



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRIA Y DOCTORADO EN
INGENIERIA

METODO CONTEXTUAL PARA FUSION DE DATOS

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN INGENIERIA

INGENIERIA ELECTRICA - TELECOMUNICACIONES

P R E S E N T A :

ALEJANDRO OSCAR CAMARA SANCHEZ

TUTOR:

DR. MIGUEL MOCTEZUMA FLORES

2005



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

JURADO ASIGNADO:

Presidente: **Dra. Lucia Medina Gomez**

Secretario: **Dr. Jorge Lira Chavez**

Vocal: **Dr. Miguel Moctezuma Flores**

1^{er}. Suplente: **Dr. Pablo Perez Alcazar**

2^{do}. Suplente: **Dra. Ana Maria Vazquez Vargas**

Lugar o lugares donde se realizó la tesis:

México D.F. Ciudad Universitaria

TUTOR DE TESIS:

Dr. Miguel Moctezuma Flores

INDICE

Capítulo I. Introducción.

1.1 Objetivo.....	4
1.2 Antecedentes del tema.....	4
1.3 Definición del problema.....	6
1.4 Metodología empleada.....	7

Capítulo II. Fundamentos de la percepción remota

2.1 Descripción general.....	10
2.2 Sistemas de percepción remota.....	11
2.2.1 Percepción remota aérea.....	11
2.2.2 Percepción remota espacial.....	13
2.3 Aplicaciones de los sistemas de percepción remota.....	14
2.3.1 Aplicaciones utilizando imágenes de la superficie de la tierra.....	17
2.3.2 Cartas de temperatura superficial del mar (TSM).....	17
2.3.3 Cartas promedio de TSM y ATSM.....	18
2.3.4 Monitoreo de laguna "La Niña".....	18
2.3.5 Localización de incendios forestales.....	19
2.3.5 Imágenes satelitales provenientes del satélite LANDSAT (resolución media).....	20
2.4 Imágenes multiespectrales.....	21
2.5 Imágenes de radar de apertura sintética SAR.....	24
2.5.1 Detección de derrames de hidrocarburos por medio del SAR.....	26
2.5.2 Detección de inundaciones por medio del SAR	28

Capítulo III. Principales satélites de percepción remota

3.1 ENVISAT.....	29
3.1.1 Características del sensor	32
3.2 Satélites ERS 1 y 2.....	34
3.2.1 Plataforma	36
3.2.2 Adquisición de imágenes de radar	37
3.2.3 Procesamiento de las imágenes	37
3.3 Satelite LANDSAT.....	39
3.3.1 LANDSAT 1,2 y 3.....	40
3.3.2 Sistema RBV (Return Beam Vidicom).....	41
3.3.3 Sistema MSS (Multispectral Scanner Subsystem - Generador de Imágenes Multiespectrales).....	42
3.3.4 LANDSAT 7.....	42
3.4 Satelite SPOT.....	44

Indice

3.4.1 Tres satélites en órbita.....	46
3.4.2 Capacidad de almacenamiento.....	46
3.4.3 La cámara HRG	47
3.4.4 La cámara HRS	47
3.4.5 Memoria electrónica	48
3.4.6 Transmisión a tierra	48
3.4.7 Las bandas espectrales	49
3.5 Satélite IKONOS.....	50
3.5.1 Características del satélite IKONOS.....	51
3.5.2 Cobertura.....	52
3.5.3 Aplicaciones y ventajas de las imágenes IKONOS.....	52
3.6 Satélite TERRA/ASTER.....	53

Capítulo IV. Análisis textural

4.1 Alternativas en el análisis textural.....	56
4.2 Matrices de co-ocurrencia en niveles de gris (GLCM).....	57
4.2.1 Relación espacial entre dos píxeles.....	58
4.2.2 Distancia entre píxeles.....	58
4.2.3 Normalización de la matriz.....	59
4.2.4 Propiedades de la matriz.....	62
4.3 Algoritmo rápido para calcular GLCM.....	63
4.4 Descriptores de textura de Haralick.....	68
4.4 Representaciones multibanda.....	71

Capítulo V Campos aleatorios de Markov

5.1 Nociones topológicas.....	85
5.2 Sistema de vecinos y tipos de cliques asociados.....	90
5.3 Campos aleatorios de Markov y distribución de Gibbs.....	94
5.3.1 Campos aleatorios de Markov (CAM).....	94
5.3.2 Distribución de Gibbs.....	95
5.4 Modelos markovianos.....	97
5.4.1 Auto Modelos.....	97
5.4.2 Modelo de Potts.....	99
5.4.3 Modelo Logístico Multinivel.....	100
5.5 Algoritmo de metrópolis a temperatura constante.....	103
5.6 Ejemplo de modelo CAM para segmentación de imágenes.....	104

Capítulo VI. Esquema de Fisión y fusión de datos por campos aleatorios de Markov.

6.1 Introducción.....	106
-----------------------	-----

Indice

6.2 Fisión de datos.....	107
6.3 Fusión de datos.....	108
6.4 Resultados de fusión de datos.....	110
6.5 Algoritmo de recocido simulado para fusión de datos.....	115

Capítulo VII. Análisis de resultados

7.1 Resultados para imágenes DF1_izq (20,30,40,60 y 80 iteraciones).....	119
7.2 Resultados para imágenes DF1_pr1 (20,30,40,60 y 80 iteraciones).....	131

Capítulo VIII. Conclusiones

8.1 Conclusiones.....	143
-----------------------	-----

Capítulo IX. Referencias

Referencias.....	149
------------------	-----

Apéndice

A Teorema de Bayes.....	151
B Construcción de la función de energía.....	156
C Algoritmo de recocido simulado.....	160
D Programa de simulación del algoritmo de recocido simulado.....	163

I. Introducción

La observación a distancia de la superficie terrestre constituye el marco de estudio de la percepción remota. Esta es una traducción latina del término inglés *remote sensing*, ideado a principios de los 60 para designar cualquier medio de observación remota, si bien se aplicó fundamentalmente a la fotografía aérea, principal sensor de aquel momento. La percepción remota no engloba sólo los procesos que permiten obtener una imagen, sino también su posterior tratamiento en el contexto de una determinada aplicación.

La percepción remota es una ciencia que estudia una escena sin que se tenga que entrar en contacto con los elementos que la componen, utilizando diferentes clases de instrumentos para obtener imágenes del entorno. Uno de los instrumentos más utilizados para obtener imágenes de percepción remota apareció gracias al desarrollo tecnológico del satélite, el cual utiliza instrumentos de muy alta resolución en diferentes bandas del espectro electromagnético para obtener imágenes de nuestro planeta y del universo.

Si hemos definido a la percepción remota como aquella técnica que permite adquirir imágenes de la superficie terrestre desde sensores instalados en plataformas distantes a ésta, estamos suponiendo que entre la Tierra y el sensor existe un fenómeno de transporte de energía, ya sea por reflexión de la energía solar o de un haz energético artificial. A su vez, es preciso que ese haz energético recibido por el sensor sea almacenado convenientemente, bien a bordo de satélites, bien en las estaciones receptoras, de tal manera que pueda interpretarse para una determinada aplicación. El análisis de la información recibida se incluye también en sentido amplio dentro del estudio de la percepción remota, aunque esté más allá de los procesos de adquisición propiamente dichos.

Un sistema de percepción remota incluye los siguientes elementos:

- 1.- Fuente de energía que emite la radiación electromagnética que detecta el sensor. Puede tratarse de una fuente externa a éste, en cuyo caso se habla de percepción remota pasiva, o de una fuente emisora de haz energético en el propio sensor, lo que se denomina percepción remota activa. La fuente de energía más importante es obviamente el sol.
- 2.- Cubierta terrestre, formada por distintas masas de vegetación, suelos, agua o construcciones humanas, que reciben la señal energética procedente de la fuente de energía y la reflejan o emiten de acuerdo a sus características físicas.
- 3.- Barredor multiespectral compuesto por un arreglo de sensores y la plataforma que lo alberga. Tiene como misión captar la energía procedente de las cubiertas terrestres, codificarla y grabarla o enviarla directamente al sistema de recepción.
- 4.- Sistema de recepción (comercialización) en donde se recibe la información transmitida por la plataforma, se graba en un formato apropiado y tras las oportunas correcciones, se distribuye a los interesados.
- 5.- Intérprete que convierte esos datos en información temática de interés, ya sea visual o digitalmente, de cara a facilitar la evaluación del problema en estudio.

El símil de la visión humana puede ayudar a entender el significado de estos componentes. El ojo humano (sistema sensor) recibe un haz energético procedente de la escena (cubierta terrestre) por reflejo de la luz solar (fuente de energía). Esta señal se transmite al cerebro (sistema de análisis) que es capaz de formar una serie de imágenes sobre el mundo real que le circunda. El individuo que observa es a la vez intérprete y usuario final de la imagen detectada, lo que permite tomar decisiones sobre su propio comportamiento. La visión humana forma un sistema de percepción remota muy sofisticado por cierto, puesto que nos permite caracterizar con mucho detalle los objetos que observamos.

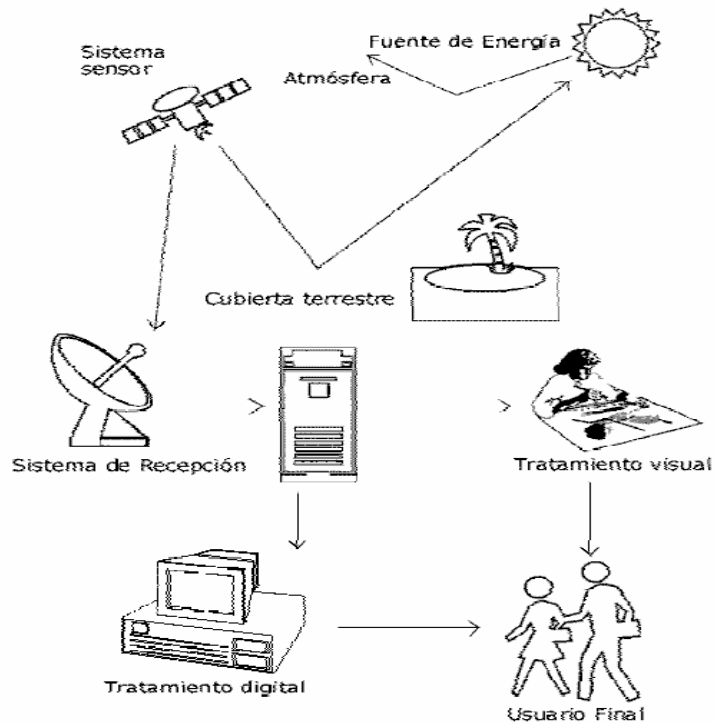


Figura 1.1 Componentes de un sistema de percepción remota

Sin embargo, nuestra percepción visual presenta también algunas limitaciones. Está restringido por la sensibilidad espectral de nuestras células sensoras, que sólo nos permiten apreciar un determinado tipo de energía, denominado por esta razón, espectro visible. Otras formas energéticas, como el calor, no son observables visualmente, por lo que parece conveniente contar con unos ojos artificiales que amplíen nuestras propias posibilidades. También, el campo de visión humano está limitado por nuestra propia estatura, o a la que podemos adquirir desde observaciones naturales, por lo que cuenta con un pequeño radio de acción. Además, se trata de una percepción oblicua, lo cual junto con el factor anterior dificultan observar fenómenos muy extendidos, como inundaciones o incendios donde las estimaciones visuales pueden ser imprecisas. Para solventar estas limitaciones, se ha acudido a sensores artificiales montados sobre plataformas situadas a una cierta altitud. Con ellos se tiene acceso a tipos de energía no visibles (ultravioleta, infrarrojo y microondas) y además desde una nueva perspectiva vertical y panorámica. Esta nueva información

expande notablemente nuestro conocimiento del medio que nos rodea, facilitando la interpretación de múltiples procesos que afectan al planeta.

El creciente empleo de distintos sensores de observación terrestre está originando no solo una enorme cantidad de información, sino también una nueva forma de estudiar la superficie terrestre. La percepción remota, junto con las técnicas añejas de almacenamiento y proceso de la información geográfica, permite disponer de una masa de datos sobre el territorio, algunos de ellos accesibles en tiempo real. Para aprovechar mejor estas técnicas, el recurrir a las computadoras resulta casi inevitable. Una vez salvado el necesario entrenamiento, estos equipos amplían notablemente nuestra capacidad de análisis de resultados, tarea en donde la inteligencia humana resulta insustituible librándola de otras labores mecánicas (acopio y ordenamiento de datos) que han consumido tradicionalmente las mejores fuerzas del investigador [1].

1.1 Objetivo

El objetivo de esta Tesis es desarrollar, a través de métodos de análisis de texturas y técnicas de fusión y fisión de datos, un esquema estadístico para la segmentación de fotografía aérea de alta resolución.

1.2 Antecedentes del Tema

El hombre siempre ha querido conocer y distinguir el ambiente que lo rodea, de aquí se desprende el estudio de los diferentes instrumentos y formas de utilizar la información de acuerdo a sus necesidades. Gracias a esto se han desarrollado campos de estudio de visión, reconocimiento de patrones y percepción remota.

El origen de la percepción remota se liga a los experimentos de Newton (1822), quien constató que un rayo luminoso (luz blanca), al atravesar un prisma, se desdobra en un espectro. Se realizaron experimentos en donde la luz blanca se ve como una síntesis de diferentes tipos de luz, una especie de vibración compuesta de muchas vibraciones

diferentes. Posteriormente, se descubrió que cada color que forma parte del espectro corresponde a una temperatura diferente.

Gracias a diferentes experimentos y el trabajo teórico de Maxwell se probó que la onda de luz es una onda electromagnética, mostrando que la luz visible es sólo una de las diferentes especies de ondas electromagnéticas.

Es evidente que la percepción remota es fruto de un esfuerzo multidisciplinario que envuelve avances en la física, fisico-química, geociencias, computación mecánica, etc.

La percepción remota es una técnica aplicada y como tal dependiente del estado tecnológico existente en cada momento. Por tal reúne aspectos muy variados óptica y detectores del sensor, vehículo que lo sustenta, sistemas de transmisión, equipos de tratamiento, etc. Las formas de la percepción remota han variado drásticamente en las últimas décadas. El crecimiento ha sido realmente vertiginoso, tanto en la cantidad como en la variedad y calidad de la información disponible para campos científicos variados.

Las primeras experiencias de la percepción remota se remontan a 1859, cuando Gaspar Félix de Tournachon obtuvo las primeras fotografías aéreas desde un globo cautivo (Hyan, 1988). Al año siguiente, James Wallace repitió la experiencia sobre la ciudad de Boston, poniéndose en evidencia el interés de la nueva perspectiva aérea para un conocimiento más detallado de la organización urbana.

En los años posteriores se produjo un importante desarrollo de la observación fotográfica desde avión, gracias a los progresos realizados en óptica y emulsiones. En 1909, Wilbur Wright adquirió la primera fotografía aérea, abriendo el camino a una larga historia de observación desde plataformas remotas. La primera cámara aérea, propiamente dicha, se desarrolló durante la 1° Guerra Mundial . Concretamente en 1915, por J.T.C. Moore Brabazon. La importancia estratégica de este escrutinio quedó pronto en evidencia, desarrollándose notablemente las técnicas de adquisición y procesado fotográfico Brookes, 1975.

La 2° Guerra mundial implicó un notable desarrollo de las técnicas de percepción remota aérea. El progreso se orientó a mejorar la óptica de las cámaras de reconocimiento, así como las emulsiones utilizadas (de esta época proceden las primeras películas en infrarrojo desarrolladas por Kodak Research Laboratories). Así mismo se introdujeron nuevos sensores como el radar y se mejoraron los sistemas de comunicación. Por otro lado, la aeronáutica prosperó notablemente, lo que aseguró mayor estabilidad a las plataformas de observación.

1.3 Definición del problema

Existen regiones del planeta de difícil acceso para la adquisición de datos. Las condiciones meteorológicas son factores determinantes en regiones que frecuentemente están cubiertas por nubes. Otro factor adicional son las condiciones de iluminación del sol en las imágenes adquiridas en el rango espectral visible e infrarrojo. Es en este contexto que las tecnologías contemporáneas de procesamiento digital de imágenes y de percepción remota se sitúan como fuentes de datos útiles a los sistemas de información ambiental.

Se propone un esquema de fusión de datos que permita integrar la información parcial proveniente de diversas fuentes en este caso los captadores seleccionados a partir del proceso de la fisión de datos el cual mas adelante se hablará en detalle, con el esquema propuesto de fisión a partir de las imágenes seleccionadas se intenta resolver el problema de los modos mezclados de la imagen original. La teoría de los campos aleatorios de Markov facilita la integración de datos parciales en un entorno de análisis contextual, razón por la cual se emplean en el esquema propuesto de fisión y fusión de datos en el presente trabajo de tesis.

Se presenta una aplicación para zonas urbanas con 2 ejemplos. En el primer ejemplo se propone la segmentación de imágenes de fotografía aérea de alta resolución de escenas urbanas del Distrito Federal. En el segundo ejemplo, se propone nuevamente la segmentación de otra escena del Distrito Federal distinta a la anterior para corroborar el algoritmo de fusión de datos.

1.4 Metodología empleada

En el presente trabajo de tesis se ha propuesto una metodología llamada fisión y fusión para segmentación de imágenes. El propósito del esquema de fisión-fusión es para solucionar el problema de que en la imagen original los modos están mezclados de manera tal que hacen difícil la obtención de un buen resultado, mientras que en los captosres (imágenes de textura) no están mezclados por lo que la forma del histograma es mejor. Con la fusión de datos lo que se pretende también es juntar informaciones parciales para obtener un resultado más completo. Por ende con la fusión de datos se llega a una aproximación en la solución al problema de los modos mezclados en la imagen original.

Una de las principales ventajas de este esquema de fisión-fusión es que da como resultado un esquema sencillo, el cual permite la obtención de datos de informaciones parciales provenientes de la imagen original para obtener un mejor resultado. Por otro lado, una desventaja de este esquema de fisión-fusión es que en las imágenes de textura (captosres) la información no siempre es muy buena, de allí que entra el factor de confiabilidad para la discriminación de cada media de clase, restándole calidad al proceso de segmentación.

Para el esquema propuesto, el paso de fisión consiste de obtener estimadores texturales también llamados descriptores de textura, los cuales son obtenidos por el método de matriz de co-ocurrencia.

Robert Haralick propuso lo que en la actualidad se conoce como el método de la matriz de co-ocurrencia, la cual es representada por la distancia y el ángulo de la relación espacial sobre la subregión de una imagen de tamaño específico. Cada elemento de esta matriz representa la probabilidad de ocurrencia de dos niveles de gris separados por una distancia y dirección dada.

El histograma describe la frecuencia de un nivel de gris dentro de la ventana de un área determinada.

Como parte del proceso de fusión se obtienen las funciones de energía por campos aleatorios de Markov. Los resultados de la teoría de campos aleatorios de Markov (CAM) nos proporcionan herramientas para codificar restricciones contextuales en la probabilidad a priori. Esta es la principal razón para el modelo de visión usando CAM.

Por otro lado, el paso de fusión toma las correspondientes funciones de energía de cada imagen de cada una de sus clases, se multiplican por un factor de confiabilidad que nos dice la calidad de la clase con respecto a las demás en los captores, posteriormente se suman sus respectivas funciones de cada captor respecto a cada clase ya multiplicada por su factor de confiabilidad y posteriormente se dividen entre una normalización de sus factores. Una vez que se realizó lo anterior, se prosigue con el modelo de Potts en un análisis en 8 conectividad, comenzando con una inicialización aleatoria en el número de clases y terminando con la imagen segmentada en las clases deseadas.

El esquema de fisión y fusión de datos empleado en el presente trabajo de tesis se presenta en la figura 1.2. En este esquema se muestran las diversas etapas que constituyen la fisión al obtener las imágenes de textura provenientes de la imagen original y posteriormente la etapa del paso de la fusión de datos para obtener finalmente la imagen segmentada en las clases deseadas.

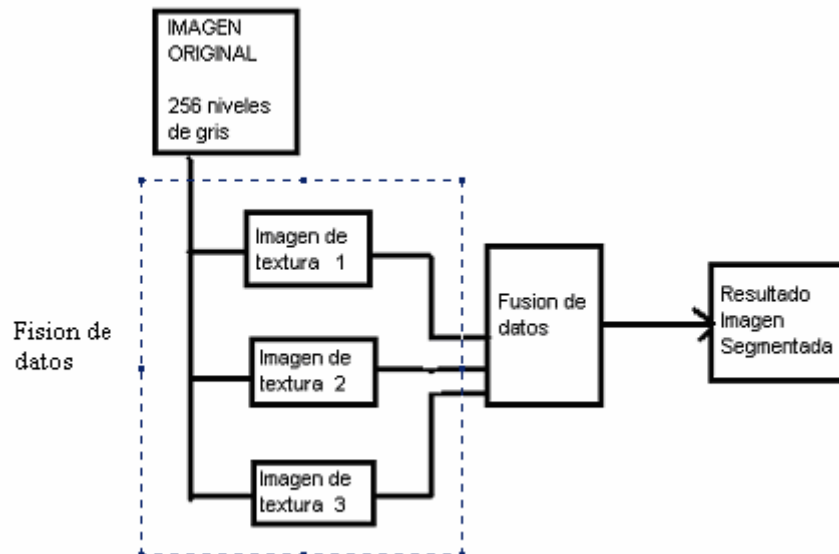


Figura 1.2 Esquema de fisión y fusión de datos

Una vez que se han seleccionado cuidadosamente las imágenes de textura que se van a fusionar, hay que identificar las clases en cada imagen, tomando en cuenta que en algunas imágenes las clases se ven mejor que en otras, para poder tomar un criterio en el factor multiplicativo de confiabilidad. Para poder fusionar los datos de cada descriptor de textura primero hay que obtener las funciones de energía. En el caso del presente trabajo de tesis por campos de Markov, las funciones de energía nos dan la probabilidad de que un píxel dado pertenezca a una clase en particular. En la figura siguiente se muestra el esquema de fusión de datos con mayor detalle, tomando en cuenta algunos de los descriptores de textura de la imagen original.

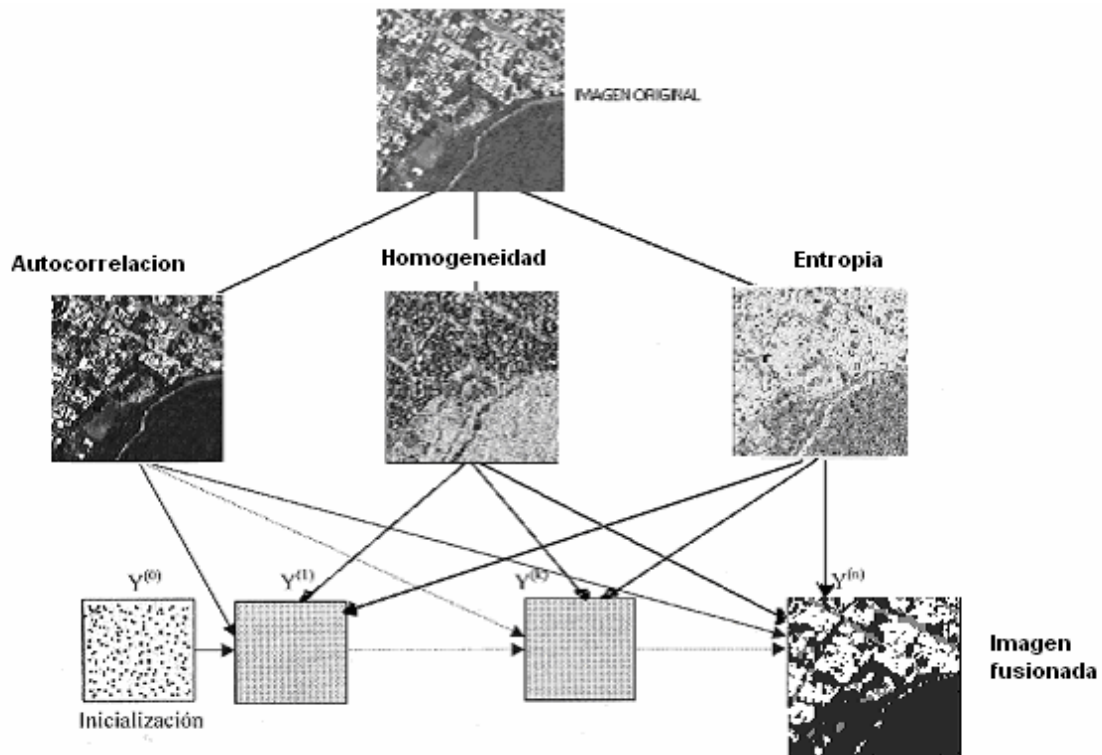


Figura 1.3 Esquema de fisión y fusión de datos detallado

II. Fundamentos de la percepción remota

2.1 Descripción general

La percepción remota (*Remote Sensing*) o **teledetección** puede definirse como la ciencia y arte de obtener información de un objeto analizando los datos adquiridos mediante algún dispositivo que no está en contacto físico con dicho objeto.

Otra definición de percepción remota puede ser: "Es la utilización de sensores para la adquisición de información sobre objetos o fenómenos sin que haya contacto directo entre el sensor y los objetos". Ud. mismo, al leer estas líneas, está ejecutando un acto de percepción remota: un ente físico, la luz que emana de la *fuentes emisora*, en este caso la pantalla de su PC, atraviesa cierta distancia hasta que es capturada por un *sensor*, sus ojos, que la envían a un *procesador*, su cerebro.



Figura 2.1 El esquema operativo de un satélite de observación

La historia de la Percepción Remota comenzó hace unos 600 millones de años, cuando alguna forma inferior de vida animal diferenció algunas de sus células, volviéndolas fotosensibles.

También durante millones de años dicho rudimento fotosensible evolucionó convirtiéndose en un poderoso y sofisticado sensor, el ojo humano. Este tuvo un imitador mecánico, la cámara fotográfica, que hizo su aparición hace algo más de un siglo y que fue muy mejorada durante la década de 1930 para ser aplicada a la fotografía aérea. La segunda guerra mundial dio un gran

impulso a la fotografía aérea, así como a otras formas de percepción remota. Sin embargo, el salto fundamental en esta disciplina se produjo en la década de los 60's[1] cuando las plataformas satelitales reemplazaron a las aéreas y los sensores electrónicos multiespectrales, acoplados a computadoras, reemplazaron las cámaras fotográficas. El esquema operativo de un *satélite de observación* se representa en forma muy simplificada, en la Figura 2.1

Los objetos terrestres, iluminados por la radiación solar, reflejan ésta luego de introducir en ella modificaciones inducidas por la misma estructura y composición de dichos objetos. La radiación reflejada es capturada por los sensores del satélite, siendo parcialmente procesada a bordo de éste y retransmitida a estaciones receptoras terrestres para su posterior procesamiento y análisis (*fuentes emisora, sensor, cerebro*) [13][1].

2.2 Sistemas de percepción remota

De acuerdo a la plataforma donde se ubique el sensor, se distinguen dos grandes tipos de sistemas de percepción remota:

- Percepción remota aérea
- Percepción remota espacial

2.2.1 Percepción remota aérea

Se considera percepción remota aérea aquella que se obtiene a partir de las plataformas:

- Globo
 - Avión
 - Helicóptero
-
-

Los sensores más usados son:

- Cámara fotográfica
- Cámara multibanda
- Cámara digital
- Cámara de video
- Barredor multiespectral
- Barredor térmico

Las fotografías aéreas son las imágenes de percepción remota más comúnmente utilizadas. Son usadas para el análisis de todo tipo de recursos naturales, cartografía y estudios de aproximación en los sectores público y privado.

