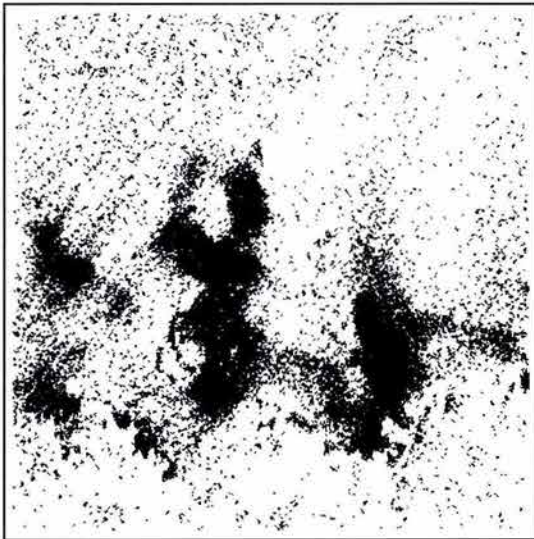


Esquema Markoviano[Propuesta]

fig. 8.74

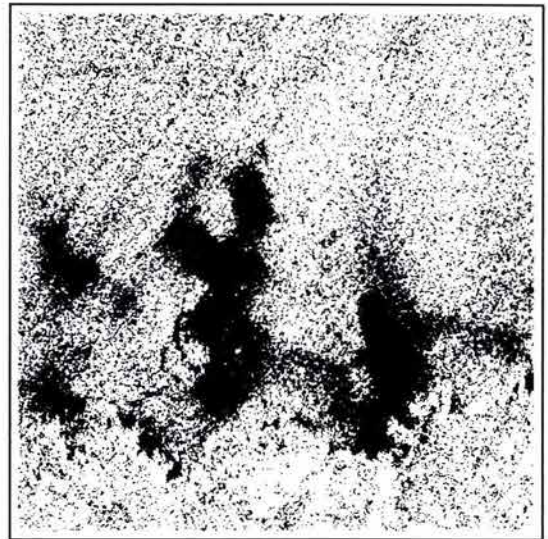
Los algoritmos de segmentación K-medias y Bayes no son contextuales, siendo la calidad de segmentación menor en comparación al esquema markoviano.