Las ventajas básicas que ofrece esta técnica son:

- Permite tener una vista de los detalles de la superficie en estudio en su contexto espacial a menor escala que la imagen satelital. Esto la transforma en una herramienta muy útil para trabajar en áreas pequeñas
- Flexibilidad. Por el hecho de usar plataformas aéreas para embarcar los sensores de imagen, es muy fácil rediseñar el plan de trabajo según las necesidades que vayan surgiendo.
- La fotografía nos da una visión instantánea de procesos bajo condiciones dinámicas (inundaciones, derrames de petróleo, tráfico, incendios, etc.). Si bien este tipo de trabajo se realiza muy bien con imágenes satelitales, 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 al de los satélites (catástrofes).

El sensor usado en esta técnica es la película fotográfica. Existen diferentes tipos de película pensados para diferentes tipos de problemas.

Las películas a las que estamos más acostumbrados son aquellas que son capaces de “ver” la imagen en forma similar a como lo hace el ojo humano. Estas películas con algunas características adicionales, son usadas en fotografía aérea y permiten analizar y resolver muchos problemas.

2.2.2 Percepción remota espacial

Se considera percepción remota espacial aquella que se obtiene a partir de las plataformas que pueden ser naves espaciales tripuladas, estaciones orbitales o satélites autónomos, que giran alrededor de la tierra en órbitas polares, oblicuas o ecuatoriales. Sus ventajas son: la visión sinóptica, la periodicidad en la toma de datos, el sincronismo con el sol y el bajo costo por hectárea de los datos generados

Algunos de los sistemas satelitales más comúnmente utilizados son:

- ERS-1 y 2 (ESA)
- SPOT 1,2,3 y 4 (Francia)
- IKONOS
- LANDSAT 5 y 7 (USA)
- RADARSAT (Canadá)
- JERS-1 (Japón)

2.3 Aplicaciones de los sistemas de percepción remota

Las aplicaciones para las imágenes satelitales son una fuente global de recursos para los gobiernos, las empresas y los ciudadanos particulares. La información que se puede extraer de la siempre cambiante faz de la tierra, tanto en sus aspectos naturales como culturales, proporciona una nueva dimensión visual de análisis que permite comprender en profundidad los fenómenos del cambio económico, la forma en que afectamos el medio ambiente y el modo como nos proyectamos hacia el futuro.

Geología

- Reconocimiento de tipos de roca.
- Cartografía de zonas geológicas
- Revisión de mapas geológicos.
- Delineación de rocas y suelos no consolidados.
- Cartografía de depósitos recientes en superficies volcánicas.
- Búsqueda de guías de superficie para mineralización.
- Determinación de estructuras regionales.
- Cartografías lineales.

Agricultura y bosques

- Discriminación de tipos de vegetación: tipos de cultivos, tipos de madera, etc.
- Determinación de la vegetación activa (proceso de creación de clorofila)
- Medición de extensión de cultivos y maderas por especies.
- Clasificación de usos del suelo.

Recursos hidrográficos

- Determinación de límites, áreas y volúmenes de superficies acuáticas
- Cartografía de inundaciones
- Determinación del área y límites de zonas nevadas
- Medida de regiones glaciales
- Medida de modelos de sedimentación y alteraciones
- Determinación de la profundidad del agua
- Delineación de campos irrigados.
- Clasificación de lagos.
- Detección de zonas de alteraciones hidrotermales.

Cartografía y planeamiento urbano

- Cartografía y actualización de mapas
- Clasificación de la fertilidad de la tierra
- Clasificación de zonas urbanas y rurales.
- Planificación regional.
- Cartografía de redes de transporte
- Cartografía de límites tierra-agua.
- Cartografía de fracturas.

Meteorología

- Predicción de desastres naturales de origen climático
- Análisis de masas nubosas y su evolución.
- Modelización climática a diferentes escalas.

Medio ambiente

- Control de superficies mineras
- Cartografía y control de contaminación de aguas
- Detección de contaminación del aire y sus efectos.
- Determinación de efectos de desastres naturales
- Control medioambiental de actividades humanas
- Seguimiento de incendios forestales y sus efectos
- Estimación de modelos de erosión.

Oceanografía y recursos marítimos

- Detección de organismos marinos vivos
- Determinación de modelos de alteración y circulación
- Cartografía térmica de la superficie del mar
- Cartografía de cambios en límites de las costas
- Cartografía de costas y áreas superficiales
- Cartografía de hielos para navegación
- Estudio de mareas y olas.

2.3.1 Algunas aplicaciones utilizando imágenes de la superficie de la tierra

El programa satelital nos provee información oceanográfica y meteorológica en tiempo real, de manera que es posible monitorear sus parámetros físicos. Para ello, muchas técnicas son empleadas en la investigación de los océanos, una de ellas es la percepción remota, mediante la cual se obtienen los datos de la superficie del planeta utilizando los sensores remotos instalados en los satélites u otros sistemas aéreos. Surgen diversas aplicaciones utilizando imágenes de la superficie de la tierra tales como:

- Cartas de temperatura superficial del mar(TSM)
- Monitoreo de la laguna “La niña”
- Localización de incendios forestales
- Detección de derrames de hidrocarburos
- Detección de zonas urbanas
- Imágenes provenientes del satélite Landsat (Imágenes de alta resolución)

2.3.2 Cartas de temperatura superficial del mar (TSM)

Las cartas de temperatura superficial del mar (TSM) son elaboradas desde la imagen original, a la cual se le somete a una calibración por el sistema, una rectificación geométrica y una corrección por interferencia atmosférica. El cálculo de la TSM se realiza utilizando el algoritmo matemático *split night multi-channel sea surface temperature*.

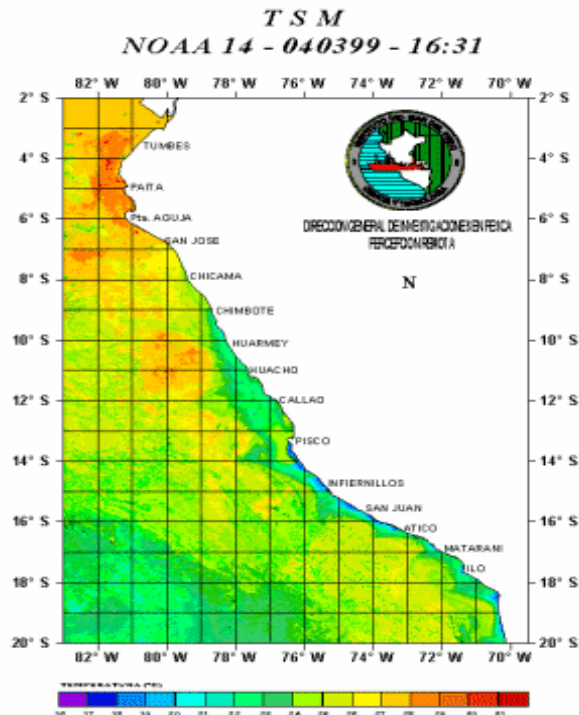


Figura 2.2 Carta de temperatura superficial del mar (TSM)

Con las cartas de TSM ha sido posible el monitoreo del evento “El Niño” en la costa peruana y en zonas específicas (puertos y caletas), desde febrero de 1998.

2.3.4 Monitoreo de Laguna "La Niña"

La laguna “La Niña” se formó como producto de las torrenciales lluvias que provocó el evento “El Niño” 1997-98 en los meses de febrero y marzo de 1998, en la zona de los desiertos de Sechura (Perú). Tiene sus orígenes en el represamiento de los enormes caudales de los ríos Piura, La Leche y Cascajal, Olmos, Motupe, sumados a ello las lagunas de Ramón, Ñapique (Perú), entre otros y por la configuración de la topografía de la zona lo que favoreció su formación. El nombre de esta laguna se debe a que se origina dentro del episodio de “El Niño”.

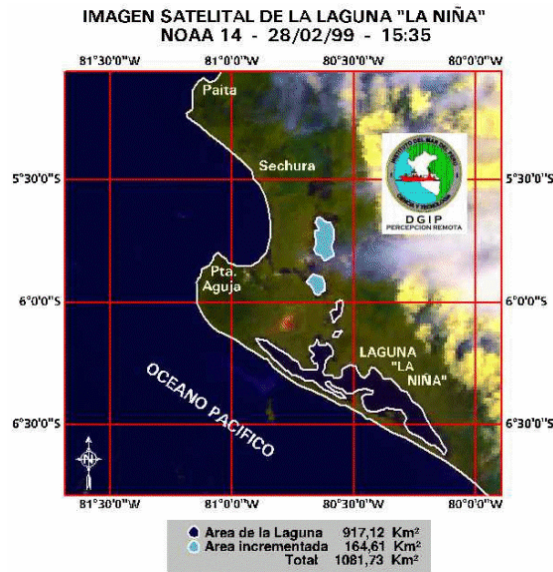


Figura 2.3 Monitoreo de la laguna la niña

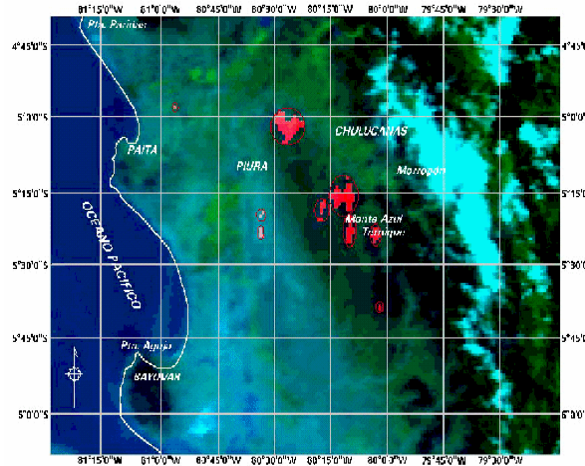
La forma y extensión de la laguna ha sido posible observarse mediante la utilización de imágenes de alta resolución de los satélites polares NOAA.

El área máxima que logró alcanzar esta laguna fue de 2,326 km en el mes de marzo y a fines de diciembre se registró un área de 1,082 km, lo que representa la pérdida de 53.48% de su extensión en siete meses de monitoreo.

2.3.5 Localización de Incendios forestales

A fines de noviembre de 1998, se presentó un incendio forestal en los bosques de Piura, el cual fue detectado por el sistema satelital HRPT, mediante los sensores AVHRR de los satélites NOAA. El siniestro ocurrido en el distrito de La Matanza, provincia de Morropón, fué monitoreado con fines de estudio y vigilancia.

La realización del presente trabajo se hizo teniendo en cuenta que, para detectar incendios por medio del sensor AVHRR del satélite NOAA, es necesario utilizar los contrastes que se producen en el canal 3 del sensor (canal térmico).



25 Noviembre, 1998 - 14:49

Figura 2.4 Localización de incendios forestales

En este canal los píxeles correspondientes a las zonas de incendios se saturan, resaltando así sobre el resto de los píxeles y permitiendo de esta forma localizar las zonas de incendios.

Las imágenes obtenidas corresponden a una combinación de las bandas 3-2-1, de la escena NOAA 14, capturadas con una antena SMARTech, y procesadas con el software *ERDAS Imagine 8.3*[13][1].

2.3.6 Imágenes provenientes del satélite LANDSAT (resolución media)

En esta sección se muestran algunas imágenes provenientes del satélite LANDSAT para que el lector pueda percibir que tipo de imágenes se obtienen de este satélite. En la figura 2.5 (a) tenemos una imagen tomada el 1 de Enero de 1997, Landsat 5, composición RGB falso

infrarrojo bandas 4,3,2 cuadrante de 90 x 90, 30 mts de resolución por pixel. En este caso, debido al gran tamaño de la imagen, su resolución fue reducida a un 10% de la original. Esto es lo que en percepción remota llamamos vista rápida o "Quick Look".

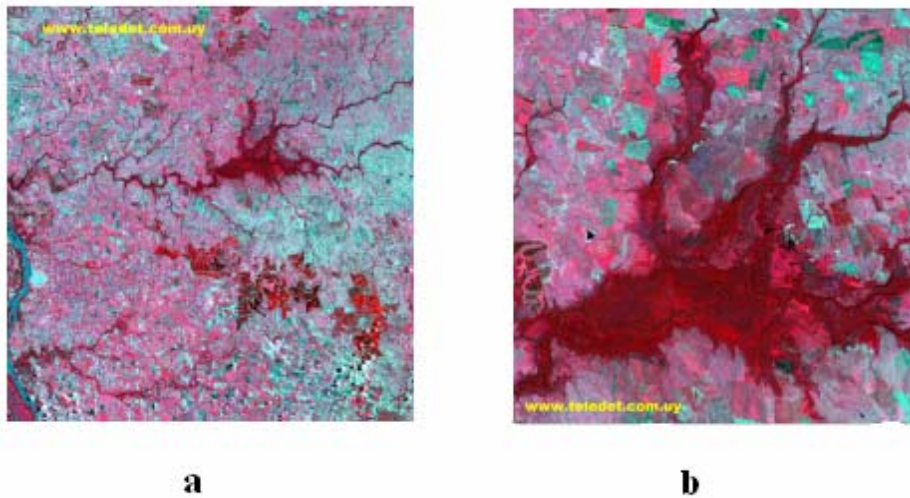


Figura 2.5 (a) zona: departamentos de Paysandú y Río Negro (Perú) (b) Subescena del cuadrante anterior, con su resolución original de 30mts por pixel: montes del Queguay (Perú)

En la figura 2.5 (b) se aprecia una composición RGB falso infrarrojo en las bandas 4,5,3 de 30 metros de resolución por píxel de la zona arrocera al sur de la ciudad de Artigas en Perú tomada el 1 de Febrero de 1999 por el satélite Landsat 5 [11].

2.4 Imágenes multiespectrales

Se les denomina "imagen multiespectral" al conjunto de imágenes digitales provenientes de los filtros de color que se hayan empleado en el sistema de percepción remota (incluso en "colores" como el ultravioleta y en el infrarrojo que son invisibles al ojo humano pero que pueden ser detectados por un sensor remoto moderno) los cuales reciben la luz de un conjunto de dispositivos de conducción localizados en un sistema óptico en el sensor opto electrónico el cual a su vez recibe información proveniente de un campo instantáneo de vista. En el siguiente

ejemplo podemos ver el tipo de detalles que se ve al volar sobre la superficie terrestre a 2500 m de altura para imágenes de fotografía aérea multispectral.



Figura 2.6 Campos cercanos al Río Negro (Uruguay), color normal.

Para otros tipos de problema es necesario recurrir a películas que "ven colores" en donde el ojo humano no ve nada. Estamos hablando en este caso de film sensible a la radiación infrarroja. El ejemplo siguiente corresponde a la misma imagen pero usando un film sensible al infrarrojo cercano. Esta foto es en falso color ya que no podemos hablar de color en el infrarrojo. Lo que se hizo fue asignar en la película el color rojo a la radiación infrarroja cercana reflejada por la superficie en estudio.

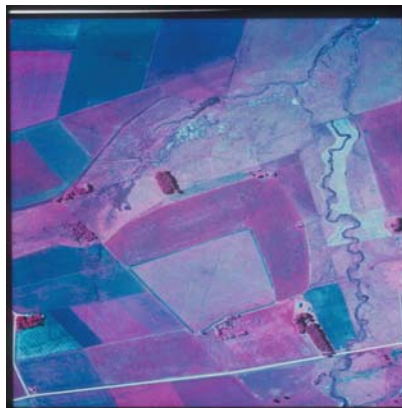


Figura 2.7 Campos cercanos al Río Negro (Uruguay). Infrarrojo.

Otra característica importante es que en este film se está filtrando completamente el color azul eliminando el efecto de dispersión de luz atmosférica conocido como "haz". Este efecto disminuye el contraste de la imagen y le da una tonalidad azul grisácea. Si comparamos ambas fotos vemos la eliminación de este efecto en la fotografía infrarroja. En la figura 2.8 es posible ver y analizar los detalles (p. ej. heterogeneidades en las áreas cultivadas) que se presentan en la fotografía infrarroja, en comparación con su correspondiente versión en color real.



Figura 2.8 Eliminación del haz en la fotografía infrarroja.

Cuando hablamos de fotografía infrarroja debemos de tener cuidado de no confundirla con la fotografía térmica. En esta técnica también estamos trabajando con radiación infrarroja pero de diferente tipo a la que es sensible la película infrarroja que se usó en este ejemplo. En esta película queda registrado un rango de radiación infrarroja que da información acerca de la naturaleza química de los objetos fotografiados. Algunos de los trabajos posibles a realizar con esta técnica son:

- + Agricultura de precisión
- + Clasificación de cultivos.
- + Detección y seguimiento de enfermedades.
- + Estrés fisiológicos de los cultivos.
- + Cálculo de áreas forestales y análisis de estado sanitario.
- + Índice de defoliación de cultivos por plagas

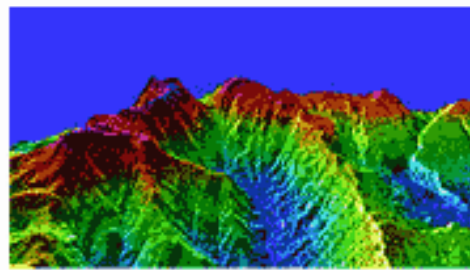
2.5 Imágenes de radar de apertura sintética SAR

El sensor radar de apertura sintética (SAR) de RADARSAT-1 capta información en la región de las longitudes de ondas de microondas y como sensor activo que es, provee su propia fuente de energía, pudiendo operar tanto de día como de noche; además de penetrar la atmósfera aún en condiciones meteorológicas extremas como nubosidad, neblina o lluvia. La capacidad de seleccionar la resolución de la imagen y el área cubierta, entre otros, le brinda ventajas comparativas con respecto a lo que son los satélites con sensores ópticos tradicionales.

El **monitoreo de eventos** puede realizarse sin problemas, dado que la periodicidad con que se pueden tomar las imágenes de una misma área es en tiempo mínimos y de 4 a 5 días, en situaciones máximas.

La **combinación de imágenes radarsat** con otras fuentes de datos como landsat, permiten resaltar importantes rasgos en las imágenes que antes eran difusas o bien interpretar nueva información, que por si solas no se puedan detectar en ambas fuentes.

La **flexibilidad de su antena**, permite obtener una gran variedad de imágenes, las cuales pueden cubrir desde grandes extensiones de terreno hasta áreas muy específicas con gran nivel de resolución de los datos. El monitoreo de la tierra con un ángulo de observación desde la derecha, le otorga la capacidad de obtener pares estereoscópicos, generando de esta manera la posibilidad de utilizar estas imágenes en la construcción de diversos productos, tales como, modelos de elevación digital, cartografía, mosaicos, entre otros.



Modelo de Elevación Digital PAPUA NUEVA GUINEA
 Dos Imágenes RADARSAT, Modo STANDARD
 posiciones 2 y 7
 (Fecha: Agosto 26 y 29 de 1996)

Figura 2.9 Modelo de elevación digital Papua Nueva Guinea

En una órbita que tiene un período de 100,7 min, el satélite tiene una capacidad de levantar 28 min (aprox.) de datos de la superficie terrestre. Este período de toma de datos corresponderá a lo definido y programado por radarsat.

La definición y características de este "calendario" de toma de datos, será el resultado de las necesidades planteadas por los usuarios que, con la ayuda e interacción de especialistas de **Radarsat** e **INCOM S.A. (Empresa de ingeniería de satelites)**, definen qué opción (ej: día, pasada ascendente y/o descendente, resolución, ángulo del haz, cobertura, y otros) de las que ofrece el satélite, es la adecuada a sus necesidades.

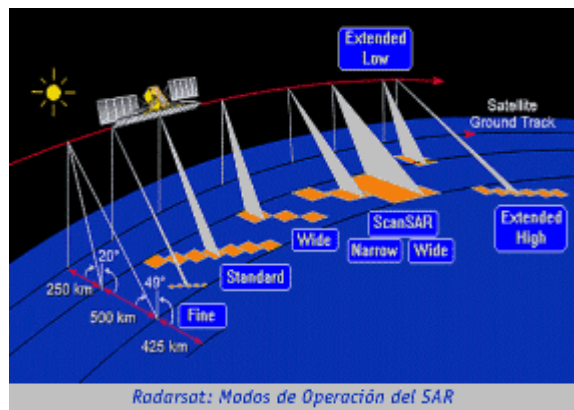


Figura 2.10 Modos de operación del SAR

Las características especiales de su diseño permiten a los usuarios elegir el tipo de imagen de entre una gran variedad de alternativas, que van desde el modo conocido como FINO, el que tiene resoluciones espaciales de 8 m y una cobertura de 50 km, hasta el modo llamado barrido amplio, el que tiene una resolución espacial de 100 m una cobertura de 500 km.

Las capacidades señaladas, garantizan a cualquier usuario la continuidad de los datos, cuya utilización se ha ido masificando en actividades tan variadas como la minería, agricultura, cartografía, hidrología, oceanografía, monitoreo, medioambiente, telecomunicaciones, meteorología, defensa, administración de desastres, emergencias y el área forestal entre otras.

Cuando un usuario requiere imágenes de Radarsat-1, lo primero que debe hacer es definir el área geográfica de su interés (el satélite puede tomar imágenes de cualquier parte del mundo) y si es posible, el tipo de imágenes que son de su interés. Es necesario destacar que existen para los diferentes modos de adquisición de las imágenes, una variedad de ángulos que pueden ser seleccionados. Esto hace que sea necesario que exista una interacción entre el usuario, **INCOM** y **Radarsat**, para definir la mejor opción en función a lo requerido. Con estos datos **Radarsat** definirá los "Path/Row" necesarios para cubrir la zona deseada en el modo de operación adecuado.

2.5.1 Detección de derrames de hidrocarburos por medio del SAR

El SAR es una herramienta excelente para monitorear y detectar aceite en superficies de agua en todos los climas. En las imágenes los derrames aparecen como filamentos conectados con áreas extensa.

Derrame petrolero en las costas de Italia

En esta imagen nocturna proveniente del SAR es posible apreciar dos derrames petroleros de barcos. Uno tiene 22 km y el otro 7 km de longitud.



Figura 2.11 Derrames petroleros en las costas de Italia

Filtración natural, mar caspio.

Este espiral está esparcido en el piso del océano, es distinto de los derrames petroleros de barcos ya que este es un punto fuente, la dirección que tiene es debido a los vientos y corrientes que prevalecen en la región.



Figura 2.12 Filtración natural, mar caspio

2.5.2 Detección de inundaciones por medio del SAR

Dos imágenes del SAR de fechas diferentes pueden ser combinadas. Asignando diferentes bandas de colores a fechas diferentes revelará cualquier cambio en reflectividad de radar en colores diferentes. Esto es muy útil para todo tipo de cambios en monitoreo, particularmente en inundaciones como en estos ejemplos, donde las capacidades de todo tipo de clima del radar pueden proveer una única fuente de información a tiempo. En las imágenes las zonas azules indican zonas de inundación [17].

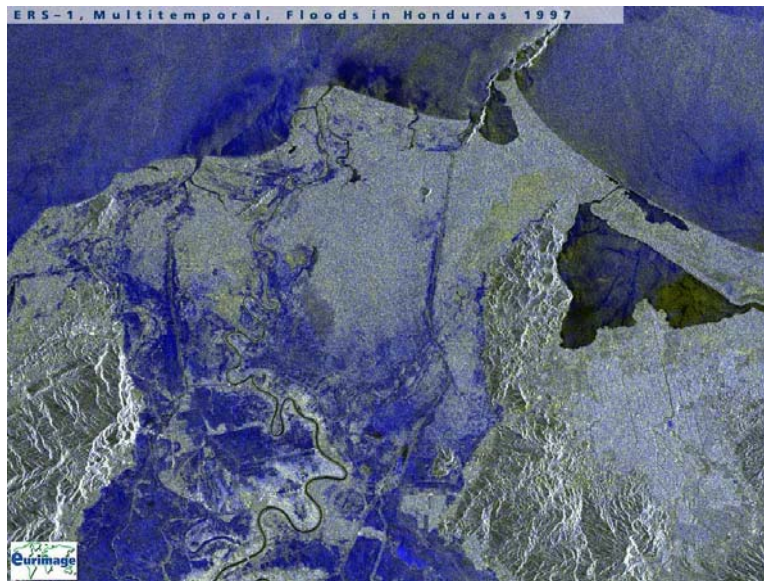


Figura 2.13 Imagen de inundación en honduras 1997

III. PRINCIPALES SATELITES DE PERCEPCION REMOTA

3.1 ENVISAT

El 1 de marzo de 2002 la agencia espacial europea (ESA) puso en órbita con éxito el satélite ENVISAT (environmental satellite). La misión ENVISAT implica la mayor contribución al programa de observación de la Tierra de la ESA y está equipado con 10 instrumentos diseñados para obtener datos sobre la atmósfera, los océanos, las zonas terrestres y las regiones polares.

Envisat, con sus 10 instrumentos a bordo, puede supervisar los fenómenos más importantes y problemáticos como la capa de ozono, el progresivo cambio de la vegetación en zonas amenazadas, el cambio del nivel y la temperatura de los océanos, las catástrofes humanas y naturales. De esta manera, el Envisat representa un buen aliado para las asociaciones de protección civil, que deben tener una visión exacta de las áreas afectadas por una catástrofe, o identificar otros lugares con mayor riesgo para desplegar sus fuerzas.

Envisat, tiene también aplicaciones comerciales. La principal aplicación en este campo, es la zona de hielo y su desarrollo en la comunicación. Supervisando el hielo, Envisat mejorará la navegación comercial en las zonas afectadas por los desplazamientos de peligrosos icebergs.

El satélite ENVISAT se encuentra en una órbita heliosincrónica a unos 800 km de altitud. Ofrece una cobertura global de la Tierra con un periodo nominal de 35 días; sin embargo, para la mayoría de los instrumentos ofrece una cobertura global cada dos o tres días, siendo más frecuentes las observaciones en las regiones situadas en latitudes altas.

La estrategia de mantenimiento de su órbita asegura que la desviación respecto a la órbita teórica de adquisición sobre el terreno (ground track) sea inferior a 1 km y que los horarios locales medios de las pasadas coincidan con el horario nominal en ± 1 minuto.

Para el envío de los datos a Tierra se utilizan estaciones terrenas habituales, pero la gran novedad estriba en el satélite de comunicaciones ARTEMIS (data relay satellite), lanzado a mediados de 2001, que actúa de apoyo para recibir los datos de ENVISAT cuando esté fuera de la cobertura de las estaciones terrenas y reenviarlos a éstas.

Para controlar la misión, la operabilidad del satélite y la provisión de servicios al usuario, ENVISAT cuenta con centros ubicados en diferentes países. El segmento de control de la órbita del satélite (FOS – Flight Operation Segment) está coordinado por el FOCC (Flight Operation Control Centre), localizado en Darmstadt (Alemania), y las estaciones de control y seguimiento de Kiruna y Svalbard.

El segmento de gestión de datos (PDS – Payload Data Segment) está coordinado por el PDCC (Payload Data Control Centre), localizado en ESRIN. Su cometido es la gestión de la programación de adquisiciones, proceso y archivo de datos, así como el soporte a los usuarios. El PDS consta de cuatro estaciones que reciben los datos directamente del satélite o a través de ARTEMIS y que pueden procesarlos en tiempo real. Así mismo, ENVISAT cuenta con centros de proceso y archivo (PAC's - Processing and Archiving Centres), distribuidos por Europa y entre los cuales se encuentra el E-PAC (Centro Español de Proceso y Archivo)[14].

Este último está englobado dentro del programa CREPAD y en él se procesan, archivan y distribuyen datos del sensor MERIS.

De forma esquemática, en la tabla 3.1 se muestra cada uno de los componentes que forman parte de la misión ENVISAT:

Instrumento	Características
ASAR	(Advanced Synthetic Aperture Radar): Para la observación de los océanos, tierras emergidas y casquetes polares con una resolución espacial de 30 a 1000 m.
MERIS	(Medium Resolution Imaging Spectrometer): Basicamente dedicado a observaciones del color del océano, sus objetivos se amplian con estudios atmosféricos y terrestres. Presenta una resolución espacial de 300 y 1200 m, en 15 bandas del espectro, programable en anchura y posición
RA-2	(Radar Altimeter 2): Permite el estudio de la topografía de la superficie de los océanos, tierras emergidas y casquetes polares, con una precisión que podría alcanzar los 5 cm
MWR	(Micro Wave Radiometer): Su objetivo principal es la medida del vapor de agua atmosférico, siendo sus datos además empleados en la corrección de la señal de RA-2 en su recorrido por la troposfera
LRR	(Laser Retro-Reflector): Instrumento activo que permite determinar con precisión la órbita del satélite
GOMOS	(Global Ozone Monitoring by Occultation of Stars): Su objetivo es el conocimiento de los procesos que ocurren en la alta atmósfera, permitiendo el estudio de la distribución del ozono durante largos periodos de tiempo a varias latitudes y altitudes, así como la medida de la distribución de aerosoles estratosféricos.
MIPAS	(Michelson Interferometer for Passive Atmospheric Sounding): Sonda atmosférica de gran resolución espectral para observaciones globales de elementos traza atmosféricos
AATSR	(Advanced Along Track Scanning Radiometer): Permite conocer, con 1 km de resolución espacial, la temperatura superficial del mar y de los continentes, el crecimiento de la vegetación, la humedad del suelo y la composición de las nubes
DORIS	(Doppler Orbitography and

	Radiopositioning Integrated by Satellite): Sistema de navegación que con una resolución espacial de 3 cm es clave para medidas de altimetría
SCIAMACHY	(Scanning Imaging Absorption Spectrometer for Atmospheric Cartography): Su objetivo principal se centra en la medida global de gases traza en la troposfera y estratosfera

Tabla 3.1 Instrumentos a bordo del ENVISAT

3.1.1 Características del sensor

El instrumento MERIS (Medium Resolution Imaging Spectrometer) es un espectrómetro programable de resolución espectral media que opera en el rango espectral correspondiente a la reflexión de la luz solar en la superficie terrestre.

A través de comandos enviados desde tierra es posible seleccionar quince bandas espectrales, cada una de ellas con un ancho y una localización programables dentro del rango que va desde los 390 a los 1040 nm.

El instrumento barre la superficie terrestre por el llamado método 'push broom'. Los arreglos de CCDs realizan un muestreo espacial en la dirección perpendicular al desplazamiento del satélite, mientras que el desplazamiento del propio satélite permite el muestreo longitudinal.

MERIS está diseñado para adquirir datos sobre la Tierra sólo cuando las condiciones de iluminación son las adecuadas. Los 68,5° de ángulo de barrido del instrumento alrededor del nadir permiten cubrir una franja de 1150km con una resolución espacial de 300 m en el nadir.

La misión MERIS: la principal misión de MERIS es la medida del color del mar tanto en los océanos como en las zonas costeras. El conocimiento del color del mar nos permite saber la concentración del pigmento clorofila, concentración de sedimentos suspendidos y cargas de aerosol sobre el medio marino. MERIS es además capaz de

medir la altura de las nubes, columna total de vapor de agua; así como, carga de aerosol sobre el suelo. Estas medidas constituyen la misión secundaria de MERIS.

Productos MERIS: a través de los datos obtenidos por MERIS se pueden procesar productos oceánicos, atmosféricos y terrestres:

- Radiancia/reflectancia emergentes del agua.
- Índices de pigmentos de algas I y II.
- Materia en suspensión.
- Substancia amarilla.
- Radiación fotosintéticamente activa (PAR).
- Espesor óptico del aerosol.
- Factor épsilon del aerosol.
- Espesor óptico de las nubes.
- Albedo de las nubes.
- Tipo de nube.
- Reflectancia de las nubes.
- Reflectancia terrestre.
- Espesor óptico del aerosol sobre tierra.
- Factor épsilon del aerosol sobre tierra.
- Factor global de vegetación MERIS.
- Columna de vapor de agua.

Datos archivados en el E-PAC: el E-PAC posee un archivo de datos MERIS desde mayo del 2002 en adelante.

Productos generados en el E-PAC, nivel de proceso y almacenamiento: en el E-PAC se procesan los datos de MERIS a partir del nivel 0 (L0) y se almacenan los datos tanto de nivel 0 como de nivel 1 (L1) y de nivel 2 (L2). Los productos geofísicos de nivel 2 generados en el E-PACs se indican a continuación:

- Reflectancia superficial normalizada.
 - Contenido de vapor de agua.
-
-

- Índices de pigmentos de algas I y II.
- Índice de vegetación en lo alto de la atmósfera (TOAVI) e índice de vegetación en la parte baja de la atmósfera (BOAVI).
- Presión en lo alto de las nubes.
- Substancia amarilla, materia en suspensión total o reflectancias rectificadas.
- Presión superficial.
- Radiación fotosintéticamente activa (PAR).
- Albedo de nubes.
- Tipo y espesor óptico de aerosol o nube.
- Flags.

Estos datos se procesan para cualquier área geográfica del planeta. Los usuarios sólo tienen que proporcionar las coordenadas geográficas de la zona requerida teniendo en cuenta que ésta se encuentre dentro del área del pase del satélite en la fecha y hora deseadas.

3.2 Satélites ERS 1 y 2

Los satélites ERS 1 y 2, fueron diseñados y puestos en órbita por la ESA (European Spatial Agency). Fueron los primeros en incorporar sensores SAR (Synthetic Apperture Radar) basados en microondas, con fines comerciales, ofreciendo nuevas oportunidades a las actividades de percepción remota. Las imágenes generadas a partir del radar, son independientes de las condiciones atmosféricas y de luz, mientras que el resto de los sensores, utilizan otro tipo de ondas aptas para aplicaciones que dependen de la atmósfera, velocidad del viento, etc.

ERS-1 fue lanzado en julio de 1991 en el cohete Ariane 4, lleva 2 radares especializados y un sensor de infrarrojo, y fue retirado del servicio el 10 de marzo de 2000, mientras que el ERS 2, fue puesto en funcionamiento en mayo de 1995 y aun sigue en operación. La ESA, puso en órbita en 2001 el satélite Envisat, que es la evolución natural del ERS 2, con un sistema SAR mejorado y un gran número de nuevos sensores. Los satélites ERS 1 y 2, se encuentran a 785 km, de distancia de la

superficie de la tierra y tienen una inclinación de $98^{\circ}52'$, con un periodo de revisita de 35 días. Sus sensores, tienen resoluciones que oscilan entre los 10 m y los 50 km y proporcionan tamaños de escena que varían entre los 5 y los 500 km. ERS-1 y ERS-2 son los satélites de observación de la Tierra más sofisticados que se hayan desarrollado en Europa. Un ERS pesa más de dos toneladas y mide unos 12 metros de largo.

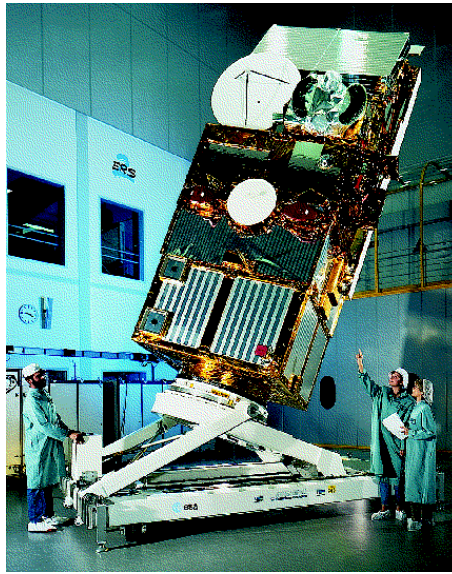


Figura 3.1 Satélite ERS-2

ERS-2 fue lanzado por la ESA en 1995. Es una copia de su predecesor con la única diferencia de que incluía un nuevo instrumento llamado GOME (Global Ozone Monitoring Equipment), diseñado para poder medir el ozono en la troposfera y estratosfera.

El instrumento de microondas activo - el mayor sistema de los que lleva a bordo el ERS - produce imágenes extremadamente detalladas de 100km cuadrados (4.5 millones de pixels por segundo, enlazados directamente a la tierra). El aparato dispone de un modo para medir continuamente la velocidad del viento así como su dirección, y tomar parámetros del oleaje en los océanos.

El radar altímetro (-segundo instrumento del ERS) mide con precisión la variación de altura del satélite. El tercer instrumento es un barredor que construye mapas detallados de la estructura térmica de los mares y océanos a partir de las medidas de temperatura de la superficie.

3.2.1 Plataforma

La plataforma desde la que se adquieren las imágenes es el satélite ERS-2, que pertenece a una misión que la ESA puso en marcha en 1990 para vigilancia de la Tierra desde el espacio. Está formada por los satélites ERS-1 (1990) y ERS-2 (1995). ERS-1 completó su misión en 1999, quedando en la actualidad sólo el ERS-2.

ERS-2 viaja en órbita cuasi-polar a una altura de 785 km, con una inclinación de 98.52°, y completa una órbita cada 100 min. (14 y 1/35 órbitas por día) de la que puede obtener una media de 10-12 mín de datos. Al ser una órbita cuasi polar, la Tierra rota gradualmente al paso del satélite dejando ver por completo su superficie en sólo tres días.

El período de revisita, tiempo que tarda el satélite en ver exactamente el mismo punto de la Tierra, es de 35 días. Las escenas (frames) ERS cubren un área de aproximadamente 100 km x 100 km y su resolución es de 30m, disponiendo de un solo ángulo de incidencia de 23°. Las imágenes procesadas pueden oscilar entre los 100 m de una imagen de baja resolución (LRI - Low Resolution Image) y los 12.5 m de una imagen de precisión corregida (PRI - Precision Resolution Image), pasando por los 20 m de la imagen con resolución completa (FRI - Full Resolution Image). ERS-2 es capaz de medir la atmósfera de la Tierra y las propiedades de su superficie con un alto grado de precisión, independientemente de la hora del día y las condiciones meteorológicas. A bordo lleva un variado número de sensores de microondas apoyados por una serie de instrumentos complementarios: El active microwave instrument (AMI) - que combina un SAR y un dispersómetro de viento, el radar altimeter (RA), el along-tracking scanning radiometer and microwave sounder (ATSR), el precise range and range-rate equipment (PRARE), el global ozone monitoring experiment (GOME) y el laser retroreflectors (LRR).

Uno de los objetivos primarios de esta misión es la vigilancia de los océanos, proporcionando datos esenciales para mejorar la representación de las interacciones océanos/atmósferas en los modelos climáticos, mejorar conocimientos sobre la

circulación oceánica y su transferencia de energía, realizar estimaciones más fiables sobre la masa de hielo en el Ártico y Antártico, y mejorar la monitorización de la polución y los procesos costeros dinámicos.

3.2.2 Adquisición de imágenes de radar

Los sistemas radar consisten en dispositivos de alta resolución que emiten una serie de pulsos en la longitud de onda de las microondas y reciben la respuesta reflejada por la cubierta terrestre. Debido a las propiedades electromagnéticas de estas ondas, son capaces de recibir datos en casi todas las condiciones atmosféricas, de día y de noche.

El pulso electromagnético enviado desde el satélite llega a la superficie terrestre y es reflejado. La dirección e intensidad del pulso reflejado depende de varias características del terreno, entre ellas su rugosidad. Así, una superficie rugosa refleja gran parte del pulso en la dirección del satélite, mientras que una superficie plana refleja especularmente la energía de manera que la parte que llega al sensor es prácticamente nula.

La detección de vertidos de hidrocarburos en el mar con imágenes radar se basa precisamente en el amortiguamiento que estos producen en la parte más superficial de las olas por impregnación capilar. Este amortiguamiento se traduce en una superficie más lisa en la zona de vertido, por lo que la respuesta que llega al sensor es menor. El resultado es que las zonas afectadas por vertidos de hidrocarburos aparecen en la imagen como zonas más oscuras [15].

3.2.3 Procesamiento de las imágenes

Las escenas ERS-2 adquiridas en la sala de recepción son transferidas a CREPAD donde se procesan y analizan mediante un software denominado SAROIL, que consiste en un procesador SAR que genera imágenes de alta precisión georreferenciadas. Dichas imágenes son analizadas por un operador que selecciona las áreas con sospecha de

vertido y genera un informe en formato html. Dicho informe es enviado a todas las organizaciones implicadas en el proyecto por correo electrónico.

A continuación se presentan 2 ejemplos de imágenes adquiridas por el satélite ERS-2 para constatar el poder del satélite al detectar derrames petroleros sobre el mar.

Satélite: ERS-2

Fecha: 18/10/2001

Localización: Oeste de la isla de Fuerteventura del Archipiélago de canarios a 100km del continente Africano

Observación: Posible vertido producido por limpieza de sentinas de un barco. La mancha negra indica que puede ser un vertido varios días anterior a la adquisición de la imagen.

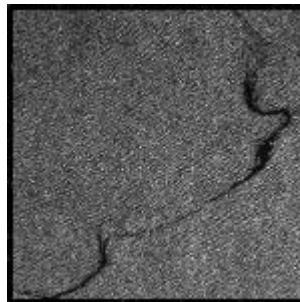


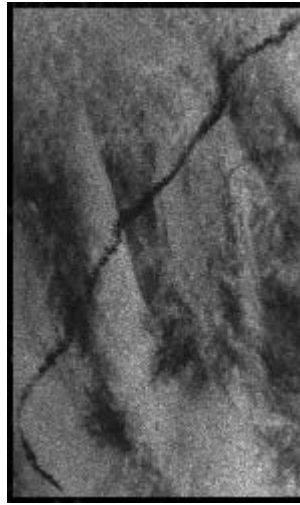
Figura 3.2 Imagen adquirida por el satélite ERS-2 oeste de la isla de Fuerteventura del Archipiélago de canarios a 100km del continente Africano

Satélite: ERS-2

Fecha: 31/10/2001

Localización: Noreste de las Islas Canarias a 100km del continente Africano

Observación: Posible vertido de petróleo. La forma alargada indica que posiblemente haya sido producido por limpieza de sentinas de un barco



*Figura 3.3 Imagen adquirida por el satélite ERS-2
noreste de las Islas Canarias*

3.3 Satélite LANDSAT

El programa LANDSAT fue desarrollado por la national aeronautics and space administration NASA y originalmente denominado earth resources technology satélite ERTS. Fue el primer programa de satélites de percepción remota para observación de los recursos terrestres, puesto en órbita terrestre. El primer satélite, de carácter experimental, fue construido para demostrar la viabilidad de mapear y monitorear rasgos de la superficie de la tierra a partir de imágenes orbitales. Este programa fue desarrollado con el objetivo de posibilitar la adquisición de imágenes de la superficie de la Tierra de manera global y repetitiva.

El primer satélite de esta serie fue lanzado el 23 de julio de 1972 con la denominación de ERTS-1. El 14 de enero de 1975 el nombre fue cambiado por LANDSAT-LANDSATellite y el 22 de enero fue lanzado el segundo satélite, entonces denominado LANDSAT-2

De la serie LANDSAT ya fueron lanzados los siguientes satélites:

Satélite	Lanzamiento	Fin operación	Sistemas sensores
ERTS-1	23/julio/1972	05/Enero/1978	MSS e RBV
LANDSAT-2	22/Enero/1975	27/Julio/1983	MSS e RBV
LANDSAT-3	05/Marzo/1978	07/Septiembre/1983	MSS e RBV
LANDSAT-4	16/Julio/1982	Fin de 1983	MSS e TM
LANDSAT-5	01/Marzo/1984	en operación	MSS e TM
LANDSAT-6	03/Octubre/1993	03/octubre/1993	ETM
LANDSAT-7	Abril/1999		ETM y HRMSI

Tabla 3.2 Satélites lanzados

Los tres primeros satélites LANDSAT tenían una fisonomía muy similar, con unas dimensiones aproximadas de 3 metros de altura por 1.5 metros de diámetro, que se convertían en 4 metros cuando se extendían los paneles solares. El peso total del sistema es de los 960Kg.

La USGS (U.S. Geological Survey's) a través de la EDC (EROS Data Center) ha administrado los archivos de datos del LANDSAT por más de dos décadas. Estos contienen una amplia información sobre la superficie terrestre.

A continuación se describen los satélites que componen la serie LANDSAT, dando una descripción de las características más importantes de los instrumentos sensores que los acompañan.

3.3.1 LANDSAT 1,2 y 3

Los tres primeros satélites de la serie LANDSAT fueron construidos a partir de una modificación del satélite meteorológico NIMBUS. Tenían una órbita circular, casi polar, sincrónica con el sol, a una altura aproximada de 920km. Durante su período de operación, los satélites LandSat realizaban una órbita completa alrededor de la Tierra cada 103minutos y 27 segundos, recubriendo 14 fajas de la superficie terrestre por día.

La configuración de estos tres satélites fue concebida de modo que cada 18 días ellos pasan sobre la misma región de la superficie de la Tierra. El horario medio de pasaje de los satélites por el Ecuador era a las 09:30, horario local.

3.3.2 Sistema RBV (Return Beam Vidicom)

Fue el primer sistema sensor proyectado para obtener imágenes a bordo del LANDSAT. Era un sistema constituido por cámaras de televisión (tres en los LANDSAT 1 y 2 y dos en el LANDSAT3) que observaba toda la escena a ser registrada en forma instantánea y operaba en una faja comprendida entre el visible y el infrarrojo próximo del espectro electromagnético. La escena terrestre registrada por la cámara representaba un área de 185km. La resolución espacial de este sistema en los LANDSAT 1 y 2 era de 80*80mts, con tres bandas espectrales (verde, rojo e infrarrojo próximo) y en el LANDSAT 3 era de 40x40mts, con una banda espectral (pancromático).

SATELITE	CAMARAS	BANDA ESPECTRAL	APLICACIONES
Landsat 1 y 2	1	.0475-.0575 (verde)	Diferenciación entre suelo y vegetación
Landsat 1 y 2	2	.0580-.0680 (rojo)	Mapeamiento de vegetación.
Landsat 3	3	0.690-0.830 (rojo)	Delineamiento de cuerpos de agua
Landsat 3	1	0.505-0.750 (pan)	Mapeamiento de áreas urbanas, vegetación, agricultura.

Tabla 3.2

Los dos primeros satélites Landsat incorporaron un sistema de 3 cámaras RBV, cada una de las cuales registraba información en una banda espectral comprendida entre el verde y el infrarrojo cercano.

Este sistema no funcionó adecuadamente en los dos primeros Landsat, por lo que fue sustituido en el tercero por otro. En este caso, el plan se dirigió más a mejorar la resolución espacial que la espectral, por lo que se montaron dos cámaras pancromáticas de alta resolución, operando simultáneamente con el MSS. Estas cámaras cubrían un abanico espectral de 0.505 a 0.705, proporcionando en cuatro adquisiciones, la misma cobertura del MSS pero con una resolución aproximada de 40m.

El registro entre ambas cámaras permitió utilizar el RBV como auxiliar del MSS, especialmente para mejorar la calidad visual de sus imágenes. Para ello se emplearon diversas técnicas de fusión, como la normalización de la bandas MSS a partir de la proporcionada por RBV o la transformación IHS.

3.3.3 Sistema MSS (Multispectral Scanner Subsystem - Generador de Imágenes Multiespectrales).

Este sistema fue colocado a bordo del LANDSAT por insistencia de los investigadores del departamento de agricultura de los Estados Unidos, que querían probar un sistema multiespectral para estudios agrícolas. Este sensor es un equipo de barrido óptico electrónico, que opera en cuatro bandas del espectro electromagnético, dos en el visible (bandas 4 y 5) y dos en el infrarrojo próximo (bandas 6 y 7). La franja terrestre observada por este sensor representaba un área de 185km y la resolución espacial era de 80x80m. Su campo total de visión es de 11.56 grados. En el LANDSAT 3 fue agregado una banda en el infrarrojo termico (banda 8), cuya resolución era de 240mts, que captaba la radiación emitida por los objetos terrestres [11].

3.3.4 LANDSAT 7

Landsat 7 fue lanzado en abril de 1999 con un nuevo sensor denominado ETM+ (Enhanced Thematic Mapper Plus). Su operación es administrada por la NASA (National Aeronautics and Space Administration), pero la producción y comercialización de imágenes depende de la USGS (United States Geological Survey).

Una imagen LANDSAT 7 ETM+ está compuesta por 8 bandas espectrales que pueden ser combinadas de distintas formas para obtener variadas composiciones de color u opciones de procesamiento. Entre las principales mejoras técnicas respecto de su antecesor, el satélite Landsat 5, se destaca la adición de una banda espectral (Banda Pancromática) con resolución de 15 metros. También cuenta con mejoras en las características geométricas y radio métricas y una mayor resolución espacial de la banda térmica para 60m. Estos avances tecnológicos permite calificar al LANDSAT 7 como el satélite más interesante para la generación de imágenes con aplicaciones directas hasta una escala de 1:100.000, principalmente en áreas rurales o territorios de grandes extensiones [11].

Las imágenes generadas por el Landsat7, adquiridas mediante el sensor ETM+ presentan una mejor relación costo-beneficio que las generadas y ofrecidas actualmente en el mercado por satélites de resolución media (15 a 30 metros) .

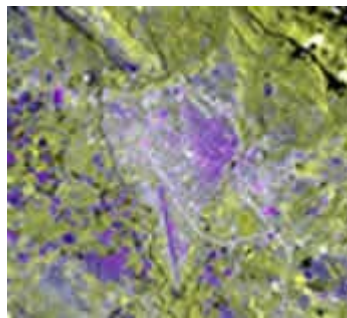


Figura 3.4 Ciudad de Tarapoto (Perú) tomada por el satélite LANDSAT



Figura 3.5 Valle de Chíncha (Perú) tomada por el satélite LANDSAT

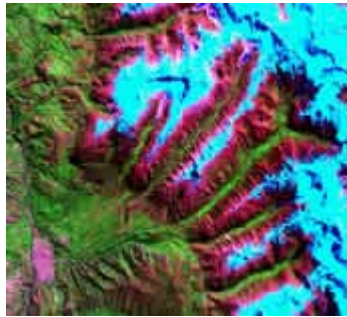


Figura 3.6 Huaraz y cordillera blanca (Perú) tomada por el satélite LANDSAT

3.4 Satélite SPOT

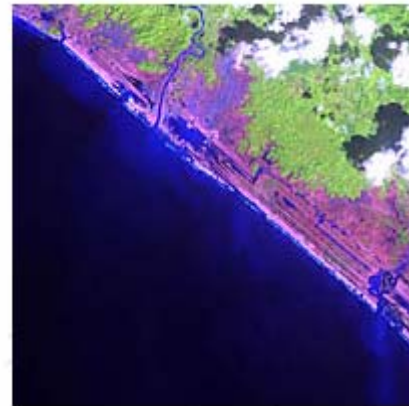
SPOT es el sistema de observación terrestre diseñado en Francia, bajo control del CNES (Centre National D'Études Spatiales - Francia), y desarrollado con la participación de Bélgica y Suecia. Desde 1986, el CNES ha lanzado 5 satélites SPOT dando una continuidad única al servicio de los clientes de SPOT IMAGE (empresa proveedora de imágenes satelitales).

Spot image es actualmente líder a nivel mundial en venta de imágenes satelitales y productos derivados de la observación satelital terrestre, con aplicación en agricultura, cartografía, catastro rural, estudios ambientales, planeamiento urbano, telecomunicaciones, forestería, geología, exploración de gas y petróleo, e ingeniería civil.

Los satélites Spot representan la referencia para la detección civil ya que permiten descubrir objetos de diez metros de diámetro, a pocos metros de distancia, con el último satélite de su flota, el Spot 5, en cada una de las imágenes de 60 kilómetros por 60 kilómetros. Las tomas de las vistas estereoscópicas a gran escala son una realidad con el Spot 5 con lo cual se puede obtener ciertas características muy buenas para procesamiento tales como retroceso, detalle, repetición y flexibilidad de empleo por una programación día a día: tantas características que hacen de estos satélites una herramienta útil próxima a las necesidades de los usuarios en las aplicaciones tan diversas como la agricultura, la cartografía, el urbanismo, la gestión de los bosques, los riesgos naturales, la geología y los océanos.



a) Imagen de diciembre 8 del 2004



b) Imagen de diciembre 29 del 2004

Figura 3.7 Fotografía tomada desde el satélite SPOT de la región de Aceh Sumatra Indonesia antes y después del tsunami

El sistema Spot está construido sobre dos operadores que fabrican sus diferentes satélites: el operador comercial y el del satélite. El papel del operador comercial, la compañía Spot image, es asegurar la programación comercial de los satélites, el tratamiento de las imágenes, la fabricación y la distribución de los productos. En este sentido, la compañía se apoya en el segmento Tierra del operador comercial (SSOC), que ella misma ha desarrollado.

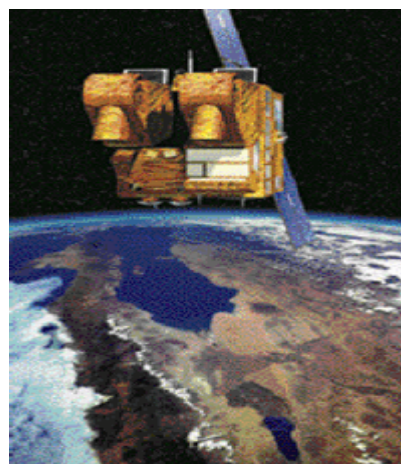


Figura 3.8 Satelite SPOT

El papel del operador de satélites, el CNES (propietaria de los cohetes Arianespace) es asegurar el control de los satélites en órbita, su mantenimiento y correcto funcionamiento del sistema. Para ello se apoya en el SSOC.

3.4.1 Tres satélites en órbita

El sistema Spot opera una constelación de tres satélites de observación: Spot 1, Spot 2 Spot 4, situados en una misma órbita más tarde en abril de 2004 se agregó a dicha constelación el Spot 5. Los dos nuevos instrumentos HRG, provenientes de los instrumentos HRVIR del SPOT 4, ofrecen una mejor resolución: 2,5 metros a 5 metros en modo pancromático y 10 metros en modo multibandas.

El funcionamiento de 2,5 metros es posible gracias a la utilización del nuevo concepto de toma de vista denominado “Supermode”. El Spot 5 embarca el nuevo instrumento de toma de vista HRS, funcionando en modo pancromático y que apunta, a la vez, hacia delante y hacia la parte trasera del satélite. HRS permite, de forma cuasi simultánea, la adquisición de parejas estereoscópicas y, así, restituye el relieve.

3.4.2 Capacidad de almacenamiento

La capacidad de almacenamiento de las imágenes a bordo se ha mejorado alcanzándose 550 imágenes diarias, en lugar de 400 del Spot 4, gracias a la utilización de una memoria de masa de 90 Gbits, asociada a una tasa de compresión de datos de 2,6 Gbits. Su gestión por ficheros ofrece una gran flexibilidad de uso y permite optimizar la programación. La velocidad de transmisión de datos de imagen hacia tierra es ahora el doble (2x50 Mbits/s). El modelo recurrente vegetación 2 y el instrumento doris están igualmente embarcados en el Spot 5. En fin, la plataforma Spot 5 ha derivado de la del Spot 4. Un nuevo sensor estelar ha sido introducido para incrementar el funcionamiento de localización de las imágenes (50 metros en lugar de 350 metros en los satélites anteriores).

3.4.3 La cámara HRG

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

3.4.4 La cámara HRS

El objetivo del instrumento HRS es adquirir las imágenes en la banda pancromática siguiendo dos ejes de vista, respectivamente, a +/- 20, en la parte delantera y trasera del satélite, en relación a su velocidad orbital. También permite adquirir rápidamente un par de imágenes estereoscópicas para elaborar, después del tratamiento en tierra, un modelo digital del terreno. Las cajas de detección delanteras y traseras con los captosres CCD asociados a su electrónica de proximidad permiten convertir la señal luminosa en una señal eléctrica.