Sitio 4



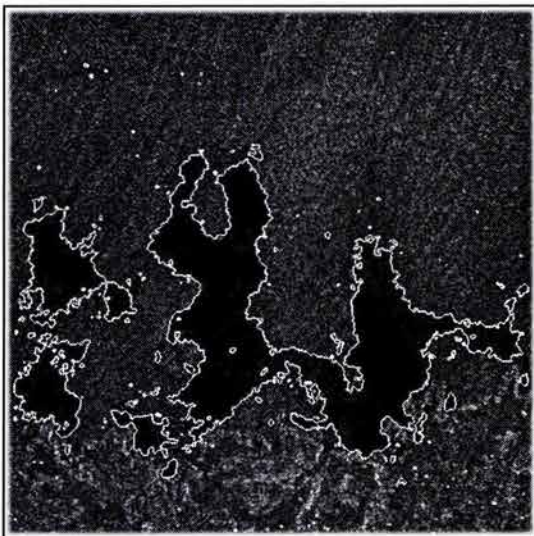
Bayes

fig. 8.75



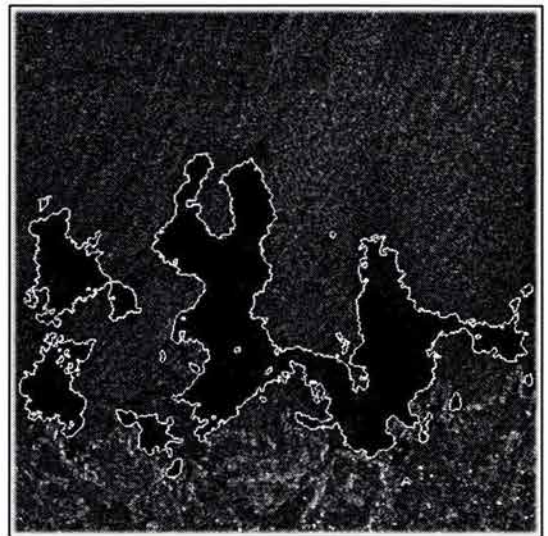
K-medias

fig. 8.76



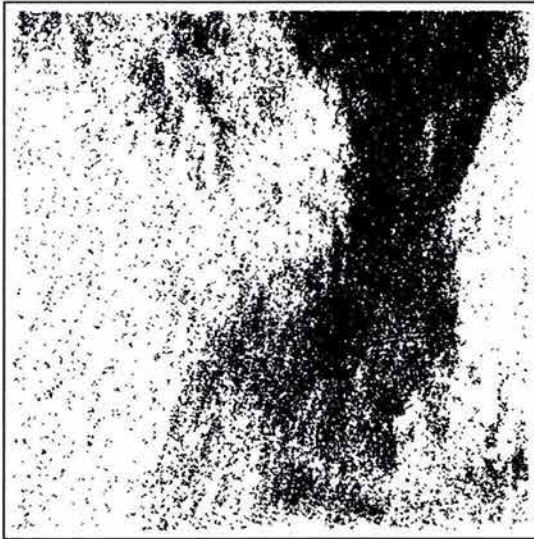
Esquema Markoviano[Shan-Yu]

fig. 8.77



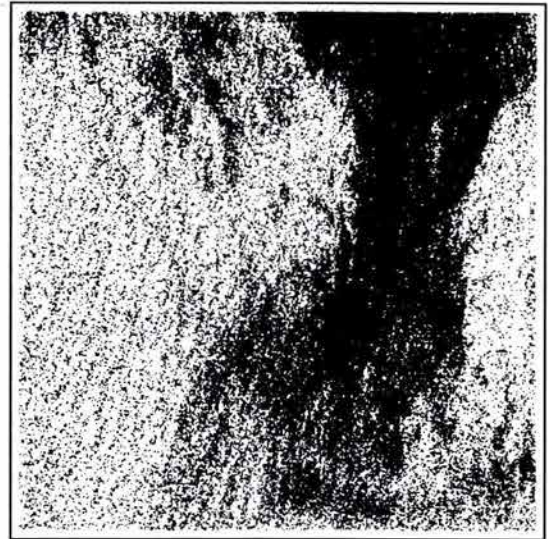
Esquema Markoviano[Propuesta]

fig. 8.78

Sitio 5

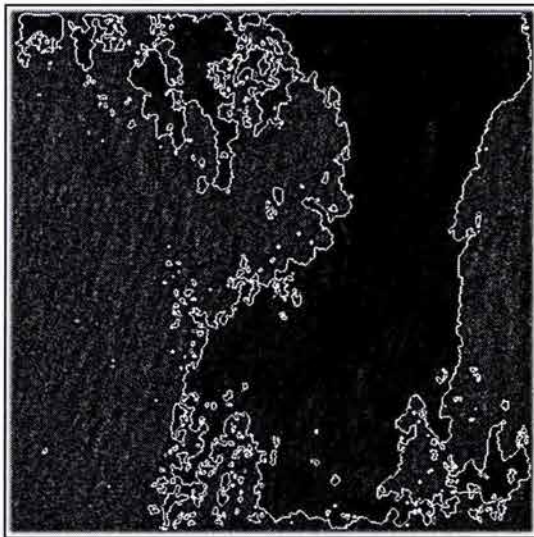
Bayes

fig. 8.79



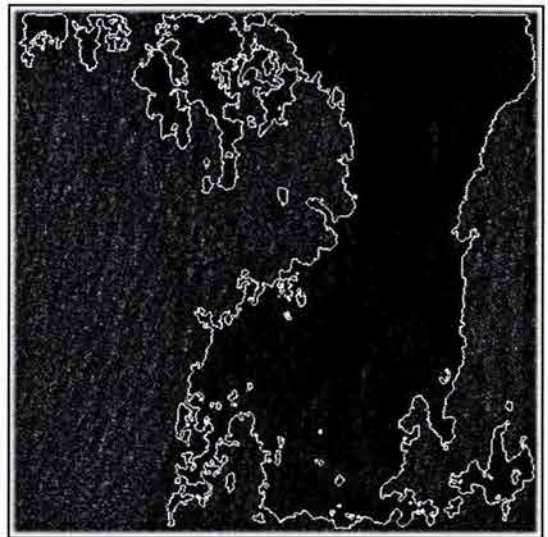
K-medias

fig. 8.80



Esquema Markoviano[Shan-Yu]

fig. 8.81



Esquema Markoviano[Propuesta]

fig. 8.82

Al disminuir el contraste de la imagen y por ende la separación entre clases, el resultado de los algoritmos de segmentación por Bayes y K-medias pierden calidad, caso contrario se presenta en el esquema markoviano.

El algoritmo de fusión tiende a disminuir el efecto de pixels y zonas aisladas de las imágenes de entrada, reduciendo las manchas ficticias.

9

CONCLUSIONES

El estudio de las imágenes de percepción remota se ha constituido como una disciplina de importancia. El análisis de estas imágenes brinda información necesaria para la interpretación y el entendimiento de los fenómenos que se presentan en la superficie terrestre.

La ventaja que presentan las imágenes adquiridas por medio del Radar de Apertura Sintética es la capacidad de adquisición bajo diferentes condiciones meteorológicas y a cualquier hora del día. Esto lo convierte en un sistema para todo tiempo. En contraparte se encuentra la naturaleza ruidosa de estas imágenes que visualmente se puede interpretar como una característica de textura.

El esquema de procesamiento para la detección remota de derrames petroleros se dividió en tres etapas:

1. La primera etapa fue el análisis de la textura de la imagen SAR. La textura era una característica de las diferentes clases que constituían una escena y por lo tanto permitía la discriminación de una clase con respecto de otra. Ante esta situación, se propuso el análisis de textura con los descriptores texturales asociados a la Matriz de Co-ocurrencia de Niveles de Gris (GLCM).

La textura se consideró como un arreglo bidimensional en una región determinada, sobre la cual se calcularon los descriptores de textura, con el fin de discriminar entre clases. Esto se basa en que la textura es esencialmente una interacción estadística de pixels, un campo aleatorio homogéneo bidimensional, que puede ser caracterizado estadísticamente.

Para realizar la selección de parámetros de la GLCM se propuso el uso de una función estadística basada en la distancia de Mahalanobis. Con esta función se calculó la distancia entre las clases de interés una vez aplicado el descriptor de textura.

De forma práctica para la elección de los parámetros se estudiaron dos opciones; primero: con una distancia interpíxel fija y modificando el tamaño de la ventana sobre la cual se calcula la GLCM, y segundo: modificando la distancia interpíxel con un tamaño de ventana fijo. Este procedimiento, por una parte, limitó la elección de los parámetros a los picos máximos presentes en las gráficas y por otra, a la calidad visual del resultado.

2. La segunda etapa del esquema propuesto fue la segmentación de la imagen SAR y la segmentación de la imagen correspondiente al descriptor de textura.

La imagen fue analizada como un Campo Aleatorio de Markov. Se consideró que la probabilidad condicional de que un *site* (pixel) determinado tuviera un nivel de gris particular, dado los valores de todos los *sites* restantes en la imagen, era idéntica a la probabilidad condicional dados los valores del campo en un conjunto pequeño de *sites*, es decir, su *vecindad*. Esta propiedad local también recibe el nombre de Markovianidad.

En el presente esquema de procesamiento, la pertenencia a una determinada clase fue dada por la probabilidad Máxima A Posteriori (MAP). La aproximación MAP contiene tres términos, el modelo de probabilidad $P(X|Y)$, la probabilidad *a priori* $P(Y)$ y la probabilidad *a posteriori* $P(Y|X)$, así como las funciones de energía respectivas.