El subsistema telemedida de la carga útil del satélite (TMCU) asegura las principales funciones:

- Compresión de datos de vídeo producidos por los instrumentos HRG y HRS.
- Formateo de los datos e inserción de datos auxiliares para facilitar el tratamiento de las imágenes en el satélite los cuales se mandan hacia la tierra.
- Almacenamiento de los datos en una memoria masiva de gran capacidad.
- Transmisión en banda X de los datos de imagen en directo o en diferido.

Cada compresor recibe los diferentes productos de imagen provenientes de los instrumentos HRG y HRS. Los flujos de vídeo son comprimidos por un algoritmo cuyos parámetros son adaptados a cada modo de toma de vista. El algoritmo ha sido seleccionado para obtener un buen funcionamiento con una tasa de compresión de 2,28 a 2,8, según el modo de toma de vista.

3.4.5 Memoria electrónica

El papel de la memoria de disco duro es almacenar las imágenes de los instrumentos HRG y HRS para, a continuación, releerlas cuando el satélite está visible desde una estación de recepción en tierra. Los componentes microelectrónicos y la digitalización han desplazado a las grabaciones magnéticas. Teniendo en cuenta la ausencia de mecanismos, la memoria eléctrica presenta numerosas ventajas en relación a las grabaciones magnéticas: gran simpleza y fiabilidad de funcionamiento y de gestión desde tierra.

La memoria del Spot 5 permite el almacenamiento a bordo de 550 imágenes, en lugar de 400 del Spot 4. En cuanto a la gestión por ficheros, ésta permite utilizar mejor la fuente de almacenamiento disponible.

3.4.6 Transmisión a tierra

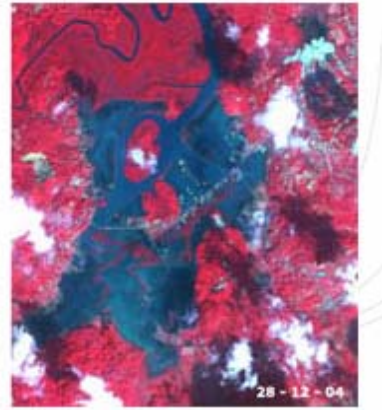
Esta memoria dispone de una gestión por archivos: los archivos grabados pueden ser releídos una o varias veces en las estaciones, en función de la demanda del usuario, independientemente del orden en que han sido almacenados. Las cámaras HRG, HRS y la memoria electrónica permiten transmitir los datos adquiridos hacia tierra. El tren de informaciones modula la señal de un emisor en banda X que, después de la amplificación, la envía directamente hacia una estación terrena a través de una antena omnidireccional a bordo. En relación a los otros satélites Spot, la velocidad de transmisión de los datos de imagen hacia tierra se ha multiplicado por dos (2x50 Mbits/s), conservando una modulación QPSK (modulación digital de cuatro

constelaciones). La capacidad de tratamiento de las imágenes (5 canales máximo) se ha superado de forma importante desde el Spot 4 (2 canales de tratamiento) y se ha introducido una gran flexibilidad porque los canales de tratamiento se manejan de forma independiente.

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



a) Imagen de febrero 25 del 2004



b) Imagen de Diciembre 28 del 2004

Figura 3.9 Imagen proveniente del satélite SPOT de la región de Andaman India antes y después del tsunami

3.4.7 Las bandas espectrales

La elección de las bandas espectrales de observación de la tierra es resultado de tres criterios: la especificidad de la señal recibida en la longitud de onda para analizar la naturaleza de los terrenos, vegetación, desierto, nieve, aguas, ciudades...etc, la

transparencia de la atmósfera y la estabilidad de la transmisión, que sólo es suficiente en algunas “ventanas” [12].

3.5 Satélite IKONOS

El Satélite IKONOS es el primer satélite de tipo comercial que posibilita la adquisición de imágenes con un metro de resolución espacial. El término “IKONOS” proviene del griego y significa “imagen”.

Estas imágenes son consecuencia directa de la liberación tecnológica promovida en 1994 por el gobierno de los Estados Unidos de América. Antes de este año esta tecnología estaba disponible para satélites con fines militares. IKONOS colecta información de cualquier área en promedio dos veces al día, cubriendo áreas de 20,000 km² en una misma pasada y produciendo como resultado imágenes de 1 metro de resolución cada tres días y de 2 m de resolución todos los días.

El satélite IKONOS pesa unos 720kg y órbita la Tierra cada 98 minutos a una altitud de casi 680km en forma sincronizada con el sol. El satélite fue diseñado y fabricado por la empresa *Lockheed martin commercial space system*. La empresa *Raytheon company* construyó los sistemas de apoyo terrestre, geoprocesamiento digital, manejo de archivos y servicio al cliente, requerido para distribuir la información captada por IKONOS.

El sensor digital de imágenes del satélite está diseñado para producir imágenes con elevado contraste, resolución espacial y precisión, brindando a los clientes un producto preciso y nítido. La compañía Estman Kodak proveyó el mecanismo electro-óptico, incluyendo su ensamblaje con el telescopio óptico, los detectores y su ajuste al plano focal.

Incluyendo, así mismo, el mecanismo de procesamiento electrónico de alta velocidad basado en un diseño efectuado por *space imaging*. Resultando todo un suceso, las imágenes de alta resolución del satélite IKONOS son una grata realidad y están revolucionando el mercado satelital.

3.5.1 Características del satélite IKONOS

- Lanzamiento del satélite: 24/09/99
- Lugar de lanzamiento: Vandenberg air force base, California USA
- Altitud: 681 km
- Inclinación: 98.1°
- Velocidad: 7km/seg
- Sentido de la órbita: descendente
- Duración de la órbita: 98 minutos
- Tipo de órbita: sincrónica con el sol
- Angulo de visada: rápida alternancia entre diferentes ángulos
- Tiempo de revista: 2.9 días
- Bandas espectrales:
 - Azul: 0.45- 0.52um
 - Verde: 0.52-0.60um
 - Rojo: 0.63 -0.69um
 - Infrarrojo próximo: 0.76 - 0.90um
- Rango dinámico: posibilita que la información sea almacenada en 11 bits por pixel, con lo cual redundante en un mayor rango dinámico que facilita el contraste y discriminación de la información. Este número de bits por pixel hace posible trabajar con paletas de 2048 tonos reales de gris. A priori hay quien puede pensar que no es significativo, ya que los sistemas de visualización permiten mostrar 256 niveles e incluso el ojo humano no es capaz de distinguir más. En resumen, se tiene 8 veces más información que en las imágenes tradicionales. Quizá para interpretación visual solo sea una buena ayuda, pero en estudios temáticos, toda esta información se convierte en una ventaja fundamental. No obstante, los productos pueden ser entregados al usuario en 8 bits por pixel.

La tabla 3.3 muestra información técnica respecto de sus bandas.

Banda	Banda inferior al 50% (nm)	Banda superior al 50% (nm)	Ancho de Banda (nm)	Central (nm)
Pan	525.8	928.5	403	727.1
MS-1 (azul)	444.7	516.0	71.3	480.3
MS-2 (verde)	506.4	595.0	88.6	550.7
MS-3 (rojo)	631.9	697.7	65.8	664.8
MS-4 (VNIR)	757.3	852.7	95.4	805.5

Tabla 3.3

3.5.2 Cobertura

El satélite IKONOS está equipado con un grabador interno que puede obtener datos de casi cualquier área de la superficie terrestre. Este grabador puede almacenar 64Gbps de datos (aproximadamente 26 imágenes completas de datos de ambos sensores, PAN y MS).

Existe un buen enlace de descarga de datos hecho por las estaciones fairbanks y tromsoc las cuales están en posiciones óptimas para el enlace gracias a su alta latitud. Además, una red de estaciones de recepción terrestre afiliadas a space imaging se han comenzado a construir para habilitar el enlace de la descarga de datos en muchas áreas. Actualmente, las estaciones completamente operativas son: Dubai (Space Imaging, medio este), Japón (Space Imaging, Japón), Korea (Space Imaging, Asia) y Ankara Turkey (Space Imaging, Euroasia).

3.5.3 Aplicaciones y ventajas de las imágenes IKONOS

Las imágenes IKONOS están revolucionando actualmente las formas de utilización de los productos satelitales. Por el hecho de ser IKONOS el primer satélite comercial que brinda un producto de 1m de resolución espacial, tenemos la posibilidad de definir los tipos de estándares para imágenes de alta resolución.

Actualmente es posible que:

- Los agricultores puedan monitorear con mayor precisión la condición y vitalidad de sus cultivos y predecir con mayor acierto sus volúmenes de cosecha; además, pueden prevenir problemas y ahorrar importantes sumas de dinero a través de su determinación en estados tempranos.
- Los científicos ambientalistas pueden predecir tendencias en áreas de elevada fragilidad ambiental.
- Los funcionarios de gobierno pueden monitorear, evaluar y planificar políticas de tipos específicos de utilización de la tierra.
- Los planeadores urbanísticos pueden evaluar los avances de planes comunales de viviendas y las compañías de seguros pueden medir y mapear daños a propiedades luego de desastres naturales.
- Los geólogos ya no necesitarán recurrir a vuelos fotográficos para interpretaciones geológico-estructurales. A partir de la posibilidad que ofrece IKONOS de generar productos con elevado nivel de resolución y recubrimiento estereoscópico, estas imágenes resultan comparativamente más ventajosas que las fotografías aéreas [10].

3.6 Satélite *TERRA/ASTER*

Dado el requerimiento de investigadores de recursos y de las geociencias en general, 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.

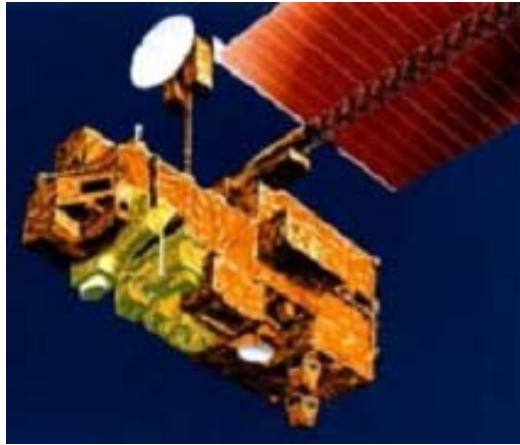


Figura 3.5 Satelite ASTER

ASTER es un sensor satelital de imágenes de alta resolución, transportado por el satélite TERRA, que fue 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 Analysis Center) para el desarrollo de sistemas de procesamiento y aplicaciones de los datos.

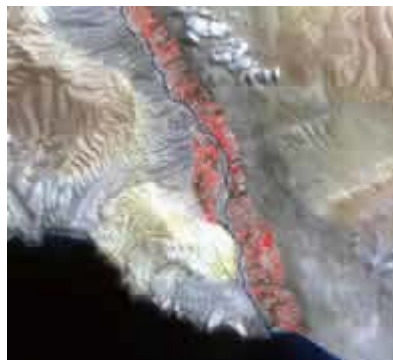


Figura 3.6 Valle de Ocoña – Arequipa (Perú)

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 metros de resolución; SWIR registra 6 bandas con 30 metros de resolución; y TIR tiene 5 bandas con 90 metros de resolución.

Además de alta precisión geométrica y radiométrica, es posible obtener pares estereoscópicos a lo largo de la trayectoria orbital. El ancho de barrido de ASTER es de 60km. 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° y $+24^\circ$ (UNIR) 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.

Consta de tres telescopios construidos en Japón, uno para cada zona espectral. Los mismos son apuntables en dirección transversal con desviaciones del nadir. Es un sistema de alta resolución espacial, espectral y radiométrica. Comprende 14 bandas espectrales, con un ancho de barrido de 60 km. No toma datos en forma continua sino durante 8 minutos por órbita. Dada su alta resolución y posibilidad de variar el ángulo de visión; el ASTER producirá imágenes estereoscópicas. Tiene variadas aplicaciones en geología y recursos no renovables, dada la elección de sus bandas espectrales [16].

IV. Análisis textural

4.1 Alternativas en el análisis textural

Textura es un término usado para caracterizar la rugosidad de la superficie de un objeto o fenómeno y es uno de los descriptores usados en el procesamiento de imágenes y el reconocimiento de patrones, ya que está relacionado con la composición de una sustancia cualquiera a través de variaciones en la superficie de los materiales, ésta es principalmente una propiedad que estudia las vecindades. La textura de una imagen es una cuantificación de la variación espacial de valores de tono que es imposible definirlo precisamente por su carácter sensorial. El uso de la textura de una imagen proviene de la habilidad innata de los humanos de reconocer diferencias texturales.

El nivel de gris de un pixel en una imagen monocromática representa un tono, mientras que un nivel de textura es la variabilidad del tono en una vecindad, es el patrón de relaciones espaciales entre los niveles de gris de los pixeles de un vecindario y puede ser descrito como suave o rugoso. El estudio de la textura depende de dos factores, la escala de variaciones, que puede ser local o regional, y la escala de observación.

El campo de estudio del análisis de texturas es muy amplio y no existe un método que cubra todos los factores de textura posibles, por esto, para realizar el análisis de detección de texturas, se han desarrollado tres modelos aplicados al procesamiento digital de imágenes: modelo estructural, estadístico y espectral.

- El modelo estadístico indica si una imagen es suave, burda, etc.
- El modelo estructural indica qué arreglos primitivos existen en la imagen, tales como regularidad de líneas paralelas, etc.
- El modelo espectral está basado sobre las propiedades de Fourier y es usado principalmente para detectar propiedades de periodicidad global en una imagen, identificando picos de una cierta energía en su espectro.

Estas técnicas pueden usarse por separado ó apoyarse unas a otras para detectar diferentes texturas.

4.2 Matrices de co-ocurrencia en niveles de gris (GLCM)

La matriz de co-ocurrencia describe la frecuencia de un nivel de gris que aparece en una relación espacial específica con otro valor de gris, dentro del área de una ventana determinada. La matriz de co-ocurrencia es un resumen de la forma en que los valores de los píxeles ocurren al lado de otro valor en una pequeña ventana.

El Método más comúnmente utilizado para medir matemáticamente la textura es la matriz de co-ocurrencia de niveles de gris o con sus siglas en inglés GLCM (Grey Level Co-occurrence Matrix), basadas en estadísticas de 2° orden. Es un histograma de los niveles de grises de dos dimensiones para un par de píxeles (píxel de referencia y vecino). Esta matriz aproxima la probabilidad de distribución conjunta de un par de píxeles.

Los pasos a seguir para obtener la matriz de co-ocurrencia son:

1. Reducir el número de niveles de gris de la imagen original.
2. Definir el tamaño de la ventana de análisis
3. Calcular las matrices de co-ocurrencia utilizando un cierto desplazamiento

En la figura 4.1 se representa la imagen prueba u original donde los valores corresponden a niveles de grises. La imagen tiene 4 píxeles de lado y 4 niveles de grises: 0,1,2 y3. Se aclara que en este caso se toman los valores de gris de la imagen en su totalidad para realizar los cálculos, dando como resultado solo un valor que representa a toda la imagen



Figura 4.1 Imagen prueba u original donde los valores corresponden a niveles de grises

4.2.1 Relación espacial entre dos píxeles

La Matriz de co-ocurrencia considera la relación espacial entre dos píxeles, llamados píxel de referencia y píxel vecino. Por ejemplo, si se escoge el píxel vecino que está situado un píxel a la derecha de cada píxel de referencia, esto se expresa como (1,0): 1 píxel en la dirección x, 0 píxel en la dirección y.

Cada píxel en la ventana se va convirtiendo sucesivamente en el píxel de referencia, empezando por el ubicado arriba a la izquierda y finalizando abajo a la derecha. Los píxeles ubicados en el margen derecho de la imagen original, no tienen vecino a la derecha por lo tanto no son usados en el cómputo.

Se pueden utilizar diferentes relaciones entre píxeles, por ejemplo:

(-1,0) un píxel a la izquierda del píxel de referencia

(1,1) un píxel a la derecha y un píxel abajo (en diagonal).

4.2.2 Distancia entre píxeles

En este documento se utiliza un píxel de separación (un píxel de referencia y su inmediato vecino). Cuando la ventana es suficientemente grande, se puede usar una separación mayor, sin que haya diferencias en la metodología de su cálculo.

Las posibles combinaciones de niveles de grises para la imagen de prueba se presentan en la Tabla 4.1, estas etiquetas no se volverán a mostrar en las matrices de co-ocurrencia.

Pixel Vecino Pixel de Referencia	0	1	2	3
0	(0,0)	(1,0)	(2,0)	(3,0)
1	(1,0)	(1,1)	(2,1)	(3,1)
2	(2,0)	(1,2)	(2,2)	(3,2)
3	(3,0)	(1,3)	(2,3)	(3,3)

Tabla 4.1 Todas las posibles combinaciones de los 4 niveles de gris de la imagen de prueba

La primera celda debe ser llenada con la cantidad de veces que ocurre la combinación 0,0. Cuantas veces en el área de la ventana de un pixel con valor de gris igual a 0 (pixel vecino), esta situado a la derecha de otro pixel con valor 0 (pixel de referencia)

Existen, por lo tanto diferentes matrices de co-ocurrencia para cada relación espacial, según se considere el vecino de arriba, al costado o en diagonal.

En la tabla 4.2 se muestra la matriz de co-ocurrencia para la relación espacial (1,0). Esta matriz se interpreta de la siguiente manera: En la imagen de prueba, dos veces el pixel de referencia es 0 y su vecino a la derecha es también 0 (primera celda). Dos veces el pixel de referencia es 0 y su vecino a la derecha es 1.

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

Tabla 4.2 Matriz de co-ocurrencia (1,0) para la imagen de prueba

4.2.3 Normalización de la matriz

En la matriz precedente, se cuentan cada pixel de referencia con su vecino a la derecha. Si el cálculo se realiza solo de este modo usando solo una dirección, entonces el número de veces que aparece la combinación 2,3 no es el mismo que la combinación 3,2 (por ejemplo el 3 esta a la derecha del 2 tres veces, pero a la izquierda solo una) por lo tanto la matriz no es simétrica respecto de la diagonal.

Sin embargo, la simetría es necesaria para el calculo, esto se logra si cada par de pixeles se cuentan dos veces: uan vez a la derecha y otra vez a la izquierda (se intercambian los pixeles de referncia y vecino en el segundo calculo).

Para obtener una matriz simétrica, la forma mas sencilla en vez de contar dos veces, es sumarle a esta matriz su matriz transpuesta. La matriz transpuesta se logra intercambiando las filas y columnas de la matriz original.

Sumando cada elemento de la matriz original y su transpuesta, se llega a la matriz simétrica de la tabla 4.3

4	2	1	0
2	4	0	0
1	0	6	1
0	0	1	2

Tabla 4.3 Matriz simétrica para una relación horizontal (derecha + izquierda) de la imagen prueba

Una vez obtenida la matriz simétrica, el paso siguiente es expresar esta matriz como probabilidad. La definición mas simple de la probabilidad es: "el número de veces que un evento ocurre, dividido por el numero total de posibles eventos" y la ecuación para su calculo es (1)

$$P_{i, j} = \frac{V_{i, j}}{\sum_{i, j=0}^{N-1} V_{i, j}}$$

Ecuación 4.1

Donde:

i es el numero de fila y j el numero de columnas

V es el valor de la celda i,j en la ventana

P_{i,j} es la probabilidad en la celda i,j

N es el numero de filas o columnas

Considerando la imagen de prueba de 4 x 4 pixeles, y la relación (1,0) el numero total de posibles pares es de 12, como muestra la figura 4.2, y para una relación horizontal (derecha mas izquierda) ese numero se duplica (24).

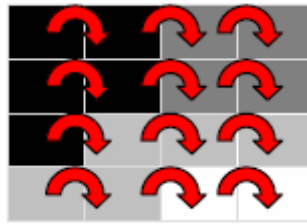


Figura 4.2 12 pares de pixeles para una relación a la izquierda (1,0). Considerando también la relación a la derecha (-1,0) para construir la matriz horizontal, los pares se duplican.

Observando la matriz horizontal de la tabla 4.3, vemos que, por ejemplo, la combinación 2,2 aparece 6 veces de las 24 posibles (12 a la derecha y 12 a la izquierda) y la combinación 2,3 solo 1 vez.

La combinación 2,2 ocurre 6 veces sobre 24 posibles, por lo que la probabilidad es de 1/4 o 0.250. Mientras que la combinación 2,3 es de 1/24 o 0.042.

La ecuación 4.1 transforma la matriz de co-ocurrencia en una aproximación de tabla de probabilidad. Decimos, que es una aproximación porque una verdadera probabilidad requiere de valores continuos y los valores de grises son valores enteros, por lo tanto discretos.

Este proceso se denomina Normalización de la matriz. Aplicando esta ecuación a la matriz simétrica de la Tabla 3 obtenemos la matriz de la Tabla 4, donde la sumatoria de todos los elementos debe ser igual a 1, pues esta normalizada

0.166 (4/24)	0.083 (2/24)	0.042 (1/24)	0 (0/24)
0.083	0.166	0	0
0.042	0	0.250	0.042
0	0	0.042	0.083

Tabla 4.4 Matriz normalizada horizontal de la matriz original

Se asume que toda la información está contenida en la matriz de dependencia espacial desarrolladas para las 4 direcciones de la figura 4.2. En general, cuanto mayor es el número de la diagonal en la matriz de co-ocurrencia, más homogénea es la textura en esa parte de la imagen que está siendo analizada.

Con respecto a la matriz de co-ocurrencia simétrica y normalizada hay algunos aspectos a resaltar:

1.- Los elementos de la diagonal representan pares de píxeles que no tienen diferencias en su nivel de gris. Si estos elementos tienen probabilidades grandes, entonces la imagen no muestra mucho contraste, la mayoría de los píxeles son idénticos a sus vecinos.

2.- Sumando los valores de la diagonal tenemos la probabilidad que un píxel tenga el mismo nivel de gris que su vecino.

3.- Las líneas paralelas a la diagonal separadas una celda, representan los pares de píxeles con una diferencia de 1 nivel de gris. De la misma manera sumando los elementos separados dos celdas de la diagonal, tenemos los pares de píxeles con dos valores de grises de diferencia. A medida que nos alejamos de la diagonal la diferencia entre niveles de grises es mayor.

4.- Sumando los valores de estas diagonales paralelas obtenemos la probabilidad que un píxel tenga 1,2,3,etc. niveles de grises de diferencia con su vecino.

4.2.4 Propiedades de la matriz

Cuadrada: El rango de los valores de los píxeles de referencia y el de los vecinos es el mismo, por lo tanto las filas y las columnas tienen idéntico número.

Es simétrica con respecto a la diagonal: Una matriz simétrica significa que los mismos valores ocurren en las celdas opuestas a la diagonal. Por ejemplo, el valor en la celda 3,2 debería ser el mismo que el valor en la celda 2,3 para que la matriz sea simétrica.

4.3 Algoritmo rápido para calcular GLCM

El método para obtener la matriz de co-ocurrencia toma en cuenta un barrido a través de toda la ventana de análisis para encontrar las probabilidades para formarla. También se puede ver que la matriz de co-ocurrencia resultante es una matriz simétrica. Por lo que para reducir el número de operaciones necesarias para obtener esta matriz, sólo se debe calcular la diagonal y la parte triangular superior o inferior de la matriz.[8].

Supóngase que tenemos una imagen de tamaño 512x512, con 256 niveles de gris. El tiempo de procesamiento utilizado normalmente es muy alto, ya que se tiene que llenar una GLCM de 256 x 256, por lo que generalmente se utilizan métodos para reducir los niveles de gris a 64, 32, etc.

Para reducir aún más el tiempo de cálculo de la matriz de co-ocurrencia lo que se hace es calcular la primera GLCM y después barrer la imagen completa en zig-zag Tabla 4.4, sumando o restando las nuevas probabilidades de aparición de los píxeles.

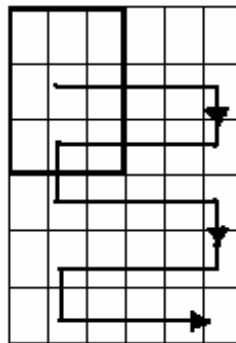


Figura 4.3 La imagen original es barrida en zig-zag

Ejemplo:

Supóngase que se quiere calcular la GLCM de la siguiente tabla, con 4 niveles de gris mostrada en la Tabla 4.5

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

Tabla 4.5 Ventana 5x5 para calcular GLCM

Si tomamos una ventana de 5 x 5, una distancia $d = 1$ y un ángulo $=0^\circ$, entonces tenemos que las primeras parejas son las mostradas en la Tabla 4.6

i	j		j	i
0	0		0	0
0	0		0	0
0	2		2	0
2	1		1	2
0	1		1	0
1	1		1	1
1	2		2	1
2	2		2	2
0	1		1	0
1	2		2	1
2	2		2	2
2	3		3	2
1	1		1	1
1	2		2	1
2	3		3	2
3	3		3	3
3	3		3	3
3	2		2	3
2	2		2	2
2	2		2	2

Tabla 4.6 Parejas de pixeles de la ventana de 5x5 de la imagen original

Acomodando estos pares en la GLCM se obtiene la matriz de 4 x 4 mostrada en la Tabla 4.7

i/j	0	1	2	3
-----	---	---	---	---

0	4	2	1	0
1	2	4	4	0
2	1	4	8	3
3	0	0	3	4

Tabla 4.7 Matriz de probabilidades

Por lo que la GLCM normalizada (con parámetro de normalización 1/40) será la mostrada en la Tabla 4.8

i/j	0	1	2	3
0	0.1	0.05	0.025	0
1	0.05	0.1	0.1	0
2	0.025	0.1	0.2	0.075
3	0	0	0.075	0.1

Tabla 4.8 GLCM normalizada de la ventana 5 x 5 original

Ahora, si movemos la ventana de análisis un pixel a la derecha queda la matriz mostrada en la Tabla 4.9

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

Tabla 4.9 Ventana 5 x 5 de la imagen original, recorrida un pixel a la derecha.

La Tabla 4.10 muestra que las parejas utilizando el método normal y el rápido serían las siguientes:

a) Método Normal

i	j		j	i.
0	0		0	0
0	2		2	0
2	1		1	2
1	1		1	1
1	1		1	1
1	2		2	1
2	2		2	2
2	2		2	2
1	2		2	1
2	2		2	2
2	3		3	2
3	2		2	3
1	2		2	1
2	3		3	2
3	3		3	3
3	2		2	3
3	2		2	3
2	2		2	2
2	2		2	2
2	3		3	2

b) Método Rápido

Restar					Sumar				
i	j		i	j	i	j		i	j
0	0		0	0	1	1		1	1
0	1		1	0	2	2		2	2
0	1		1	0	3	2		2	3
1	1		1	1	3	2		2	3
3	3		3	3	2	3		3	2

Tabla 4.10 métodos: a) normal y b) rápido.

La GLCM por el método rápido después de hacer las restas y sumas de las parejas de pixeles respectivamente es presentada en la Tabla 4.11

a) Resta					b) Suma					c) Normalización.				
i/j	0	1	2	3	i/j	0	1	2	3	i/j	0	1	2	3
0	2	0	1	0	0	2	0	1	0	0	0.05	0	0.025	0
1	0	2	4	0	1	0	4	4	0	1	0	0.1	0.1	0
2	1	4	8	3	2	1	4	10	6	2	0.025	0.1	0.25	0.15
3	0	0	3	2	3	0	0	6	2	3	0	0	0.15	0.05

Tabla 4.11 Restas, sumas y obtención de GLCM normalizada

Al llegar al final derecho de la matriz, como se muestra en la Tabla 4.12, se baja un pixel la ventana.

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

Tabla 4.12 Nueva matriz 5 x 5

Y se agregan y quitan las combinaciones listadas en la tabla 4.13 para obtener el resultado final

Restar				sumar			
i	j	i	j	i	j	i	j
2	1	1	2	3	3	3	3
1	1	1	1	3	1	1	3
1	1	1	1	1	1	1	1
1	3	3	1	1	1	1	1

Tabla 4.13 Pares a sumar y restar de la matriz GLCM anterior

De igual forma se continúa barriendo la imagen original hasta terminar con la última ventana de 5 x 5. Al calcular cada una de las matrices de co-ocurrencia, se utilizan los descriptores de textura propuestos originalmente por Haralick. Y así obtener la imagen resultado del descriptor correspondiente.

4.4 Descriptores de textura de Haralick

Al tener la matriz de co-ocurrencia, se utilizan los siguientes descriptores de textura.

Energía: Este descriptor esta dado por la ecuación siguiente:

$$E = \sum_i \sum_j P(i, j)^2$$

Ecuación 4.2

De esta ecuación se puede definir que cuando todos los valores de P(i,j) son semejante, el valor de la energía será menor, por el contrario si ocurre que en la diagonal principal se dan mayores picos de intensidad el descriptor se maximizará. Este descriptor da una idea de la suavidad de la textura, y esto se refleja en la ubicación de sus probabilidades en la matriz de co-ocurrencia.

Contraste: El contraste de una textura proporciona información acerca de las variaciones bruscas de color en la imagen. La extracción de este descriptor está dada por la siguiente expresión matemática.

$$C = \sum_i \sum_j (i - j)^2 P(i, j)$$

Ecuación 4.3

Como se puede observar en esta expresión, el valor del contraste aumentará si existen más elementos de la matriz de co-ocurrencia alejados de la diagonal principal. El efecto de los componentes de la diagonal es totalmente despreciado. Puesto que este se debe a los puntos cuyos pixeles son de colores iguales y como se busca una información acerca del contraste de la textura estos elementos no representan relevancia.

Correlación: Es la medida de la probabilidad que mide la relación entre las diferentes intensidades de los colores. Matemáticamente esta definida como:

$$Co = \sum_i \sum_j (i - \mu_x)(j - \mu_y) P_{ij} / \sigma_x \sigma_y$$

Ecuación 4.4

Debido al carácter estadístico que posee este descriptor, su valor está ligado al valor de sus medias estadísticas. El concepto de media y varianza se aplica al histograma de segundo orden, y la combinación en la media en i y en j, es referida a la diagonal principal de la matriz de co-ocurrencia.

El descriptor de correlación se verá aumentado cuando la distancia de los elementos a la media sea mayor.

Homogeneidad: Este descriptor proporciona información sobre la regularidad local de la textura. La expresión matemática para calcular este descriptor sobre la matriz de co-ocurrencia es la siguiente:

$$H = \sum_i \sum_j \frac{1}{1 + (i - j)^2} P(i, j)$$

Ecuación 4.5

Este descriptor aumentará su valor cuando la distancia i-j sea mínima. Lo que indica que mientras los elementos de la matriz de co-ocurrencia estén más próximos a la diagonal principal mayor será el valor de la homogeneidad.

Varianza

$$v = \sum_i \sum_j (i - \mu)^2 P(i, j)$$

Ecuación 4.6

Entropía: Este descriptor es una medida de la aleatoriedad contenida en la matriz de co-ocurrencia. La obtención de éste descriptor esta dada por la siguiente fórmula:

$$Ent = - \sum_{i=1}^N \sum_{j=1}^N P(i, j) \log P(i, j)$$

Ecuación 4.7

Donde $P(i,j)$ es cada elemento de la matriz, e i,j varían desde 0 hasta el número de niveles de gris. A medida que todos los elementos de la matriz son similares, este descriptor aumenta su valor, siendo máximo en el caso que todos los elementos de la matriz fueran iguales.

Autocorrelación

$$Ac = \sum_i \sum_j (i * j) P(i, j)$$

Ecuación 4.8

Disimilaridad: Este descriptor también conocido como momento diferencial de orden k , con $k=1$, provee información dependiendo del parámetro k . La tendencia de este descriptor en general, es aumentar de valor cuando los elementos de la matriz de co-ocurrencia se encuentran alejados de la diagonal principal.

$$Dis = \sum_i \sum_j |i - j| * P(i, j)$$

Ecuación 4.9

Cluster Shade: Este descriptor mide el grado en que los pixeles se inclinan a favor de un lado u otro de la media estadística. Su expresión matemática es la siguiente:

$$Sg = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 P(i, j)$$

Ecuación 4.10

Cluster Prominence

$$Pg = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 P(i, j)$$

Ecuación 4.11

Máxima Probabilidad: Este descriptor suministra información sobre la respuesta más fuerte de un par de pixeles en la matriz de co-ocurrencia. La expresión matemática de este descriptor es la siguiente:

$$Max = \max_{i,j} P(i, j).$$

Ecuación 4.12

El valor de este descriptor será mayor cuando en la matriz de co-ocurrencia se observe un pico de intensidad, el mayor de estos elementos dará el valor máximo de probabilidad.

Donde μ_x y μ_y son los valores de la media, y σ_x y σ_y son las desviaciones estándar definidas a continuación:

$$\mu_x = \sum_i \sum_j i * P(i, j)$$

$$\mu_y = \sum_i \sum_j j * P(i, j)$$

$$\sigma_x = \sum_i \sum_j (i - \mu_x)^2 * P(i, j)$$

$$\sigma_y = \sum_i \sum_j (j - \mu_y)^2 * P(i, j)$$

4.5 Representaciones multibanda

Configuración de un sistema de reconocimiento de patrones: Un esquema modular de un sistema de reconocimiento de patrones se muestra en la Figura 4.1. Este esquema no debe verse como los pasos a seguir en la construcción de un sistema de reconocimiento de patrones, sino más bien desde un punto de vista funcional: la entrada es un patrón natural y el resultado es una etiqueta. Además, no debe entenderse que todos los sistemas de reconocimiento de patrones deben incorporar todas estas unidades, ni siquiera que éstas deban estar tan claramente separadas. Se entiende que los componentes ya están diseñados y operativos. A continuación se profundizará en los diferentes módulos funcionales de este sistema general.

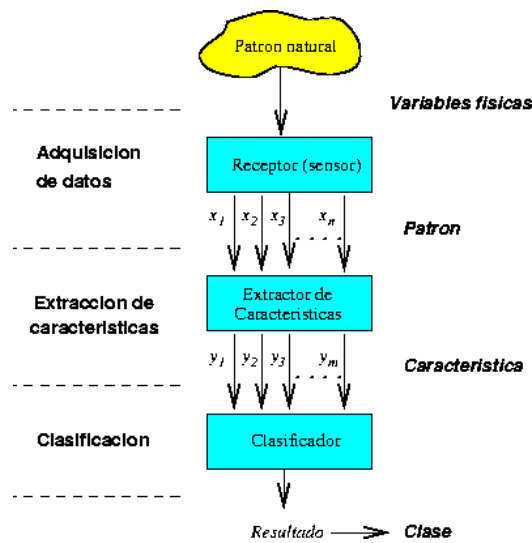


Figura 4.1 Etapas de un sistema de reconocimiento de patrones

Adquisición de datos. Representación de los patrones

La entrada a un sistema de reconocimiento estadístico de patrones es un vector numérico que contiene los valores muestreados y cuantificados (o binarizados) de una serie de señales naturales.

De una manera más formal, suponiendo patrones n -dimensionales, un **patrón** X es una variable aleatoria n -dimensional compuesta por n componentes, x_1, x_2, \dots, x_n , tales que x_i pertenece al conjunto de niveles de gris G_i para $i = 1, 2, \dots, n$.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Ejemplo 1

En el modelo de color RGB cada color aparece descompuesto en sus componentes espectrales primarias: rojo (R), verde (G) y azul (B). El modelo está basado en un sistema de coordenadas cartesiano. El subespacio de interés es el cubo que se muestra en la Figura 4.2

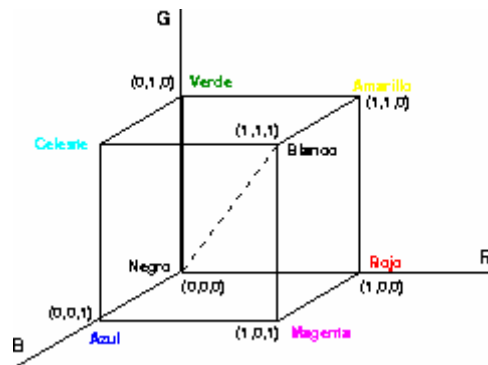


Figura 4.2 Cubo de color RGB. La diagonal principal indica los niveles de gris, desde el negro en el origen $(0, 0, 0)$ al blanco en el punto $(1,1,1)$.

Por conveniencia se supone que todos los colores han sido normalizados y escalados; es decir, se supone que todos los valores de rojo, verde y azul están en el rango $\{0-255\}$. Los puntos de la diagonal pueden representarse “visualmente” por 256 niveles de gris, escalados desde el negro (nivel de gris 0 ó $(0,0,0)$ en RGB) al blanco (nivel de gris 255 ó $(255, 255, 255)$ en RGB).

Las imágenes en el modelo de color RGB están formadas por tres planos de imágenes independientes, cada una de ellas asociada a cada color primario. La Figura 4.3 muestra una imagen en color de verduras y frutas junto con sus tres bandas de color rojo, verde y azul. Obsérvese que el mayor nivel de gris en cada banda corresponde al color predominante en cada objeto.

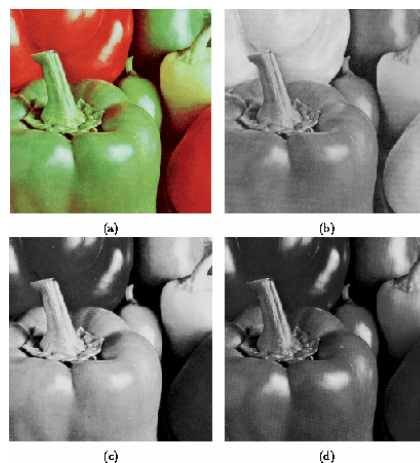


Figura 4.3 a) Imagen original a color. b) Plano de rojo. c) Plano de verde y d) plano de azul.

En una imagen digital monobanda (Figura 4.3 b,c y d) un pixel tiene asociado un único valor (nivel de gris) mientras que en una imagen multibanda un pixel se representa mediante un vector. Si suponemos que los planos de color son individuales y cada patrón es un pixel de la imagen, entonces cada patrón puede representarse por un vector dimensional ($n = 3$) ($[x_1, x_2, x_3]$).

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ donde } \begin{cases} x_1 \text{ es el nivel de gris en el plano R} \\ x_2 \text{ es el nivel de gris en el plano G} \\ x_3 \text{ es el nivel de gris en el plano B} \end{cases}$$

En este caso, por la notación introducida anteriormente, $G_i = G = \{0, 1, \dots, 255\}$ donde $i = 1, 2, 3$. Así, un patrón de rojo intenso podría tener asociado el vector $[250, 5, 2]^T$ mientras que otro de verde intenso podría tener asociado el vector $[0, 255, 2]^T$.

Cuando son introducidas en un monitor RGB, las tres imágenes se combinan en la pantalla de fósforo para producir una imagen de color compuesta. Por tanto, el uso del modelo RGB para el procesamiento de imágenes tiene sentido cuando las imágenes expresadas en términos de los tres planos de colores. Alternativamente, la mayoría de las cámaras en color que se usan para adquirir imágenes digitales utilizan el formato RGB.

Ejemplo 2

Una imagen de teledetección es una imagen multibanda de una misma escena tomada sobre la superficie terrestre. Se dice multibanda porque simultáneamente se registran varias imágenes de la misma escena, cada una de ellas con un sensor diferente. Estos sensores pueden estar montados sobre un satélite artificial orbitando sobre la Tierra (teledetección espacial) o simplemente sobre un avión. Al igual que ocurre con una imagen RGB, un pixel de una imagen multibanda de teledetección se representa mediante un vector (Figura 4.4).

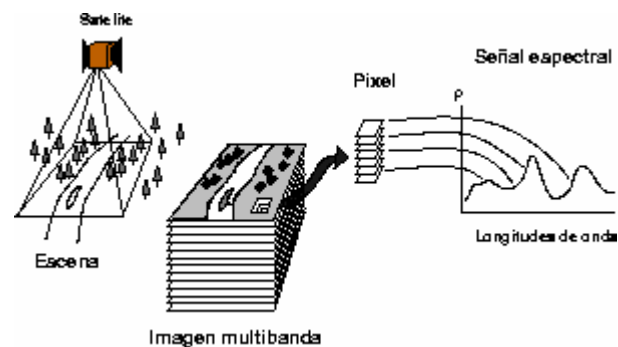


Figura 4.4 Variabilidad de los valores de una clase

El fundamento físico de la teledetección pasiva es la medición de la energía reflejada por los objetos situados en la superficie terrestre. El Sol emite radiación electromagnética (con diferentes longitudes de onda) que incide sobre estos objetos. Por el principio de conservación de la energía, parte de esta energía se absorbe, parte se transmite y parte se refleja. La proporción de energía reflejada se llama **índice de reflectividad** y se denota con el símbolo ρ

En una determinada longitud de onda λ , la cantidad de energía reflejada depende del tipo de material que estemos considerando y para un mismo material, la cantidad de energía reflejada depende de la longitud de onda considerada (Figura 4.5). Para cada material se conoce su **curva de reflectividad** (curva λ - ρ), en ocasiones se les han llamado *la huella digital de la naturaleza*.

De lo que disponemos es de un conjunto de n imágenes de la misma escena y cada una de éstas es una imagen digital de niveles de gris ($G_i = G = \{0, 1, \dots, 255\}$ para $i = 1, 2, \dots, n$ que representan valores muestreados, donde el nivel de gris 255 indica máxima reflectividad y el nivel de gris 0 indica mínima reflectividad. Así, por ejemplo, una imagen que corresponde a la captada por el sensor sensible al rango de longitudes de onda del espectro visible ($0.4\mu\text{-}0.7\mu$) mostrará las zonas de hielo (alta reflectividad en el espectro visible) con niveles de gris muy altos, cercanos a 255 (blanco) mientras que las zonas de agua (baja reflectividad en esta zona del espectro) las mostrará con niveles de gris muy bajos, cercanos a 0 (negro). Por lo tanto, cada pixel puede representarse mediante un vector con n componentes (Figura 4.5), en el que cada componente x_i indica el valor muestreado y digitalizado de la reflectividad (relativa) registrada por el

sensor correspondiente a la banda i . En la Figura 4.5 mostramos dos imágenes multibanda, captadas por los sensores MSS y TM que proporcionan imágenes de dimensionalidad 4 y 7, respectivamente.

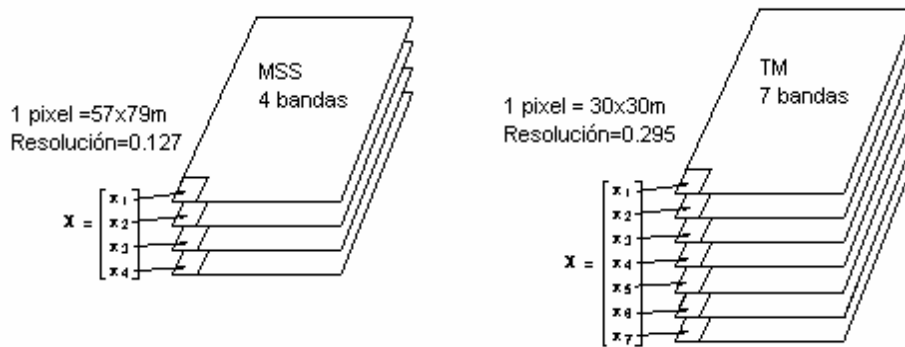


Figura 4.5 De píxeles a patrones

Una imagen MSS ó TM puede visualizarse en “falso color” asignando tres de sus bandas a cada uno de los planos de color RGB. Esta composición de color puede ser útil para conseguir una visualización “realista” de una imagen de teledetección a partir de tres bandas de niveles de gris.

Espacio de representación.

Un patrón también no es más que un punto en el *espacio de representación de los patrones*, es un espacio de dimensionalidad determinada por el número de variables consideradas. Esta aproximación concluye que los patrones pertenecientes a una misma clase están cercanos entre sí en el espacio de representación mientras que aquellos que pertenezcan a clases diferentes deben estar en diferentes regiones del espacio de representación cercanos a sus correspondientes clases, tal y como se muestra en la Figura 4.6.



Figura 4.6 Ejemplo de patrones en el espacio de representación

Un patrón se representa como un punto en el espacio de patrones P . El *espacio de patrones* P es un espacio de dimensionalidad determinada por el número de variables consideradas y se define como el conjunto de todos los valores posibles que puede tomar el patrón X , esto es,

$$P = \times_{n=1}^i G_i$$

donde x denota el producto cartesiano.

En el caso de imágenes de teledetección, valores típicos de x_i (función de la reflectividad espectral) están en el conjunto $G_i = G = \{0, 1, \dots, 255\}$ para $i = 1, 2, \dots, n$.

Similaridad entre patrones.

La tarea fundamental de un sistema de reconocimiento de patrones (clasificador) es la de asignar a cada patrón de entrada una etiqueta. Dos patrones diferentes deberían asignarse a una misma clase si son *similares* y a clases diferentes si no lo son. La cuestión que se plantea ahora es la definición de una medida de similaridad entre patrones.

Supongamos un sistema de adquisición perfecto (sin ruido). Podemos asegurar que:

1. La adquisición repetida del mismo patrón debería proporcionar la misma representación en el espacio de patrones. Por ejemplo, una cámara debería proporcionar siempre la misma imagen de la misma escena si las condiciones externas no cambian.

2. Dos patrones diferentes deberían proporcionar dos representaciones diferentes.
3. Una ligera distorsión aplicada sobre un patrón debería proporcionar una pequeña distorsión de su representación.

Se supone que el proceso de adquisición es biunívoco y continuo. Estas consideraciones sugieren que si las representaciones de dos patrones están muy cercanas en el espacio de representación, entonces los patrones deben tener un alto grado de similaridad. No obstante, no puede afirmarse tajantemente que a mayor distancia mayor disimilaridad ya que la medida (absoluta) de distancia depende de la escala en la que se cuantifiquen las variables asociadas al patrón.

Variabilidad entre patrones.

La suposición de un sistema de adquisición perfecto no deja de ser eso, una suposición. Los sistemas de adquisición introducen cierta distorsión o ruido, lo que produce una *variabilidad* en la representación de los patrones. Aunque es posible controlar eficientemente en muchos casos esta distorsión mediante el calibrado de los sistemas de adquisición aparece otra fuente de variabilidad por la propia naturaleza de los patrones.

Con mucha frecuencia, patrones de una misma clase difieren, incluso significativamente. Un ejemplo sencillo es el de los sistemas de reconocimiento de caracteres OCR que pueden interpretar diversos tipos de letra, incluso caracteres escritos (bajo fuertes restricciones), en este caso la variabilidad de los patrones se debe a factores tales como el granulado, color y calidad del papel o el tipo de tinta empleado.

Ejemplo 3

Las dos imágenes mostradas en la Figura 4.7 forman parte de una imagen multibanda captada por d sensores acoplados a un avión. El sensor A, sensible a la radiación electromagnética en el rango *visible* (película fotográfica en color, por ejemplo), proporciona la imagen mostrada en la Figura 4.7A. El sensor B, sensible a la radiación

electromagnética en el rango *infrarrojo*, proporciona la imagen mostrada en la Figura 4.7B.

Nos centraremos en dos tipos de objetos: los correspondientes a *zonas de vegetación*, por una parte, y los correspondientes a *zonas asfaltadas*, por otra.

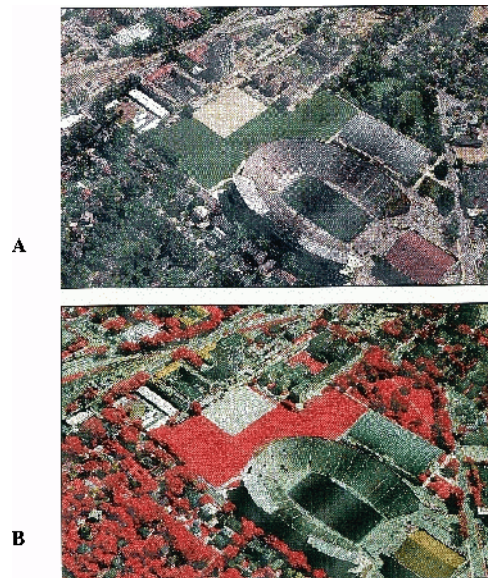


Figura 4.7 A) Imagen tomada por un sensor sensible a la banda del visible. B) Imagen tomada por un sensor sensible a la banda del infrarrojo

Resulta evidente que existe cierta variabilidad entre puntos que pertenecen a la misma clase. Observar que: 1) el prado que está por encima del fondo cerrado del estadio presenta diversas tonalidades de verde, 2) lo mismo ocurre con los árboles, además de alternar verdes con negros (sombras), 3) las zonas asfaltadas no son homogéneas. En definitiva, no todos los patrones de una clase son iguales: la incidencia de la luz solar (sombras), la distribución no uniforme de la vegetación, la presencia de objetos que reflejan la luz sobre objetos y distorsionan su color real, etc., hace que las imágenes captadas no sean uniformes.

Las curvas de reflectividad (ver ejemplo anterior) para las clases de interés se muestran en la Figura 4.7A. De éstas se deduce la variabilidad que encontramos en los datos reales. Si representamos cada patrón (cada pixel) en el espacio bidimensional formado por los valores registrados en las dos bandas, el resultado será el mostrado en la Figura 4.7B.

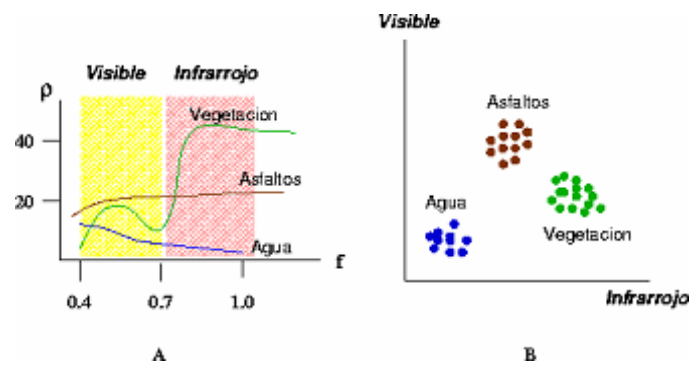


Figura 4.8 A) Curvas de reflectividad. B) Representación espectral.

La variabilidad *intrínseca* de los patrones hace que las representaciones tengan la forma de *nubes de puntos* en lugar de puntos individuales. Esta última representación es cuando no existe variabilidad entre los patrones. Estas agrupaciones de puntos sugieren que los patrones de una misma clase (similares) se representan cercanos (relativamente) en el espacio de representación, mientras que patrones de diferentes clases (diferentes) se representan lejanos (relativamente) en ese espacio. Si no fuera así, nos encontraríamos con imágenes como la mostrada en la Figura 4.9, donde mostramos el prado y una zona asfaltada completamente uniformes, algo absolutamente irreal.

Finalmente, para demostrar la capacidad de discriminación cuando se consideran diferentes bandas (atributos), observe con detenimiento las imágenes mostradas en la Figura 4.7 y en la curva de reflectividad de la vegetación (Figura 4.8A). Las zonas de vegetación aparecen más claras (mayor reflectividad) en la zona del infrarrojo y destacan en la Figura 4.7 B. Sin embargo, observe que el terreno de juego no aparece de la misma tonalidad que los prados o el resto de vegetación la explicación es simple: se trata de césped artificial o sea una textura sintética.



Figura 4.9 Clases con patrones homogéneos

Selección y extracción de características

El problema a resolver es el de extraer la información *relevante* para la clasificación a partir de la suministrada por los sensores (datos en bruto). De forma general, este problema puede plantearse como sigue. Dado un conjunto de patrones n -dimensionales

$$X = [x_1, x_2, \dots, x_n]^T$$

se trata de obtener un nuevo conjunto (características) d -dimensionales

$$Y = [y_1, y_2, \dots, y_d]^T$$

donde $d \leq n$.

Este objetivo puede abordarse de dos formas:

1. **Reduciendo la dimensionalidad de los datos.** Si los patrones son de alta dimensionalidad, el costo computacional asociado a la clasificación puede ser muy alto. Muchos clasificadores están basados en cálculos de distancias y estos cálculos pueden depender de forma cuadrática respecto a la dimensionalidad de los patrones. Otra consideración computacional hay que tomar en cuenta el espacio de almacenamiento adicional que supone guardar los valores de nuevas variables. Además, algunas de las variables pueden ser redundantes.

Las técnicas dedicadas a *seleccionar* las variables más relevantes son llamadas *selección de características* y reducen la dimensionalidad de los patrones. Este proceso puede esquematizarse como se indica en la Figura 4.10A, en el que un módulo selector recibe patrones n -dimensionales (en el ejemplo, $n = 100$) y proporciona como resultado las d variables más significativas (en el ejemplo, $d = 3$) de acuerdo a algún criterio de optimización. En este caso $d < n$ y el conjunto de las d variables seleccionadas es un subconjunto del conjunto original de variables.

2. **Cambiando el espacio de representación.** El objetivo es obtener una nueva representación de los patrones en la que los agrupamientos aparezcan separados si son de diferente clase y que haya un agrupamiento por clase. Esto puede conseguirse aplicando alguna transformación sobre los datos originales. Estas transformaciones

suelen ser transformaciones *lineales* y el objetivo suele ser minimizar la varianza (la varianza nos dice que tan separados esta los datos unos de otros).

Estas técnicas reciben el nombre de *extracción de características* y producen un nuevo conjunto de variables. Este proceso puede esquematizarse como se indica en la Figura 4.10B, en el que un módulo extractor recibe patrones n -dimensionales (en el ejemplo, $n = 100$) y proporciona como resultado nuevos patrones n -dimensionales de acuerdo a algún criterio de optimización. Es posible que las nuevas variables estén implícitamente ordenadas, por lo que proporcionan, adicionalmente un procedimiento de selección. En este caso $d = n$ y las variables seleccionadas *no* forman un subconjunto del conjunto original de variables.

Los vectores resultantes del proceso de selección o extracción se denominan *características*, aunque nosotros utilizaremos el término patrón a no ser que queramos resaltar explícitamente que se ha producido un proceso de selección/extracción de características.

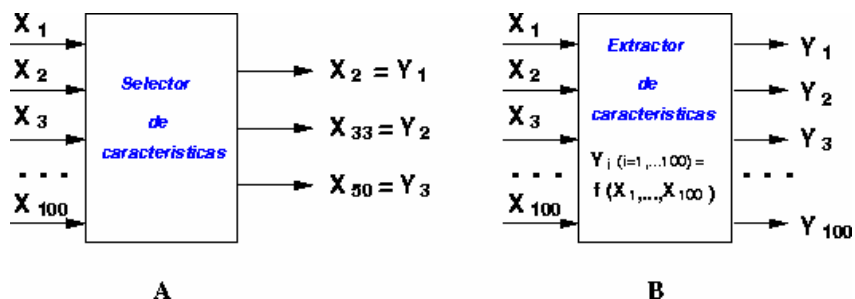


Figura 4.10 A) Selección de las 3 variables más significativas entre un conjunto de 100. B) Transformación de las 100 variables originales

Módulo de clasificación: El objetivo final de un sistema de reconocimiento de patrones es el etiquetar de forma automática patrones que desconocemos su clase. Suponemos que el sistema dispone de un módulo de adquisición de datos y se han seleccionado previamente las variables más significativas.