El modelo de probabilidad $P(X|Y)$ tuvo asociado el modelo de energía $U(X|Y)$. Para definir el modelo de energía $U(X|Y)$ se utilizaron ventanas de entrenamiento y así obtener una aproximación de la media correspondiente a cada clase.

Al modelo de probabilidad *a priori* $P(Y)$ se le asocia la función de energía *a priori* $U(Y)$. La teoría de Campos Aleatorios de Markov proporcionó las herramientas para codificar las restricciones contextuales en el modelo de energía *a priori*. Para tal fin, solamente se analizaron *cliques* formados por dos *sites*. Se consideró que el Campo Aleatorio de Markov era homogéneo e isotrópico, por lo que el parámetro β asociado a los *cliques* de segundo orden era independiente de la posición y orientación del *clique*.

El análisis de la vecindad asociada al pixel estudiado se restringió a dos casos: *8-conexidad* y *4-conexidad*. El análisis en *8-conexidad* es el más utilizado en la segmentación de imágenes, este análisis presentó el inconveniente de no conservar estructuras finas, proporcionando una descripción general de la mancha de petróleo. Por contraparte, el análisis en *4-conexidad* proporcionaba detalles y estructuras finas de la mancha de petróleo, pero introducía manchas falsas.

La energía *a posteriori* se formuló como la suma de la función de energía *a priori* y la función de energía conjunta. Finalmente, para encontrar la configuración óptima del problema (imagen segmentada), fue necesario minimizar la energía *a posteriori*, para tal fin, se utilizó el algoritmo de *Recocido Simulado*, el cual es un algoritmo que permite encontrar el óptimo global del problema.

Como se observó, el empleo del esquema de segmentación por la teoría de Campos Aleatorios de Markov permite introducir información contextual útil al proceso para el análisis de datos ruidosos, lo que conduce a tener resultados más homogéneos.

Ante los resultados obtenidos por la segmentación bajo diferentes variantes del análisis de *conexidad*, se pudo observar que presentaban información redundante y complementaria entre sí, por lo que era necesario fusionar.

3. El tercer paso del esquema de procesamiento fue la fusión de los resultados obtenidos anteriormente.

Uno de los esquemas de fusión empleado es generalmente usado en el análisis de zonas donde se tiene un conocimiento previo en mapas o descripciones cartográficas. Ante esto y debido a la naturaleza impredecible de la forma de petróleo, se elaboró una aproximación burda de la zona. La desventaja de este algoritmo consiste en que proporciona falsas manchas, ya que no se considera la posición de la información complementaria con respecto a la información redundante.

La contribución técnica del segundo algoritmo de fusión consistió en la robustez para el análisis contextual de imágenes que provenía de diferentes fuentes y que las llevaba a un resultado mas completo, es decir, permitía obtener zonas homogéneas libres de falsas manchas y preservaba en su mayor parte las estructuras de la mancha de petróleo.

10

ANEXO

El *Prestige*, fue un buque petrolero de una compañía nigeriana con bandera por "conveniencia" de Bahamas, que transportaba 67 mil toneladas de combustible y se partió en dos la madrugada del martes 19 de noviembre de 2002 en las inmediaciones de las costas gallegas, en la llamada Costa da Morte (fig. 10.1 y 10.2), una zona que había sufrido en los años anteriores al menos tres grandes desastres ecológicos. No fue suficiente la operación especial de las Fuerzas de Salvación Marítima españolas ni el auxilio europeo para evitar que el barco, con medidas de seguridad obsoletas y más de 26 años de antigüedad, sucumbiera ante el temporal que había azotado en los días anteriores a la región atlántica española.

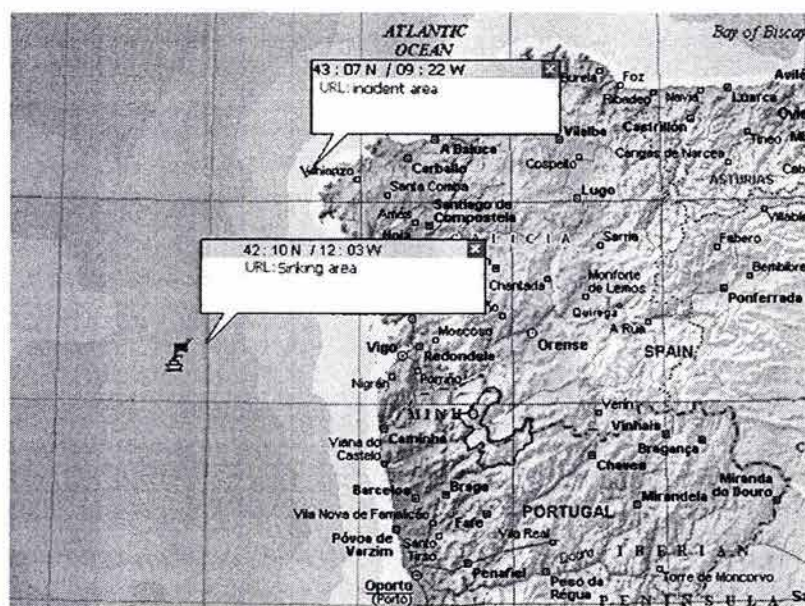


fig. 10.1

Lo paradójico de este desastre ecológico, comparable si acaso con el que sufrió el Reino Unido en 1997, cuando se derramaron miles de toneladas de petróleo en sus costas, es que el fallo -o posible fallo- estaba detectado: el *Prestige*, un buque petrolero construido en Japón 26 años antes, no contaba con el "doble casco" de protección que se exige a los barcos que transportan hidrocarburos en algunos países, entre ellos todos los que forman la Unión Europea.

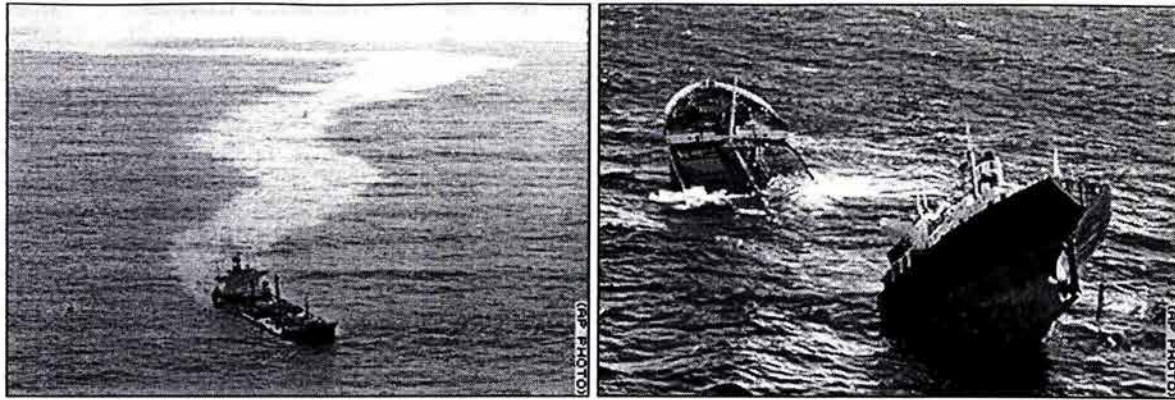


fig. 10.2

En el caso de esta nave, siete días antes había sufrido una vía de agua a causa del temporal frente a las costas gallegas, la tripulación, salvo el capitán y tres marineros más, abandonaron sin más el buque; el barco estaba viejo y sobrecapado. La cantidad de combustible transportado era un riesgo latente, que finalmente se convirtió en tragedia después de casi una semana.

El suceso, después del hundimiento del buque a 250 kilómetros del Atlántico gallego, fue una tragedia peor que el derrame del petrolero estadounidense *Exxon Valdez*, que en 1989 chocó contra un arrecife y derramó unas 42 mil toneladas de crudo en la zona pesquera más rica de Alaska, el estrecho Prince William y provocó la muerte de unas 250 mil aves marinas.