El conjunto de clases: En primer lugar debe establecerse claramente el objetivo final del sistema: ¿cuál es el conjunto de etiquetas? o dicho de otra forma ¿qué salidas debe proporcionar el sistema?

Si suponemos que todos los patrones a reconocer son elementos potenciales de J clases distintas denotadas ω_j $j = 1, 2, \dots, J$, llamaremos

$$\Omega = \{ \omega_1, \omega_2, \dots, \omega_J \}$$

al *conjunto de las clases informacionales*. Conviene tener en cuenta que una clase informacional es la denominación que se da a una clase conocida y con significado. Por ejemplo, un conjunto de clases informacionales sensato para el problema expuesto en el ejemplo sería:

$$\Omega = \{ \text{agua, vegetación, suelo} \}$$

Usualmente se habla simplemente de clase prescindiendo de la coletilla “informacional” aunque esta distinción es muy importante a la hora de abordar las técnicas de agrupamiento.

Resulta conveniente ampliar el conjunto Ω incorporando una nueva clase, llamada la *clase de rechazo*. Así, se define la clase de rechazo (ω_0) como una clase que se asigna a todos los patrones para los que no se tiene una certeza *aceptable* de ser clasificados correctamente en alguna de las clases de Ω . Se dice que

$$\Omega^* = \{ \omega_1, \omega_2, \dots, \omega_J, \omega_0 \}$$

es el *conjunto extendido de clases informacionales*.

El clasificador: Una vez establecido el conjunto de clases se procede a la construcción del clasificador. Como puede intuirse, la construcción del clasificador no es una tarea trivial ni directa e involucra una serie de etapas:

1. La elección del modelo.
2. Aprendizaje (entrenamiento del clasificador).
3. Verificación de los resultados.

Es muy importante señalar que estas etapas no deben verse de forma secuencial. Puede ocurrir que en un momento dado hay que volver atrás para replantearse alguno de los pasos dados, incluso el conjunto de clases informacionales.

Un *clasificador* o regla de clasificación es una función *d-dimensional*: el espacio de patrones P definida sobre los patrones X .

Ejemplo 4

Supongamos el problema de clasificación mostrado en la Figura 4.11, en el cual se trata de clasificar una imagen multiespectral de cuatro bandas. La imagen tiene 256 filas y 256 columnas por lo que disponemos de un conjunto de 65536 patrones para clasificar de dimensionalidad 4. Se ha establecido que el conjunto de clases informacionales (Ω) consta de 5 clases y una vez construido (de alguna forma) el clasificador, el resultado es un conjunto de 65536 etiquetas, una por cada patrón a clasificar.

En el caso de imágenes digitales el resultado de la clasificación es más apreciable: puede presentarse en forma de imagen de etiquetas, asignando un color a cada una. Además, si se conserva la misma estructura espacial en la imagen de etiquetas que en la imagen multibanda de entrada, el resultado se puede interpretar visualmente, lo que introduce la posibilidad de realizar una valoración *subjetiva* de la clasificación.

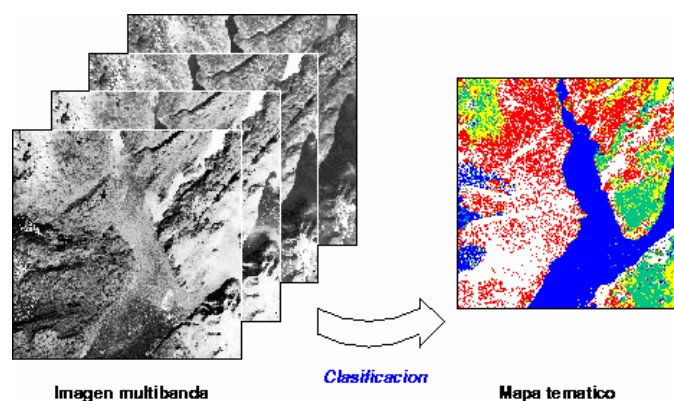


Figura 4.11 Clasificación: de la imagen multibanda al mapa temático

V Campos aleatorios de Markov

5.1 Nociones topológicas

Los modelos de campos aleatorios de Markov (CAM) son atractivos por las siguientes razones:

- Hacen más fácil el cálculo de medidas de ejecución cuantitativas para caracterizar cómo trabajan los algoritmos de análisis de imagen.
- Pueden ser usados para incorporar información contextual a priori, o bien restricciones, de manera cuantitativa.
- Los algoritmos basados en CAM tienden a ser locales y conducen a una implementación hardware paralelo de manera natural.

Los modelos de visión estocásticos completos basados en campos aleatorios de Markov están formulados dentro del campo de trabajo Bayesiano. La solución óptima de un problema se define como el estimador de probabilidad máxima a posteriori (MAP), el mejor que se puede obtener a partir de observaciones aleatorias. La probabilidad a posteriori se define usando la regla de Bayes, a partir de un modelo a priori y un modelo de probabilidad. El modelo de probabilidad está relacionado con 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 CAM nos proporcionan herramientas para codificar restricciones contextuales en la probabilidad a priori. Esta es la principal razón para el modelo de visión usando CAM.

Del uso común que los distintos autores hacen de los CAM, la implementación de un modelo de campos aleatorios de Markov podemos resumirla en 5 pasos:

1. Plantear el problema de visión como uno de etiquetado en el que la configuración de etiquetas representa una posible solución.

2. Después plantearlo como un problema de etiquetado Bayesiano en el que la solución óptima se define como la configuración de etiquetas MAP.
3. Usar la distribución de Gibbs para caracterizar la distribución a priori de las configuraciones del etiquetado.
4. Definir la función de energía de la distribución de Gibbs.
5. Usar un algoritmo (por ejemplo, simulated annealing) para minimizar la función de energía de la distribución de Gibbs y obtener así la configuración de etiquetas MAP.

El problema de segmentación de una imagen puede considerarse de una forma más simple como un problema de etiquetado. El problema de etiquetado consiste en asignar una etiqueta del conjunto de etiquetas a cada estado, y su solución consiste en la asignación de un conjunto de etiquetas a los píxeles (sites) de la imagen. Sea S un conjunto discreto de valores con m píxeles:

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

Ecuación 5.1

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

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

Ecuación 5.2

El conjunto de etiquetas L puede ser un conjunto discreto o continuo. En el caso continuo el conjunto de etiquetas puede corresponder a \mathbb{R} (conjunto de números reales) o a un intervalo de \mathbb{R} .

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

Ecuación 5.3

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

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

Ecuación 5.4

El conjunto de etiquetas puede ser caracterizado en términos de su continuidad. Además de la continuidad, otra característica esencial de un conjunto de etiquetas es su orden. En un conjunto discreto como, por ejemplo, la intensidad del tono de gris (0,1,...,255), podemos ordenar los elementos según los valores de intensidad [0,1,2,...,255].

En un conjunto de etiquetas ordenado podemos definir una medida numérica (cuantitativa) de semejanza entre 2 etiquetas. Para un conjunto de etiquetas sin relación de orden, la medida de semejanza es simbólica (cualitativa), normalmente tomando un valor de igual o no igual. Ordenar las etiquetas no solamente es importante para clasificar problemas de etiquetado, sino que afecta también a la elección de algoritmos de etiquetado y por lo tanto a la complejidad computacional.

Se llamará etiquetado a:

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

Ecuación 5.5

Donde a cada píxel de S se le asigna una única etiqueta de L con la función mapping o de mapeo:

$$f : S \rightarrow L$$

Ecuación 5.6

En términos de campos aleatorios de Markov, se llama configuración al etiquetado. De esta forma, una configuración de etiquetas puede corresponder a una imagen, a un mapa de bordes de la imagen, a una interpretación de una imagen con características de objeto, a una interpretación de los objetos que aparecen en una escena, etc.

Cuando todos los píxeles tiene el mismo conjunto de etiquetas L, el conjunto de todas las configuraciones posibles, es el producto cartesiano:

$$\mathfrak{S} = \underbrace{L \times L \times \dots \times L}_{m \text{ veces}} = L^m$$

Ecuación 5.7

Donde m es el tamaño de S

En términos de homogeneidad y de continuidad, podemos clasificar un problema de etiquetado como uno correspondiente a una de las cuatro categorías siguientes:

PE1: píxeles homogéneos con conjunto de etiquetas continuo.

PE2: píxeles homogéneos con conjunto de etiquetas discreto.

PE3: píxeles no homogéneos con conjunto de etiquetas continuo.

PE4: píxeles no homogéneos con conjunto de etiquetas discreto.

Las dos primeras categorías caracterizan el procesado de bajo nivel realizado sobre las imágenes observadas y las 2 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 píxeles S corresponde a los píxeles de la imagen y el conjunto de etiquetas L a un nivel 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 en la restauración continua, se intenta estimar la señal de la imagen real. La diferencia es que cada píxel de la imagen resultante aquí asume un valor discreto y, por lo tanto, L es conjunto discreto de etiquetas.

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

ser por ejemplo, un tono de gris, un color o una textura. Los píxeles dentro de cada región tienen asignadas una etiqueta única. La suposición principal en estos problemas es que la señal es suave o está formada por trozos suaves. Esto es complementario a la suposición sobre los bordes, en los que pueden ocurrir cambios bruscos.

La detección de bordes es también un PE2. A cada píxel (más concretamente, entre cada par de píxeles vecinos) se le asigna una etiqueta entre (borde y no-borde) si existen diferencias significativas entre las propiedades del par de píxeles vecinos.

El agrupamiento perceptual es un PE3. Los píxeles corresponden a características segmentadas inicialmente (puntos, líneas y regiones) que están ordenadas no homogéneamente. Las características fragmentarias se deben organizar en grupos con significado perceptual.

El reconocimiento y emparejamiento de objetos basados en características es un PE3. Cada píxel indica una característica de la imagen tal como un punto, un segmento de la línea o región. Las etiquetas son un conjunto discreto por naturaleza y cada una de ellas indica una característica del modelo. La configuración resultante es un “mapping” a partir de las características de la imagen, llegando a aquellos del objeto modelo.

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

Para un problema de etiquetado discreto con m píxeles y M etiquetas, existen M^m configuraciones posibles. Para un problema de etiquetado continuo, el número de configuraciones posibles es infinito.

5.2 Sistema de vecinos y tipos de cliques asociados

Los campos aleatorios de Markov son una parte de la teoría de la probabilidad que proporciona una herramienta para analizar dependencias espaciales o contextuales de fenómenos físicos.

Los píxeles de S están relacionados mediante sistema de vecinos, definido por:

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

Ecuación 5.8

Donde N_i es el conjunto de píxeles vecinos a i para los que se cumple que:

1.- Un píxel no es vecino a si mismo

$$i \notin N_i$$

2.- La relación de vecindad es mutua entre el píxel j y el píxel i

$$i \in N_j \Leftrightarrow j \in N_i$$

Para una rejilla (lattice) regular S , el conjunto de vecinos del píxel i está definido como el conjunto de píxeles próximos dentro de un radio r .

$$N_i = \{j \in S \mid [dist(pixel_j, pixel_i)]^2 \leq r, j \neq i\}$$

Ecuación 5.9

Donde $dist(pixel_j, pixel_i)$ es la distancia euclídea entre el píxel j y el píxel i y r es un número entero.

Al sistema de vecinos de primer orden, se le llama sistema de 4 vecinos, en cada píxel (interior) tiene cuatro píxeles vecinos, como se muestra en la Figura 5.1(a). El sistema de vecinos de segundo orden, llamado sistema de 8 vecinos, hay ocho vecinos para cada píxel (interior) (ver Figura 5.1(b)). Se podría seguir mostrando sistemas de vecinos de orden n , pero los más conocidos y usados son los tres órdenes mostrados en la Figura 5.1.

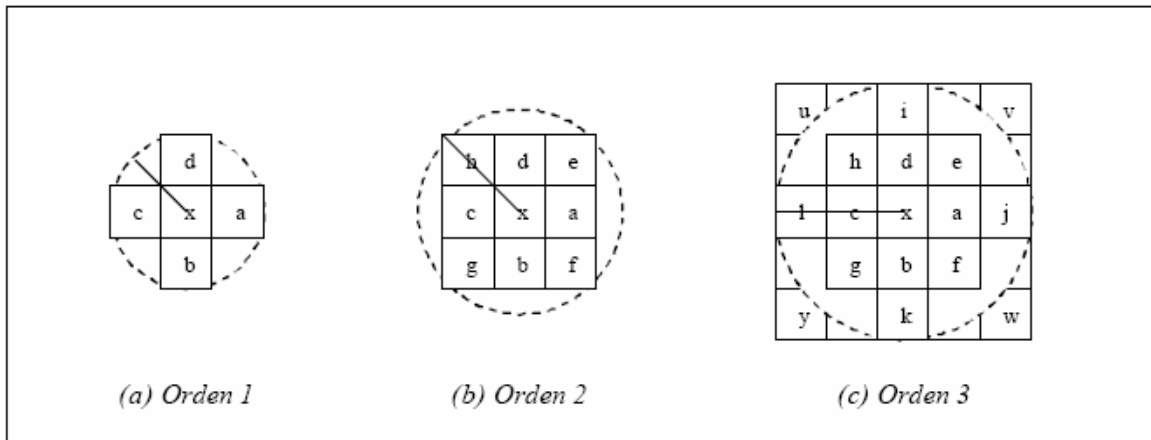


Figura 5.1 Vecinos de rejilla regular S

Teniendo como base estas definiciones se introduce el concepto de conexidad, que constituye una relación de vecindad espacial entre un píxel y sus vecinos próximos; es decir, un píxel i y su vecino i' tienen 4-conexidad si i' pertenece al conjunto de N_i de primer orden, y dos píxeles i e i' tienen 8-conexidad si i' pertenece al conjunto de N_i de segundo orden y así sucesivamente.

Cuando se ha especificado un orden entre los elementos de S , el conjunto de vecinos se puede determinar más fácilmente. Por ejemplo, sea $S = \{1, 2, \dots, m\}$ un conjunto ordenado de píxeles cuyos elementos son los índices de los píxeles de una imagen 1D, entonces un píxel interior $i \in \{2, \dots, m-1\}$ tiene 2 vecinos $N_i = \{i-1, i+1\}$ y los píxeles de la frontera (los 2 extremos) tienen un vecino cada uno, $N_1 = \{2\}$ y $N_m = \{m-1\}$. El conjunto de píxeles de una rejilla regular rectangular $S = \{(i, j) / 1 < i, j < n\}$ corresponde a los píxeles de una imagen en 2 dimensiones de tamaño $n \times n$. Cada píxel interior, representado por (i, j) , tiene 4 vecinos $N_{i,j} = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}$, los píxeles de la frontera tienen 3 y los píxeles de las esquinas tienen 2.

Para un conjunto irregular de píxeles S , el conjunto de vecinos de i , N_i , se define como:

$$N_i = \left\{ j \in S \mid \left[dist(característica_j, característica_i) \right]^2 \leq r, j \neq i \right\}$$

Ecuación 5.10

corresponde a estados próximos en un radio r , N es el estado que se representa por un píxel en el conjunto de vecinos de i , N_i . La función $\text{dist}(\text{característica}_j, \text{característica}_i)$ necesita ser definida apropiadamente para puntos que no tienen características. En general el conjunto de vecinos N_i para un conjunto irregular de estados S tiene dimensión y tamaño variables.

Un clique C de (S, N) se define como un subconjunto de píxeles en S , que pueden estar constituidos por un solo píxel, dos píxeles, tres píxeles o más, representandose por C_1, C_2, C_3, \dots , respectivamente,

$$\begin{aligned}
 C_1 &= \{i \mid i \in S\} \\
 C_2 &= \{(i, j) \mid j \in N_i, i \in S\} \\
 C_3 &= \{(i, j, k) \mid i, j, k \in S \text{ son vecinos dos a dos}\} \\
 &\vdots
 \end{aligned}$$

Ecuación 5.11

El conjunto de todos los cliques para $\{S, N\}$, de una rejilla regular, está determinado por su tamaño, dimensión y orientación. La figura 5.2 muestra distintos tipos de cliques para un sistema de vecinos de primer y segundo orden. A medida que aumenta el orden del sistema de vecinos, el número de cliques crece rápidamente y también el costo computacional asociado.

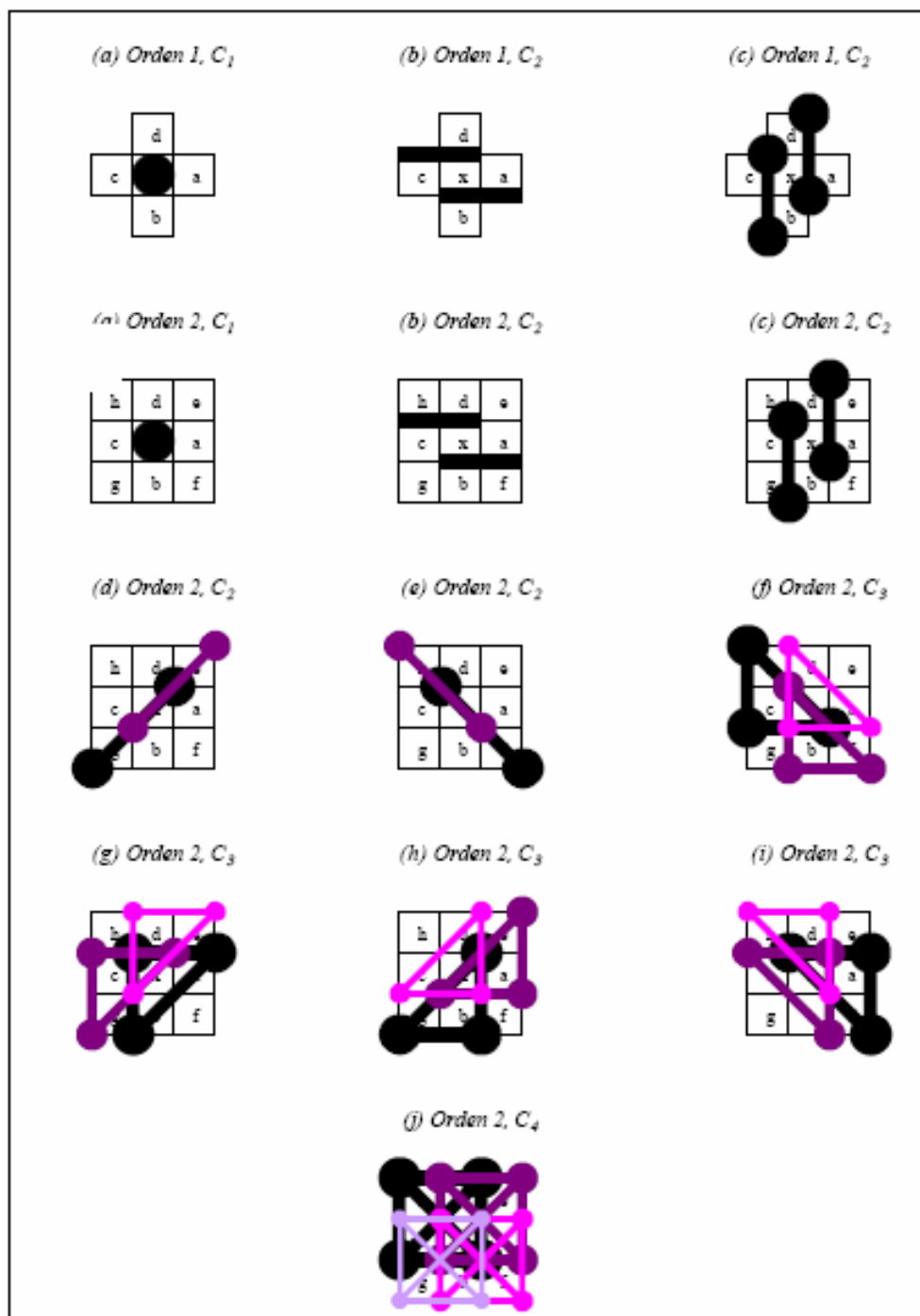


Figura 5.2 Cliques en una rejilla regular S

5.3 Campos Aleatorios de Markov y Distribución de Gibbs

5.3.1 Campos aleatorios de Markov (CAM)

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 . Se llamará campo aleatorio a la familia de variables aleatorias F . Se usará la notación $F_i = f_i$ para denotar que el suceso F_i toma el valor f_i .

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

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

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

Una familia de variables aleatorias F se dice que es un campo aleatorio de Markov sobre S , con respecto al sistema de vecindad \mathbf{N} , si y solo si se cumple que:

$$1. \quad P(F = f) > 0 \quad \forall f \in \mathfrak{F}$$

Ecuación 5.12

$$2. \quad P(F_i = f_i \mid F_j = f_j, j \in d, j \neq i) = P(F_i = f_i \mid F_j = f_j, j \in N_i)$$

Ecuación 5.13 (Propiedad Markoviana)

La condición 1 es para exigir que F sea un campo aleatorio, y es llamada positividad. La condición 2 es llamada característica local. Se dice que la probabilidad de un suceso local a L condicionado a todos los sucesos restantes es equivalente a aquella condicionada a los sucesos vecinos de L .

Finalmente, se define un Campo Aleatorio de Markov discreto sobre una trama finita como una colección de variables aleatorias, que corresponden a píxeles de una rejilla, 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 píxeles, que son los vecinos de ese píxel determinado.

5.3.2 Distribución de Gibbs

Diremos que un conjunto de variables aleatorias F es un campo aleatorio de Gibbs en S , con respecto a N (sistema de vecinos), si y solo si sus configuraciones siguen una distribución de Gibbs. Una distribución de Gibbs tiene la siguiente función de densidad:

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

Ecuación 5.14

Donde,

$U(f)$ = Función de energía

Z = constante de normalización

T = temperatura

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

Ecuación 5.15

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

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

Ecuación 5.16

Se considera entonces que $V_c(f)$ es el potencial del clique, y que $U(f)$ es la suma de todos los potenciales de clique, $V_c(f)$ para todos los cliques posibles C . El valor de $V_c(f)$ depende de

la configuración local del clique c , considerándose a la distribución Gaussiana como un caso especial de la familia de distribuciones de Gibbs.

Se dice que un campo aleatorio de Gibbs es homogéneo si $V_c(f)$ es independiente de la posición relativa del clique C en S , y se dice que es isotrópico si la región es independiente de la dirección. Es más simple especificar una distribución de campo aleatorio de Gibbs si es homogéneo o 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, o patrón f . Las configuraciones más probables son aquellas con energías más bajas. La temperatura T controla la forma de distribución. Cuando la temperatura es alta, todas las configuraciones tienden a estar igualmente distribuidas; pero cuando la temperatura es cercana a cero, la distribución se concentra alrededor del mínimo global de energía. Teniendo T y $U(f)$, se puede generar una “clase de patrones” muestreando el espacio de configuraciones de acuerdo con $P(f)$. Para problemas de etiquetado discretos, la función clique potencial se puede especificar por un número de parámetros y la configuración local tiene un número finito de píxeles, por lo tanto un número finito de valores. En el caso de etiquetado continuo la configuración puede variar continuamente.

Por conveniencia, la energía de una distribución de Gibbs se puede expresar como la suma de varios términos correspondientes a cliques de distintos tamaños, tal como se indica a continuación:

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

Ecuación 5.17

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

Un caso importante es cuando se consideran los cliques de tamaño dos como máximo, entonces

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

Ecuación 5.18

Si se considera que $\{i,j\}$ y $\{j,i\}$ son dos cliques distintos en C_2 (ya que los estados del clique están ordenados), la probabilidad condicional puede escribirse como:

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

Ecuación 5.19

Un campo aleatorio de Markov se caracteriza por sus propiedades locales (Markovianas) mientras que un campo aleatorio de Gibbs se caracteriza por sus propiedades globales (distribución de probabilidad Gibbs). El teorema de Hammersley-Clifford[7] establece la equivalencia entre estas dos propiedades: F es un campo aleatorio de Markov en S con respecto a N . Sin embargo, la prueba de que un campo aleatorio de Markov es un campo aleatorio de Gibbs es mucho más complicada y un resultado sobre la representación única de campos aleatorios de Gibbs proporciona tal prueba.

5.4 Modelos de campos aleatorios de Markov

5.4.1 Auto Modelos

En este modelo se usan dos tipos de estados (trabaja en 2 estados), es por ello que también se conoce como modelo de Ising, debido a que se basa en el fenómeno del mismo nombre.

Las restricciones de contexto entre dos etiquetas son un primer nivel a tener en cuenta en la información del contexto. Son muy utilizadas debido a su simplicidad y bajo costo computacional. Se codifican en la función de energía de Gibbs como funciones clique potenciales para cada par de píxeles, como se muestra:

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

Ecuación 5.20

A esta función se le llama energía de segundo orden porque se calcula para cada clique formado por dos estados. Se puede especificar un campo aleatorio de Gibbs o un campo aleatorio de Markov mediante la selección apropiada de V_1 y V_2 .

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

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

Ecuación 5.21

El modelo anterior recibe el nombre de auto-modelo y se puede considerar como un modelo auto-logístico si las f_i toman valores en el conjunto de etiquetas discreto $L = \{0,1\}$ o $L = \{+1,-1\}$. La correspondiente función de energía es de la forma:

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

Ecuación 5.22

Donde $\beta_{i,j}$ puede ser interpretado como el coeficiente de interacción. Cuando el sistema de vecinos que se considera es el de 4-vecinos en una rejilla 2D, entonces el modelo auto-logístico se reduce al modelo de Ising y la probabilidad condicional para este modelo es:

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

Ecuación 5.23

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

5.4.2 Modelo de Potts

En el caso de imágenes de niveles de gris, es posible modelar las interacciones entre los píxeles a efecto de segmentar en varias clases. La finalidad del modelo de Potts es asociar a cada píxel una etiqueta de un conjunto finito. Denotemos por X la imagen de datos originales y por Y la imagen segmentada, el proceso de segmentación consiste en maximizar la probabilidad condicional $P(X|Y)$, que de la aproximación de la regla de Bayes se tiene (Apéndice A):

$$P(Y|X) \doteq P(X|Y) P(Y)$$

Ecuación 5.24

$P(Y)$ está referido como el modelo a priori mientras que $P(X|Y)$ está referido como el término de manejo de datos. En estudios prácticos, el modelado probabilístico se basa en el análisis de energía de los cliques. Esto produce funciones de energía relacionadas a la llamada función potencial. El primer término probabilístico (el modelo a priori) está a menudo compuesto de un modelo de Potts. La función de energía asociada con este modelo es la siguiente:

$$U(Y) = - \sum_{\{i,j\} \in C} \beta \delta_{y_i y_j}$$

Ecuación 5.25

Los cliques tienen dos elementos, definiendo interacciones estrechamente pares. El coeficiente β define la probabilidad de homogeneidad de la solución, esto es, el mayor de

esos términos, el más probable de los pixeles adyacentes que pueden tener la misma etiqueta.

Los términos de manejo de datos están definidos por las funciones de costo dependiendo de la etiqueta l y denotada por f^l . Los parámetros inducidos son directamente extraídos de los datos. El potencial asociado a los cliques de primer orden se define como:

$$U(X|Y) = \sum_{c=\{i\}} \sum_l f^l(x^i) \delta_{y_i} = 1$$

Ecuación 5.26

Donde x_i e y_i son los datos y los valores de las etiquetas respectivamente en el sitio i

El problema de segmentación consiste en minimizar la energía global:

$$U(X|Y) + U(Y)$$

Ecuación 5.27

En la estructura del MAP, el proceso de minimización está desarrollado por una técnica estocástica como el Recocido Simulado (convergiendo próximamente con un mínimo global). Remarcando la minimización puede incluso ser desarrollada por técnicas determinísticas como los Modos de Iteración Condicional (rápido, pero converge solo aproximándose al mínimo local)

5.4.3 Modelo logístico multinivel (MLL)

El modelo auto-logístico se puede generalizar al modelo logístico multinivel (MLL, por sus siglas en inglés, MultiLevel Logistic), también llamado proceso de Strauss y modelo generalizado de Ising. En este modelo existen $M > 2$ etiquetas en el conjunto de etiquetas $L = \{1, \dots, M\}$. El clique potencial depende del tipo c (tamaño, dimensión y orientación posibles) del clique y de la configuración local. Para los cliques que contienen más de un píxel se definen las funciones clique potenciales MLL como:

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

Ecuación 5.28

Donde ζ_c es el potencial para los cliques de tipo c ; para los cliques de un solo estado depende de la etiqueta asignada al píxel:

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

Ecuación 5.29

Donde α_I es el potencial para la etiqueta I . La figura 5.3 muestra los tipos de cliques y parámetros asociados en el sistema de 8-vecinos (segundo orden).