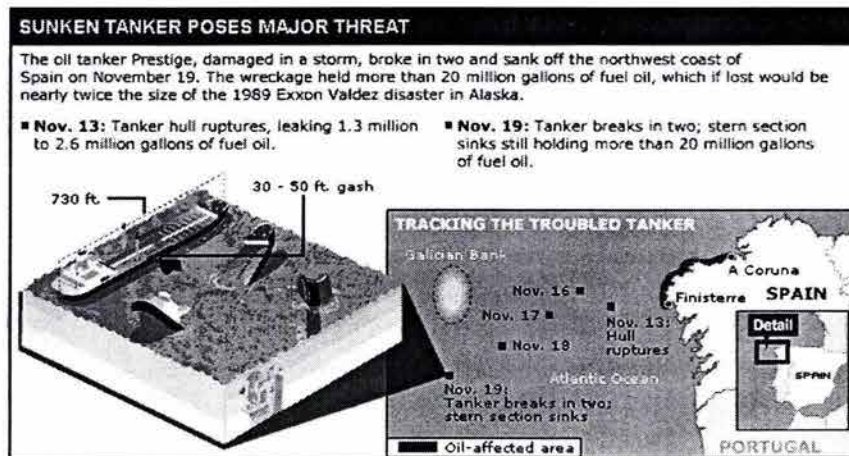


fig. 10.3

Los expertos españoles confiaban en que el crudo solidificara por las bajas temperaturas que se registran a 4 mil metros de profundidad, aún así, el desastre se convirtió en tragedia para las miles de familias de pescadores después de que, en plena madrugada, una inmensa marea negra irrumpiera de lleno en cerca de 300 kilómetros de las costas de Galicia (fig. 10.3).

Durante los primeros cuatro días en que el buque *Prestige* anduvo a la deriva, solamente se derramaron 7 mil de las 67 mil toneladas que transportaba la nave. El gobierno español, estimó que la catástrofe ecológica superaba los 40 millones de dólares en daños.

Las consecuencias de esta tragedia, que en décadas el ecosistema subsanará, son incalculables, sobre todo porque resulta muy difícil saber qué incidencias provocará en este paraje natural, uno de los más bellos y agrestes de España. Una gran extensión de sus costas servía de morada a diversas especies: berberechos, almejas, mejillones, angulas, etcétera. Aves y peces sucumbieron con las plumas y escamas ennegrecidas por la contaminación causada por el crudo. Algunos efectos se ilustran en la fig. 10.4.

Pese al panorama, más bien adverso, las asociaciones de pescadores gallegos salieron a trabajar como cualquier otro día. Para los que tuvieron que ir mar adentro la faena resultó trágica, ya que la mayoría de los peces capturados estaban impregnados con el petróleo del buque, lo que significa que el "producto" no supera las medidas de control sanitario.