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

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

Ecuación 5.30

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 (Figura 5.3) y cuatro diferentes β_c . Cuando el modelo es isotrópico los cuatro toman el mismo valor. Debido a su simplicidad, el modelo descrito se usa extensamente para modelar regiones y texturas.

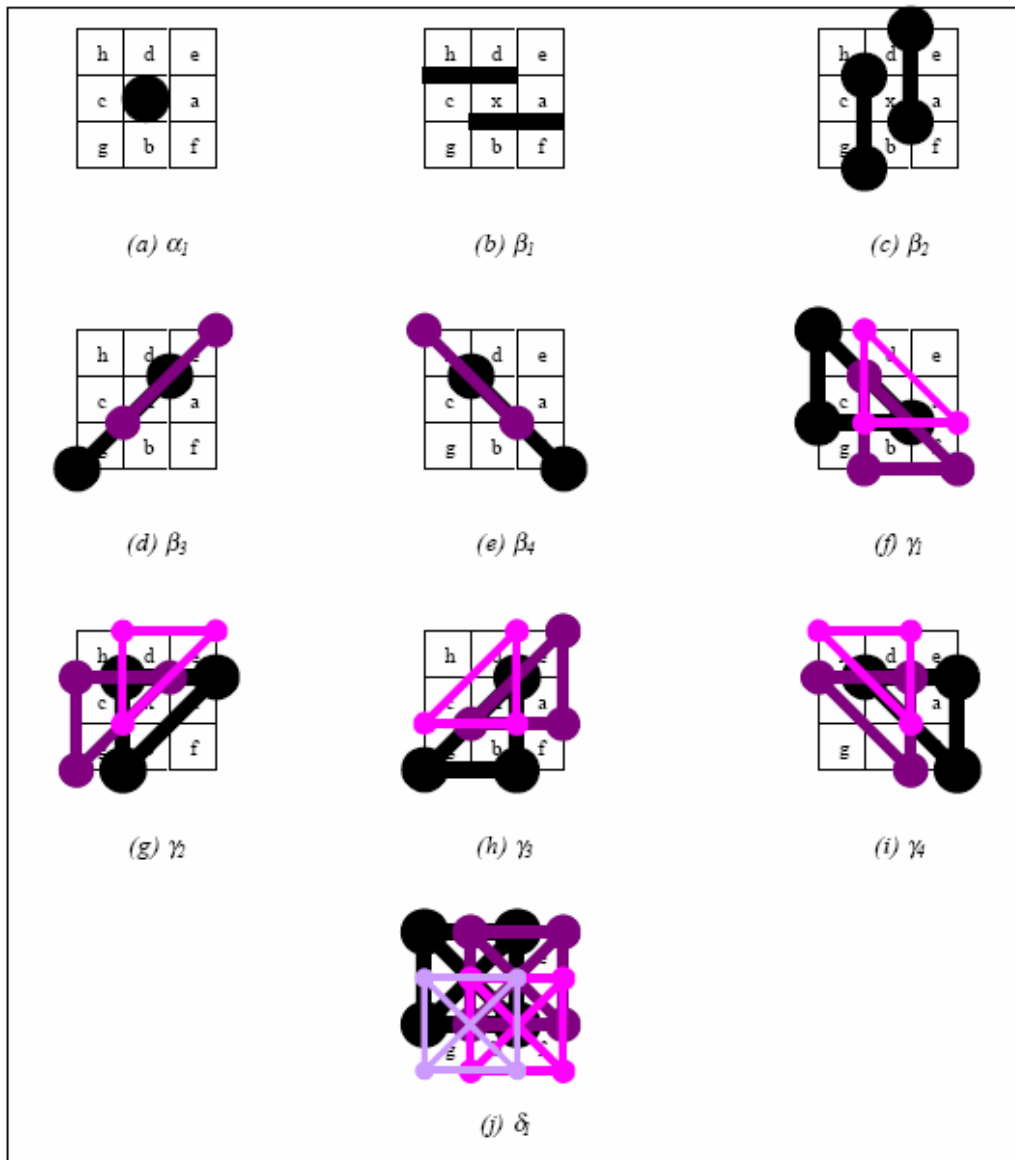


Figura 5.3 (Tipos de cliques y parámetros asociados en sistemas de 8-vecinos (segundo orden). Los estados son los puntos y las líneas muestran las relaciones entre vecinos)

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

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

Ecuación 5.31

Donde $n_i(I)$ es el número de estados en N_i que están etiquetados por I . Particularmente cuando sólo hay dos etiquetas, 0 y 1, el modelo se reduce al modelo de Ising.

5.5 Algoritmo de metrópolis a temperatura constante

Se trata de un método de relajación en el que se efectúan barridos sobre la imagen, y en cada píxel se tiene la posibilidad de efectuar cambios en el estado del píxel. Estos cambios involucran una variación de la energía del campo $U(x)$.

Sea X es la imagen original y Y es la imagen de clases

Algoritmo:

- Inicialización de la solución $x(0)$

- En cada etapa (n):

Escoger al site $r = s^n$ cuyo valor es x_r^n y su configuración de vecindad es V_r^n

Efectuar el sorteo de una variable de distribución uniforme $\xi \in \Omega$.

Calcular la variación de la energía local asociada

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

Ecuación 5.32

En este último paso puede existir una de dos eventualidades:

a) $\Delta U \leq 0$: puesto que la energía disminuye, el cambio es aceptado incondicionalmente, entonces:

$$x_r^{(n+1)} = \xi.$$

Ecuación 5.33

b) $\Delta U > 0$: el cambio es aceptado o rechazado efectuando un sorteo aleatorio con las probabilidades respectivas:

$$p = \prod (x_r^{(n+1)} = \xi) = \exp(-\Delta U)$$

Ecuación 5.34

Sea ξ_2 la variable obtenida en el sorteo aleatorio,

$$X_r^{(n+1)} = \xi \text{ si } \xi_2 < p.$$

Ecuación 5.35

5.6 Ejemplo de Modelo CAM para segmentación de imágenes

La segmentación de imágenes es un proceso que divide una escena en un conjunto de regiones disjuntas basándose en características similares como la intensidad, color o textura. La segmentación es el primer nivel en la tarea de entendimiento de la imagen y afecta severamente al proceso posterior de interpretación de la imagen, proporcionando estructuras útiles tales como regiones y bordes.

Como se había mencionado previamente, para segmentar una imagen, la imagen en cuestión tiene que ser modelada en diferentes tipos de regiones, cada uno de los cuales debe seguir una distribución normal caracterizada por un valor diferente de μ_i y varianza σ^2 . Si tomamos en cuenta la distribución de Gibbs de la región en proceso se establece $T = I$, se puede escribir una medida de probabilidad generalizada de que el píxel x pertenezca a algún tipo de región:

$$P(f_i = I | f_{N_i}) \propto \exp \left\{ -\frac{1}{2\sigma_i^2} (f_{N_i} - \mu_i)^2 - \sum_{c \in C} V_c(f_i) |_{f_i=I} \right\}$$

Ecuación 5.35

Este modelo se compone de dos partes: la primera obligando a que la intensidad de los píxeles que pertenezcan a una misma región sea casi idéntica y la segunda, estableciendo la continuidad espacial a través de la dependencia de las características de los píxeles vecinos al píxel central, y establece continuidad espacial. De acuerdo a la definición de la función de energía, es claro que si el número de píxeles que pertenecen a los tipos de regiones i en el vecindario son exactamente los mismos, el segundo término de la ecuación anterior tendrá el mismo valor para todos los tipos de regiones. Esto no cambiará el orden relativo

de los valores de probabilidad y la característica del píxel bajo consideración será determinada solamente por estos datos. Por el contrario, si estos números son diferentes unos de otros, el valor de probabilidad correspondiente al máximo de ellos obtendrá un incremento mayor y el píxel central será más probable que sea asignado al tipo de región afín a éste. Si el valor del término dependiente llega a ser insignificante, en el caso de que β sea muy grande por ejemplo, la característica de este píxel será decidida principalmente por la característica de sus vecinos. En pocas palabras, la fuerza espacial tendrá efecto sólo cuando el número de píxeles pertenecientes a algún tipo de región i en el grupo de vecinos del píxel central difieran uno de otro y esto causará que la característica de cada píxel en la segmentación de la imagen sea más parecida con las de la mayoría de sus vecinos. Finalmente, tomando la probabilidad logarítmica de la forma anterior, puede obtenerse fácilmente un tipo de clasificación de máxima verosimilitud.

VI. Esquema de fisión y fusión de datos por campos aleatorios de Markov.

6.1 Introducción.

La proliferación de sensores y captores han suministrado diversas fuentes de información de una escena o lugar, por ejemplo, en teledetección las imágenes del SAR o de satélites multi-canal suministran diferentes imágenes de la misma escena. La teoría de la decisión es ampliamente aplicada en procesamiento de imágenes para solucionar el problema de fusión de los datos. Entre las diferentes herramientas tenemos, fuzzy sets, Dempster and Shafer's theory [3] y aproximaciones probabilísticas. En este capítulo nos enfocaremos en aproximaciones probabilísticas usando campos aleatorios de Markov (CAM). Los campos aleatorios de Markov son muy usados para problemas de procesamiento de imágenes tales como segmentación de imágenes. En lugar de una fusión directa de datos, se propone un paso inicial de fisión el cual consiste en crear nuevas fuentes de información a partir de datos originales. De esta manera obtenemos numerosos captores (imágenes de textura), cada uno de ellos conteniendo información simple concerniente a una o algunas clases. De hecho, la información de cualquier captor puede ser fácilmente recuperada de su histograma. Una de las aplicaciones de esta metodología podría ser en la segmentación de regiones urbanas, no urbanas, vegetación y sombras. Primero se definen las diferentes clases (regiones urbanas, no urbanas, vegetación y sombras) y después se realiza el paso de fisión, estimando parámetros asociados con un campo aleatorio de Markov (CAM). Las respuestas de estos estimadores son usadas como nuevos captores en el paso de fusión. La fusión de datos es llevada a cabo usando el algoritmo de recocido simulado junto con el modelo de Potts para la probabilidad a priori (visto en el capítulo 5) con potenciales pieza lineales para el manejo de datos.

6.2 Fisión de datos

El primer paso en nuestro planteamiento consiste en extraer información simplificada de los datos originales. Con ésto se crean nuevos captores, que representan el resultado de la estimación de un parámetro textural. Tales captores son discriminados de una o varias de sus clases, considerando los modos de sus histogramas. Esto puede ser visto como un paso de pre-procesamiento, separando los diferentes tipos de información contenida en la imagen original. De esta manera el proceso de fisión será simplificado y los parámetros asociados estimados más fácilmente. Las interacciones entre pixeles adyacentes (vertical y horizontal) son modelados por una distribución gaussiana. Otro término gaussiano es añadido para cada pixel. El Hamiltoniano asociado con este campo está dado por:

$$H(\Sigma) = \frac{U(\Sigma)}{T} = \frac{1}{T} \left(\sum_{c=\{i,j\}} (\sigma_i - \sigma_j)^2 + \lambda \sum_{c=\{i\}} (\sigma_i - \mu)^2 \right)$$

Ecuación 6.1

que depende de los tres parámetros T , λ y μ . El parámetro T se refiere a la temperatura, puede ser interpretado como varianza condicional, mientras λ es similar a un peso de correlación y μ es la media local. Ciertamente, las propiedades Markovianas en las probabilidades condicionales locales implican que:

$$P(\sigma_i/\sigma_j, j \in V_i) = \mathcal{N} \left(\frac{\lambda\mu + 4m_i}{4 + \lambda}, \frac{T}{2(4 + \lambda)} \right)$$

Ecuación 6.2

donde V_i es la serie de cuatro vecinos de i , m_i es el promedio de estos vecinos y $N(a,b)$ denota una distribución normal de media a y varianza b . Así pues obtendremos 11 captores en total por los descriptores de Haralack utilizando el método de matriz de co-ocurrencia visto en el capítulo IV, de entre estas once imágenes seleccionamos las más relevantes para la discriminación de las clases definidas. Los captores escogidos son entonces los de homogeneidad, entropía y autocorrelación, ya que estos captores son en los que se distinguen mejor las clases (vegetación, urbano y calles). Podemos entonces simplificar usando las operaciones lógicas, dándonos una idea para la discriminación de los potenciales de los sistemas escogidos (captors). En la presente tesis se escogieron

los captosres de homogeneidad, entropía y autocorrelación los cuales se muestran a continuación:

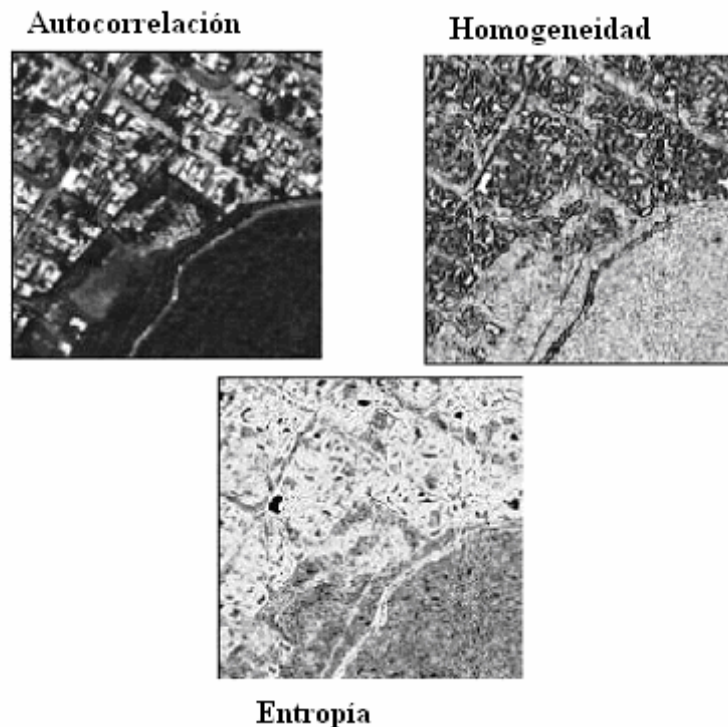


Figura 6.1 Captosres seleccionados obtenidos durante el paso de la fisión

6.3 Fusión de datos

Para tratar el paso de fusión usamos una aproximación de regulación usando un campo aleatorio de Markov en un esqueleto Bayesiano. Sea Y la imagen segmentada y los X_i los diferentes captosres. El proceso de segmentación consiste en maximizar la probabilidad condicional $P(Y/X_i, i \in (0, \dots, n))$ la cual, de la regla de Bayes, es proporcional a $P(Y/Y_i, i \in (0, \dots, n)/Y) * P(Y)$. $P(Y)$ está referido como el modelo a priori. Los diferentes captosres X_i son las variables independientes condicionales de Y , para que la probabilidad $P(X_i, i \in (0, \dots, n)/Y)$ pueda ser escrita como un producto $\prod_i P(X_i/Y)$. Estos términos son referidos como los términos de datos manejados.

El modelo a priori esta compuesto de un modelo de Potts de 8 conexidad. La función de energía asociada con este modelo puede ser escrita como:

$$U(Y) = \sum_{c=\{i,j\}} \sum_{l,l'} \gamma_{l,l'} \delta_{y_i=l} \delta_{y_j=l'}$$

Ecuación 6.3

Los parámetros asociados están entonces definidos por la matriz de regularización $\Gamma = (\gamma_{l,l'})$ donde los índices representan a las diferentes clases, δ_y es la función delta de Kronecker. La matriz de regulación es igual a γ veces la matriz identidad, de aquí obtenemos el modelo clásico de Potts[7]. Los términos de la diagonal reflejan las constantes de homogeneidad, esto significa que entre mayor son estos términos más parecidos son el que dos pixeles tengan la misma etiqueta. El segundo término nos permite favorecer o penalizar la conectividad entre 2 clases diferentes.

El paso de fisión, extrae información discriminante de cada clase, permite una simple formulación para los términos manejados. Necesitamos definir funciones pieza lineales $f(i,j)$ extraídas directamente de los histogramas en lugar de usar un modelo probabilístico complejo para los datos originales tal como se muestra en la Figura 6.2.

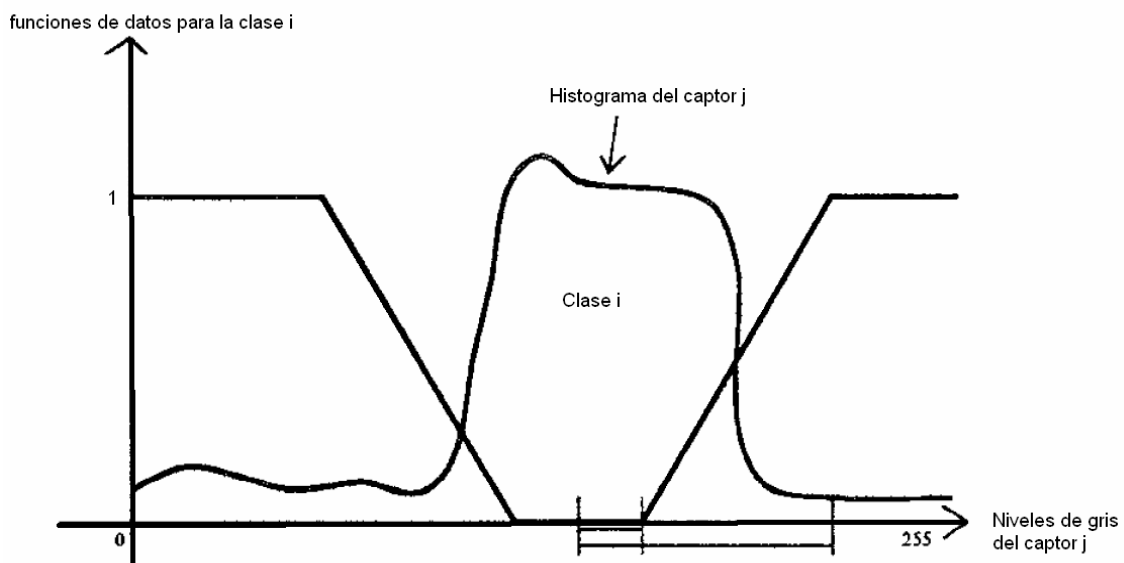


Figura 6.2 Función de datos manejados $f(i,j)$ y parámetros asociados

De esta manera definimos funciones para cada clase y cada captor. El potencial resultante $V_s(j)$ es una relación lineal entre esas funciones. Considere un captor i y un modo de histogramas correspondientes a una clase dada j . Si el valor de un pixel dado en el captor yace en el modo, este captor i vota por la clase j y la función $f(i,j)$ es igual a 0. Fuera del modo, la función es igual a 1, penalizando la clase para ese pixel. Entre

estos dos valores extremos, una pendiente lineal indica que la decisión depende principalmente del otro captor o de la información contextual. Los captores son más o menos exactos. De esta manera, podemos definir un factor confidencial multiplicativo $\lambda_{i,j}$, el cual es igual a 1 si hay una equivalencia entre un modo dado y la clase considerada. Este factor es igual a 0.5 si hay sólo una implicación e igual a 0 si los valores de los píxeles en el captor para la clase son confusos en el histograma [4]. Un factor de normalización en el denominador se usa para evitar favoritismos o penalizaciones en una clase dada por estos pesos. El potencial resultante puede ser escrito:

$$V_{site\ s} = \sum_{captors\ i} \sum_{classes\ j} \frac{\lambda_{i,j} f^{i,j}(x_s) \delta_{y_s=j}}{\sum_{captors\ i} \lambda_{i,j}}$$

Ecuación 6.4

6.4 Resultados de fusión de datos

El algoritmo resultante para esta aproximación de fisión y fusión esta contemplado en la siguiente figura:

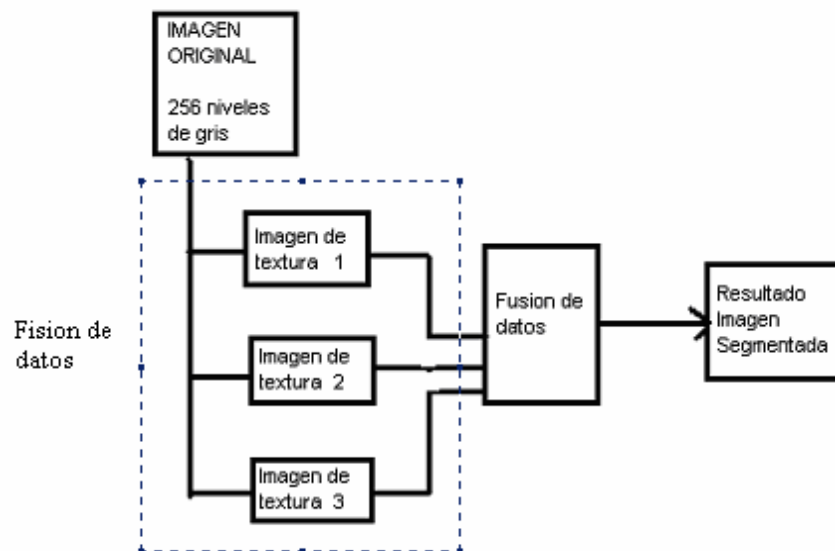


Figura 6.3 Esquema de fisión y fusión de datos

La minimización de la energía usada en el paso de fusión se lleva a cabo usando recocido simulado. En la práctica usamos un decremento geométrico de la temperatura de .05% al valor actual de la temperatura en cada iteración. El resultado es obtenido

después de 80 iteraciones que representan 24 horas del tiempo de un CPU (con procesador Pentium IV de 1.7GHz y 640MB de RAM) para una imagen de 512 x 512 pixeles. La figura siguiente representa el resultado final.

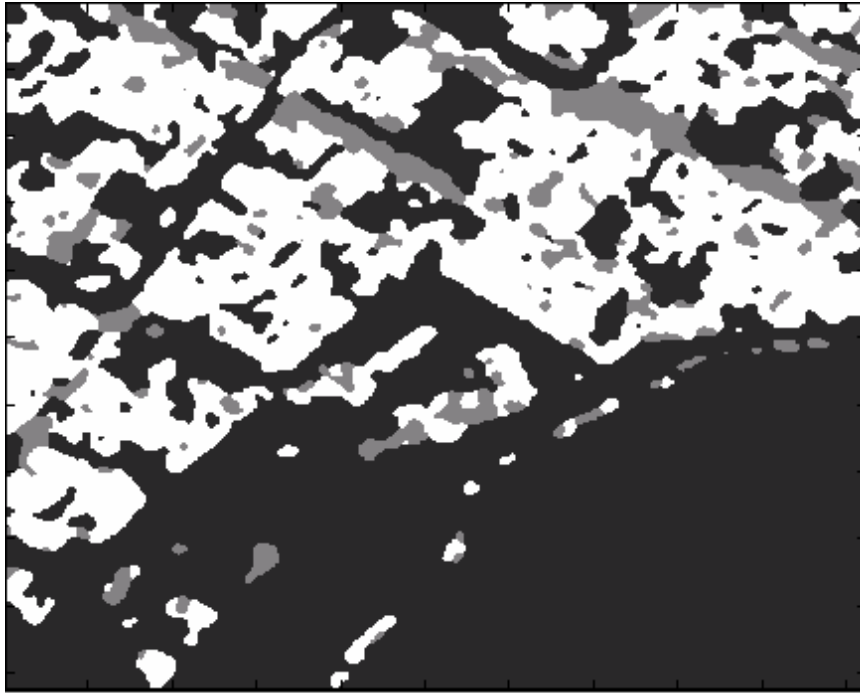


Figura 6.4 Resultado de la segmentación de fisión y fusión

La fisión permite la discriminación entre los diferentes tipos de clases. Además, la exactitud de los estimadores permite limpiar desperfectos en el paso de fusión.

En el presente trabajo de tesis se ha propuesto una metodología llamada fisión y fusión para segmentación de imágenes [3] . El paso de fisión, consiste en la obtención de estimadores texturales, realizado en base al modelo de campo aleatorio de markov gaussiano; mientras que el paso de fusión usa el modelo de Potts y el criterio de metrópolis. El paso de fisión simplifica el problema debido a que estudia los modos del histograma de los nuevos captores. Su aplicación en la segmentación de zonas urbanas prueba el poder de la aproximación markoviana, tanto en un modo analítico (durante el paso de fisión) como en un modo sintético (durante el paso de fusión).

El esquema que se presenta esta tesis es la obtención de los descriptores de textura a través de la matriz de co-ocurrencia y seleccionar sólo aquellos que se consideren óptimos para la fusión de datos; es decir sólo se consideran aquellos en donde las clases

sean mejor identificables del resto. En algunas imágenes se distinguirán muy bien lo urbano pero la vegetación no muy bien, mientras que en otras se distinguirán muy bien la vegetación y las calles pero las casas no. Es aquí donde entra el factor de confiabilidad, el cual expresa el criterio que se está tomando para las clases del siguiente modo (1= equivalencia , 0.5= implicación y 0 = no equivalencia).

Una vez que se han seleccionado cuidadosamente las imágenes de textura a fusionar hay que identificar las clases en cada imagen. Tomando en cuenta que en algunas imágenes las clases se ven mejor que en otras, por ejemplo en la imagen de homogeneidad la clase de vegetación se ve mejor que en la imagen de entropía, pero resulta que en la imagen de entropía la clase de calles se aprecia mejor que en la de homogeneidad y que en la de autocorrelación. Teniendo en cuenta lo anterior y visualizando muy bien las clases en las imágenes escogidas se procede a tomar un criterio en el factor multiplicativo de confiabilidad.

Se define una probabilidad a posteriori usando la regla de Bayes para representar el resultado de la segmentación, a partir de un modelo a priori y un modelo de probabilidad:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Ecuación 6.6 Regla de Bayes

Empleando la aproximación de la regla de Bayes (Apéndice A) tenemos:

$$P(Y|X) \doteq P(X|Y) P(Y)$$

Ecuación 6.7 Aproximación de la regla de Bayes

A través de aplicar los campos aleatorios de Markov a la aproximación de la regla de Bayes se llega a la siguiente expresión:

$$U(Y|X) \doteq U(X|Y) + U(Y)$$

Ecuación 6.7

Esta expresión se trata de minimizar con recocido simulado, ya que esta es la que modela el sistema a analizar

- Se utiliza el algoritmo de metrópolis de temperatura variable con el fin de minimizar la energía $U(\mathbf{X}|\mathbf{Y})$ ó $U(\mathbf{X}|\mathbf{W})$. .

- Para el caso de la energía $U(\mathbf{Y})$ ó $U(\mathbf{W}_n)$ se considera el modelo de Potts (visto en el capítulo V), el cual se aplica en vecindades en 8-conexidad.

El modelo de probabilidad está relacionado con la forma de observar los datos, dependiente del dominio. El modelo a priori expresa ciertas restricciones a priori. Los resultados de la teoría CAM nos proporcionan herramientas para codificar restricciones contextuales en la probabilidad a priori.