fig. 10.4

REFERENCIAS

- [BOW02]Bow Sing-Tze, *Pattern Recognition and Image Preprocessing*, 2da. edición 2002, Ed. Decker, New York.
- [CHUV90]Chuvieco E., *Fundamentos de Teledetección Espacial*, 1era. edición 1990, Ed. RIALP.
- [CLAU01]Clausi D.A., Zhao Y., *Rapid determination of Co-occurrence Texture Features*, pág. 1880-1882, IEEE IGARSS 2001.
- [DOWN01]Downsland K., Adenso-Díaz B., *Diseño de Heurísticas y Fundamentos de Recocido Simulado*, Revista Iberoamericana de Inteligencia Artificial, No. 20, pág. 34-52, 2001.
- [GONZ73]Gonzalez R. C., Woods R. E., *Digital Image Processing*, Ed. Edisson-Wesley, pág. 610-621, Nov. 1973.
- [HALL97]Hall D., Llinas J., *An Introduction to Multisensor Data Fusion*, Proceedings of the IEEE, Enero 1997, Vol. 85, No. 1.
- [HARA73]Haralick R.M., Shanmugam K., Dinstein I., *Texture Features for Image Classification*, IEEE Trans. on Syst. Man Cybern., vol. 3, pág. 610-621, Nov. 1973.
- [LI01]Li Z. S., *Markov Random Field Modeling in Image Analysis*, Ed. Springer Verlag, 2001.
- [LI97]Li Z. S., *Energy Minimization and Relaxation Labeling*, Journal of Mathematical Imaging and Vision, No.7, 1997.
- [LILL97]Lillesand T. M., *Remote Sensing and Image Interpretation*, 2da. edición, Ed. Jhon Wiley & Sons, 1997.
- [MARA94]Maravall Gómez-Allende, *Reconocimiento de Formas y Visión Artificial*, Ed. Addison-Wesley Iberoamericana, 1era. edición, 1994.
- [MOCT00]Moctezuma M., *Apuntes de Análisis de Señales Aleatorias*, UNAM-FI, 2000.
- [MORA03]Morales D., Moctezuma M., Parmiggiani F., *Fast Texture Analysis Algorithm and Applications for Image Segmentation*, CIARP 2003.
- [PARK96]Parker J. R., *Algorithms for Image Processing and Computer Vision*, 1era. edición, Ed. Wiley and Sons Inc., pág. 151-175. 1996.

-
- [PIEL02]Piella G., Heijmans H., *Multiresolution Image Fusion Guided by a Multimodal Segmentation*, Proceedings of ACIVS 2002, Ghent, Belgica, 2002.
- [PLAT02]Platonov A., *Aplicación de Imágenes de Satélite SAR en los Estudios de Contaminación Marina y de Dinámica de las Aguas en el Mediterráneo Noroccidental*, Ed. Universidad Politécnica de Cataluña, 2002.
- [PUJO00]Pujol M., Rizo P., *Aplicación de los Modelos de Campos Aleatorios de Markov en Visión Artificial*, Revista Electrónica de Visión, Departamento de Ciencia de la Computación e Inteligencia Artificial, Número 4, Universidad de Alicante, Octubre 2000.
- [ROZA01]Rozalija S., *Computing Textural Features Based on Co-occurrence Matrix for Infrared Images*, Image and Signal Processing and Analysis 2001.
- [SABI96]Sabins F., *Remote Sensing: Principles an Interpretation*, 3era. edición, Ed. Freeman. 1996.
- [SHAN97]Shan-Yu., Berthod M., Giraundon G., *Towards Robust Analysis of Satellite Images Using Map Information- Application to Urban Area Detection*, Rapport de Recherche, INRIA, Nov. 1997.
- [SHAN95]Shan-Yu.: *Improving Satellite Image Analysis Quality by Data Fusion*. IEEE IGARSS pág. 2164-2166. 1995.
- [SOH99]Soh L. K., Tsatsoulis T., *Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-occurrence Matrices*, IEEE Trans. on Geoscience Remote Sensing, vol. 37, pag. 780-795, Mar. 1999.
- [TAPI01]Tapiador F.J., Casanova J.L., *Fusión de Datos en Teledetección*, Revista de la Asociación Española de Teledetección 2001.
- [UNSE86]Unser M., *Sum and Difference Histograms for Texture Classification*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, pag. 118-125. 1986.
- [ZHAN94]Zhang D.B., Van Gool, Oosterlinck A., *Coastline Detection From SAR Images*, technical report, Departement Elektrotechniek, Katholieke Universiteit Leuven, 1994.

<http://www.bmp.com.pe/newgeo/productos/imagenes/opticas/altar/ikonos/ikonos.htm>

<http://www.bmp.com.pe/newgeo/productos/imagenes/opticas/bajar/aster/aster.htm>

<http://samadhi.jpl.nasa.gov/msl/QuickLook/>

<http://envisat.esa.int>

<http://www.agriculturadeprecision.org/percrem/AgosatChile.htm>

<http://www.iti.upv.es/~evidal/students/app/tema4/t4app2p.pdf>

<http://chue.ing.ula.ve/DOCENCIA/POSTGRADO/CURSOS/MSS05/SA-mio.html>