Para fusionar los datos de cada descriptor de textura, primero hay que obtener las funciones de energía de cada clase en particular (en el caso del presente trabajo de tesis es por campos de Markov) y posteriormente multiplicarlas por su factor de confiabilidad (λ). Después se suman las correspondientes funciones de energía de cada clase de cada captor ya multiplicadas por su factor de confiabilidad para obtener una nueva función de energía que involucre los captosres y sus respectivas medias es decir, la función de energía de la clase1 (W1) del captor 1 (X1) con la función de energía de la clase1 (W1) del captor 2 (X2) y así sucesivamente; es decir $\lambda_o * U(X_o|W1) + \lambda_1 * U(X1|W1) + \lambda_2 * U(X2|W1) + \dots + \lambda_n * U(Xn|W1)$. A continuación se normaliza el resultado dividiéndose por un valor igual a la suma de los correspondientes factores de confiabilidad empleados para cada caso ($\lambda_o + \lambda_1 + \lambda_2 + \dots + \lambda_n$) y, por último se minimiza la nueva función de energía a través del modelo de Potts, con un número de iteraciones igual a 80 y una temperatura inicial $T = 2.5$, hasta obtener la imagen segmentada en clases (resultado final). El proceso se muestra en la Figura 6.5

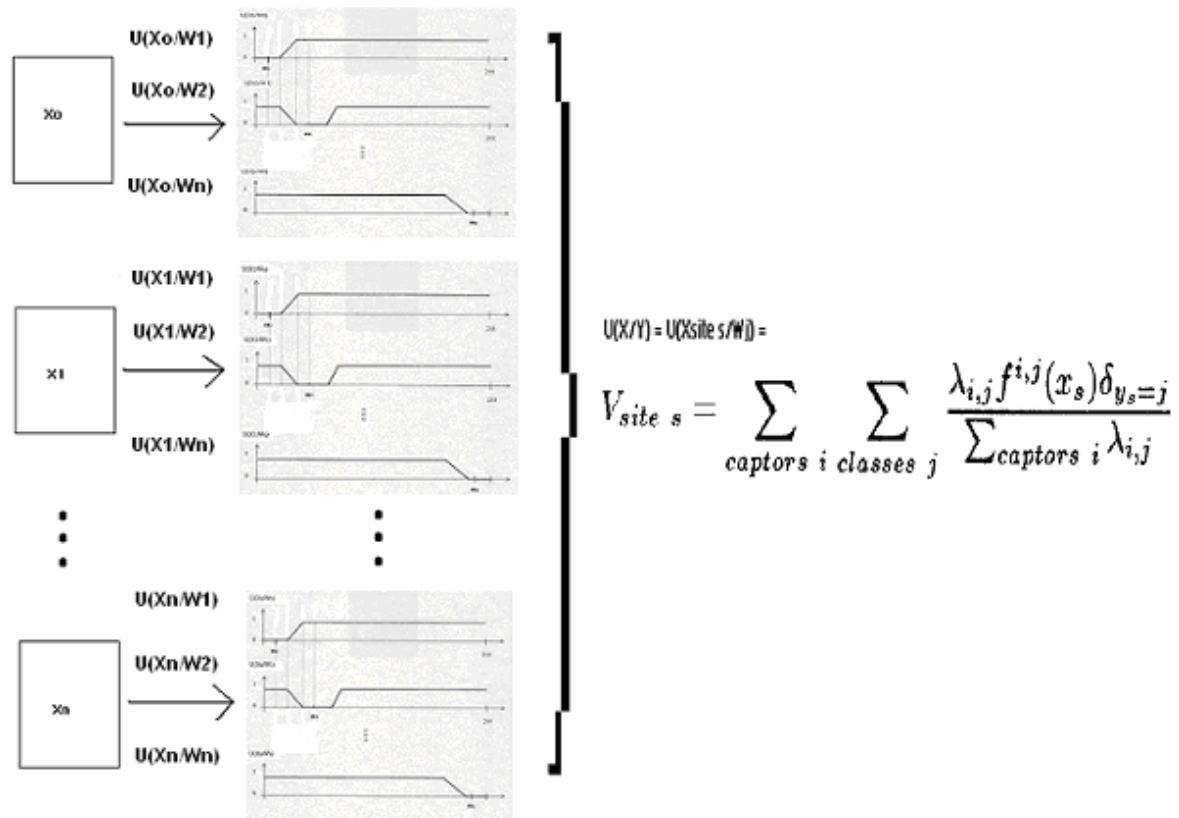


Figura 6.5 Esquema de fusión de datos (desglosado)

Donde:

$U(\mathbf{X}|\mathbf{Y}) = U(\mathbf{X}_{site\ s}|\mathbf{W}_j)$ = representa la función de energía condicional en el píxel X ó site s sobre las imágenes de textura con respecto a la clase W_j

$f_{i,j}(x_s)$ representa a la función de energía condicional $U(\mathbf{X}_i|\mathbf{W}_n)$

de manera general, tenemos

$$U(\mathbf{Y}|\mathbf{X}) = U(\mathbf{X}|\mathbf{Y}) + U(\mathbf{Y}) \quad \text{ó} \quad U(\mathbf{W}_j|\mathbf{X}_{site\ s}) = U(\mathbf{X}_{site\ s}|\mathbf{W}_j) + U(\mathbf{W}_j)$$

En términos de probabilidad es:

$$P(\mathbf{Y}|\mathbf{X}_i, i \in \{0, \dots, n\}) = P(\mathbf{X}_i, i \in \{0, \dots, n\}|\mathbf{Y}) * P(\mathbf{Y})$$

Donde:

$$P(X_i, i \in \{0, \dots, n\} | Y) = \prod_i P(X_i | Y)$$

Para las clases $W_j, j \in \{0, \dots, n\}$

Para los captosres (imágenes de textura) $X_i, i \in \{0, \dots, n\}$

Para el píxel o site s , la ecuación anterior evaluada con respecto a la clase W_n y sobre todos los captosres, en términos de energía queda expresada como:

$$U(W_n | X) = U(X_0 | W_n) + U(X_1 | W_n) + \dots + U(X_n | W_n) + U(W_n)$$

Para: $U(X_0 | W_n) + U(X_1 | W_n) + \dots + U(X_n | W_n)$ equivale a $\prod_i P(X_i | Y)$

Donde $U(W_n)$ ó $U(Y)$ se evalúa mediante el modelo de Potts en un análisis de 8 conexidad comparando el site s con sus vecinos.

Para poder obtener la imagen final se emplea el algoritmo de recocido simulado partiendo de una imagen aleatoria .

6.5 Algoritmo de Recocido Simulado para fusión de datos

Sea X la imagen original (en niveles de gris) e Y la imagen segmentada (por ejemplo, en cuatro clases), la simulación toma como datos iniciales a X y a la imagen de clases distribuida de manera aleatoria, originada en la inicialización, $Y(n)$. La Figura 6.6 muestra la esquematización del algoritmo de simulación, para el caso de n iteraciones:

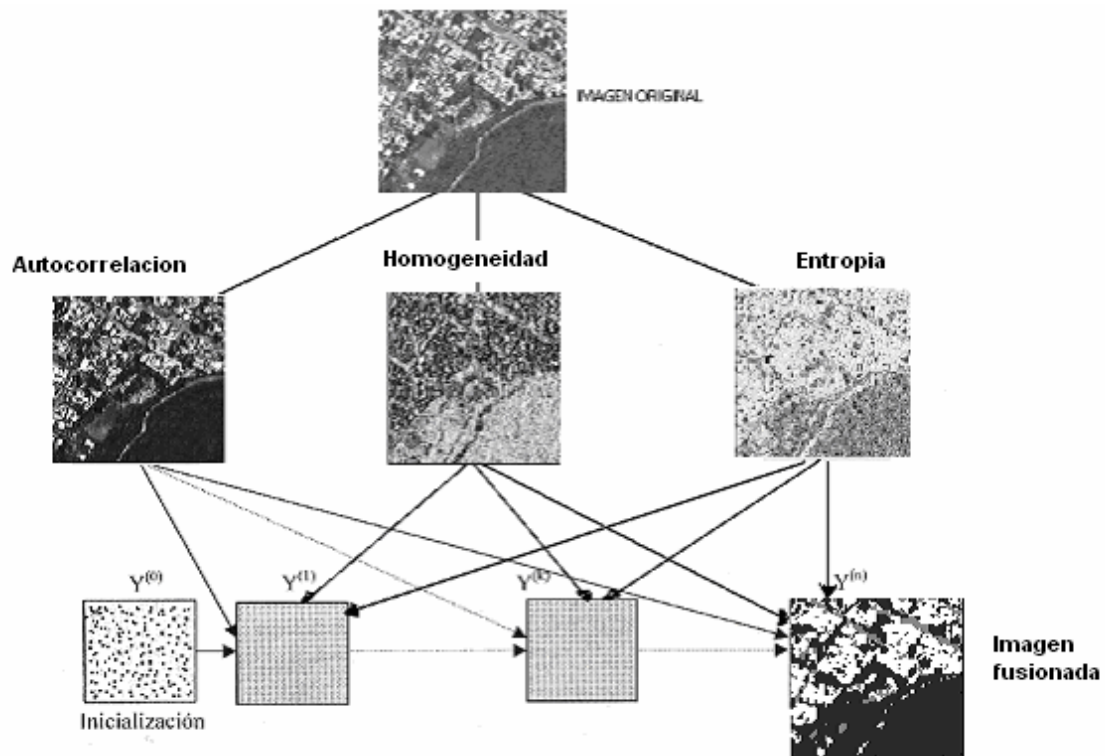


Figura 6.6 Esquema de fisión y fusión de datos

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 ξ).
- Escoger una temperatura inicial $T(0)$ suficientemente elevada (por ejemplo, $T(0)=2.5$).

2. Hacer un contador con un número específico de iteraciones $N=1.....n$

2.1 Barrer secuencialmente la imagen

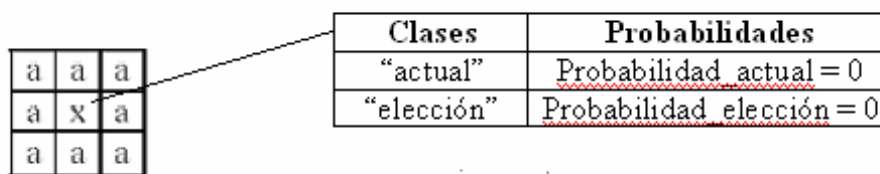
En cada etapa (n)

- Escoger un site: $r=s(n)$, su valor es $Y_r(n)$ y su configuración de vecindad es $V_r(n)$
- Efectuar el sorteo de una variable aleatoria (clase “elección”) con distribución uniforme en el rango del número de clases ξ (solo se varia elección y se compara contra actual).

2.2 Criterio de metrópolis y modelo de Potts

Se utiliza el algoritmo de metrópolis de temperatura variable con el fin de minimizar la energía $U(\mathbf{X}|\mathbf{Y})$ ó $U(\mathbf{X}|\mathbf{W}_j)$.

Para el caso de la energía $U(\mathbf{Y})$ ó $U(\mathbf{W}_n)$ se considera el modelo de Potts (sección 5.3.2), aplicado a vecindades de 8-conexidad.



Análisis 8-conexidad

- 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:

- $\text{prob_actual} = \text{prob_actual} + \text{potencial}[\text{actual}][i][j]$
- $\text{prob_elección} = \text{prob_elección} + \text{potencial}[\text{elección}][i][j]$

donde $\text{potencial}[\text{actual}][i][j]$ y $\text{potencial}[\text{elección}][i][j]$ representa la energía condicional $U(\mathbf{X}|\mathbf{Y})$ e $[i][j]$ son las coordenadas del píxel analizado.

Para el caso de la fusión aquí presentada tendremos:

$\text{potencial}[\text{actual}][i][j] =$

$$\lambda_1[U(\mathbf{X}_0|\mathbf{W}_n)] + \lambda_2[U(\mathbf{X}_1|\mathbf{W}_n)] + \dots + \lambda_n[U(\mathbf{X}_n|\mathbf{W}_n)] / (\lambda_1 + \lambda_2 + \dots + \lambda_n)$$

$\text{potencial}[\text{elección}][i][j] =$

$$\lambda_1[U(\mathbf{X}_0|\mathbf{W}_n)] + \lambda_2[U(\mathbf{X}_1|\mathbf{W}_n)] + \dots + \lambda_n[U(\mathbf{X}_n|\mathbf{W}_n)] / (\lambda_1 + \lambda_2 + \dots + \lambda_n)$$

- Si $\text{prob_elección} \leq \text{prob_actual}$: $Y[i][j] = \text{“elección”}$

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 de $[0,1]$
- Si $\xi_2 < p$ entonces $Y[i][j] = \text{“elección”}$
- Si no $Y[i][j] = \text{“actual”}$

3. Regresar al punto 2.1 hasta terminar un barrido

4. Decrementar la temperatura

$$T(n) = T(n-1)*0.95$$

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

VII. Análisis de resultados

7.1 Resultados para la imagen DF1_izq (20,30,40,60 y 80 iteraciones)

En el presente trabajo las imágenes mostradas a continuación son el resultado de la fusión de datos a partir de 3 de sus descriptores de textura, los cuales fueron obtenidos de la imagen original bajo un esquema de fusión de datos por matriz de co-ocurrencia. Estos descriptores de textura son el de homogeneidad, el de autocorrelación y el de entropía, los cuales fueron seleccionados por sus características particulares de definición de medias y dependiendo de cada caso. A continuación se muestran los 11 descriptores de textura de Haralick obtenidos por matriz de co-ocurrencia para la 1° imagen del distrito federal (DF1_izq) y posteriormente los 3 que se seleccionaron de estos 11 figura 7.1. Los cuales se obtuvieron reduciendo a 32 niveles de gris, con un desplazamiento horizontal de 2, y para un tamaño de la ventana de 5x5.



Figura 7.1 Imagen del Distrito Federal (DF1_izq)

Descriptores de textura obtenidos por matriz de co-ocurrencia 11 en total de estos se seleccionaron 3 para el proceso de fisión de datos. Los parámetros de entrada del programa se muestran a continuación.

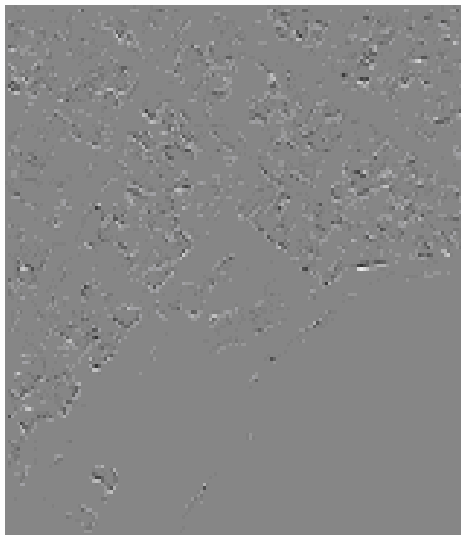
Imagen	No. De niveles de gris reducidos	Desplazamiento Horizontal	Tamaño de la ventana
DF1_izq	32	2	5x5



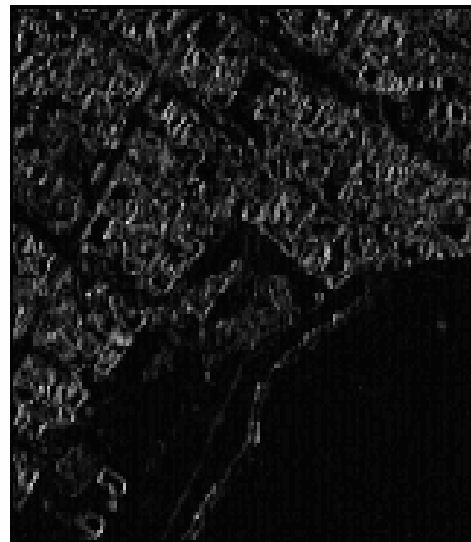
Autocorrelacion



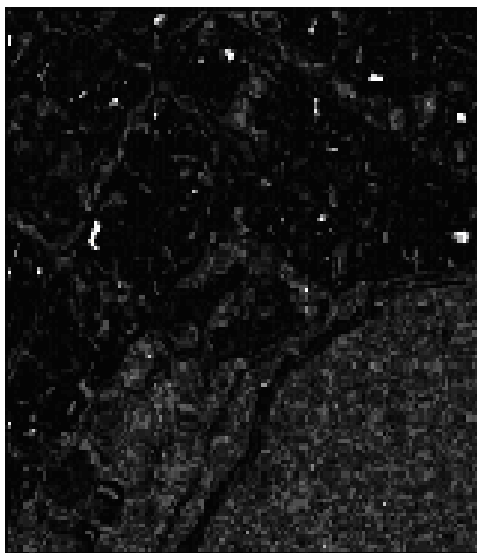
Cluster prominence



Cluster shade



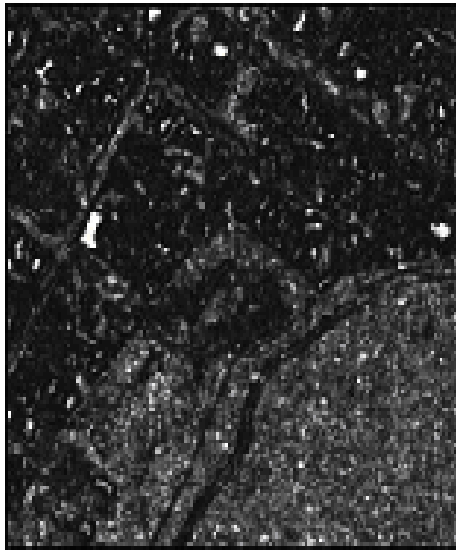
Contraste



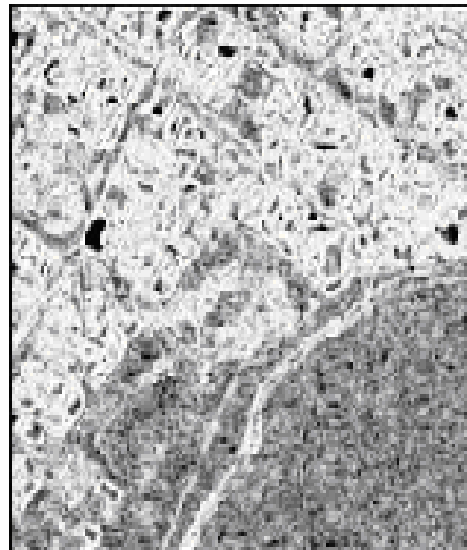
Correlacion



Dissimilarity



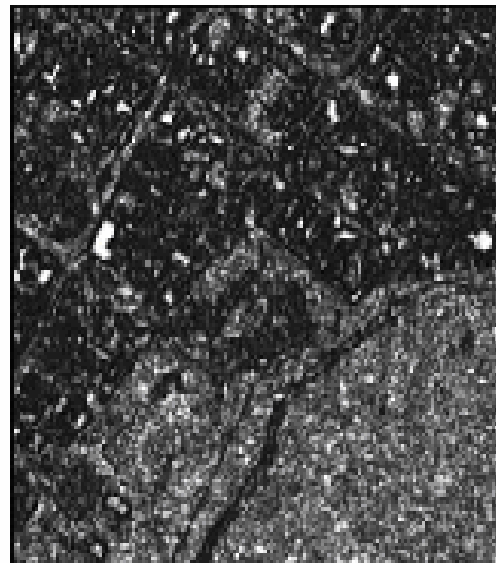
Energía



Entropía



Homogeneity



Probabilidad maxima



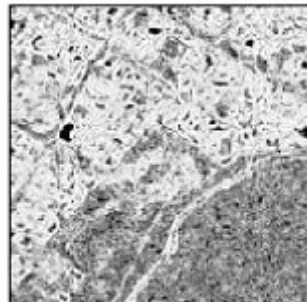
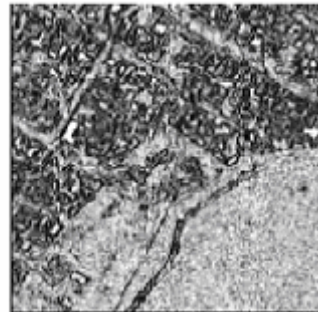
Varianza

De los cuales los que se seleccionaron fueron los correspondientes a autocorrelación, entropía y homogeneidad, mostrados a continuación.

Autocorrelación



Homogeneidad



Entropía

Para la fusión de datos se tomó en cuenta el factor de confiabilidad dependiendo de la clase que se aprecia mejor en cada imagen de textura con respecto a las demás. Estos factores (λ) se multiplicarán por cada clase a identificar en cada imagen de textura se presentan a continuación.

Para la presente imagen del distrito DF1_izq (Figura7.1) sólo se identificaron 3 clases, correspondientes a vegetación, urbano y calles.

Imagen de textura autocorrelación

Clase	Valor del factor de confiabilidad (λ)
Vegetación	0.5
Urbano	1
Calles	0.5

Imagen de textura homogeneidad

Clase	Valor del factor de confiabilidad (λ)
Vegetación	0.5
Urbano	0.5
Calles	0.5

Imagen de textura entropía

Clase	Valor del factor de confiabilidad (λ)
Vegetación	0.5
Urbano	0.5
Calles	1

Los resultados de la fusión de datos se muestran a continuación. Estos resultados son productos del algoritmo de recocido simulado para fusión de datos para diferente número de iteraciones, partiendo de una imagen aleatoria y haciendo un barrido de toda la imagen en un análisis en 8-conexidad, con una temperatura de $T = 2.5$. En estos resultados se muestra la imagen en niveles de gris junto con la de colores falsos, de tal manera que se puedan identificar con mayor facilidad las clases (calles, urbano y vegetación). En las imágenes de las figuras 7.2 a la 7.12 se muestra la secuencia que se va tomando al recocer la imagen a partir de la imagen aleatoria, se muestra que gradualmente a medida que se incrementa el número de iteraciones (20,30,40,60 y 80) la aleatoriedad se va disminuyendo dándonos una imagen final segmentada en clases producto de la fusión de datos

Imagen aleatoria

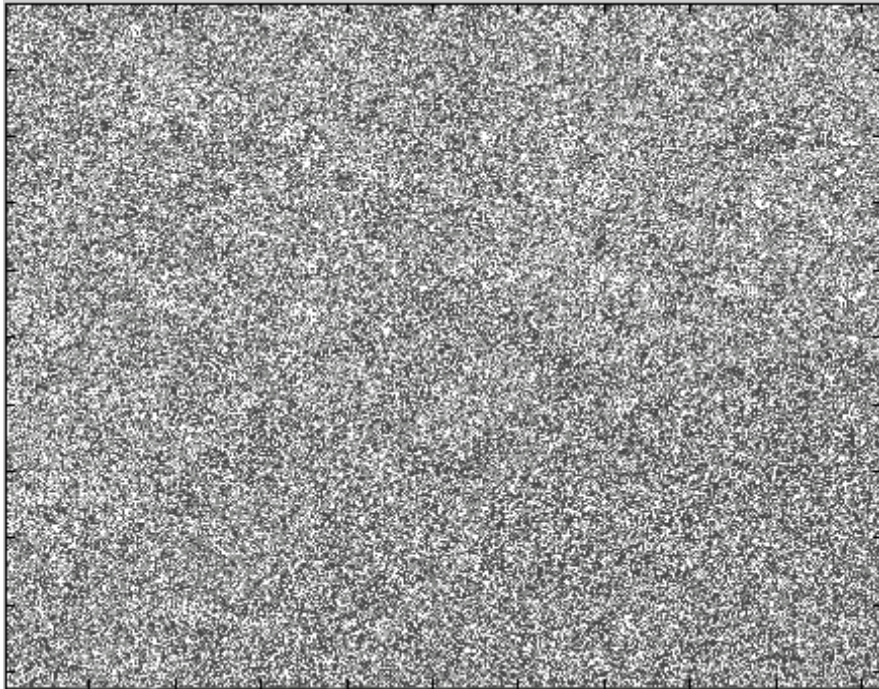


Figura 7.2 Imagen aleatoria

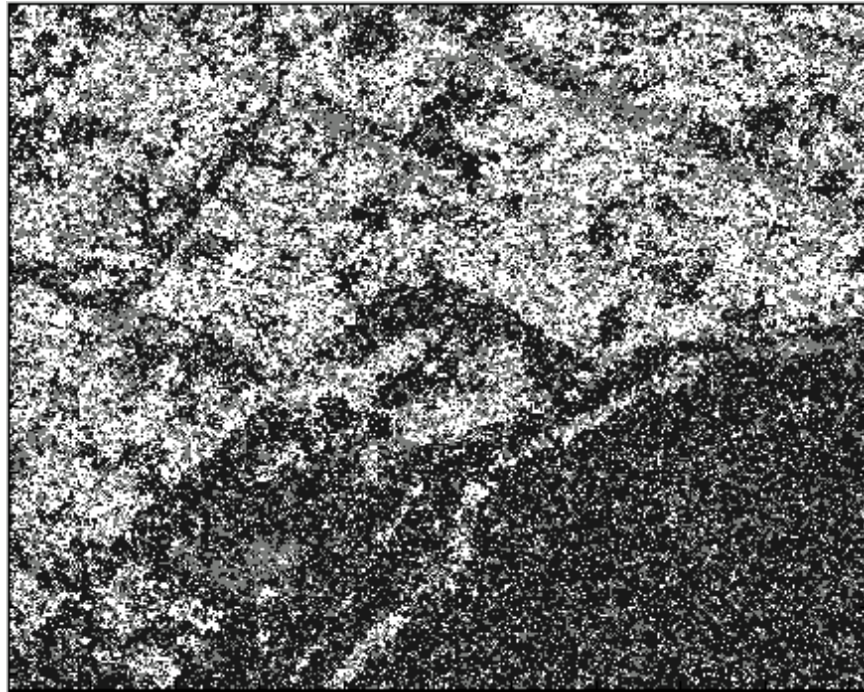


Figura 7.3 Resultado para 20 iteraciones del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad o recociendo la imagen

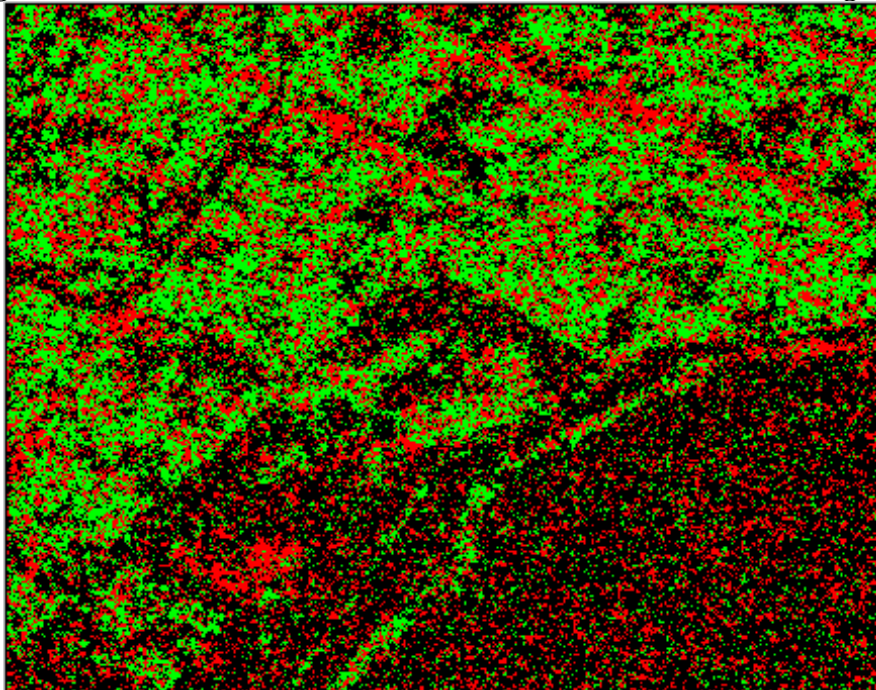


Figura 7.4 Resultado para 20 iteraciones (imagen en colores) falsos del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad o recociendo la imagen de la inicialización aleatoria

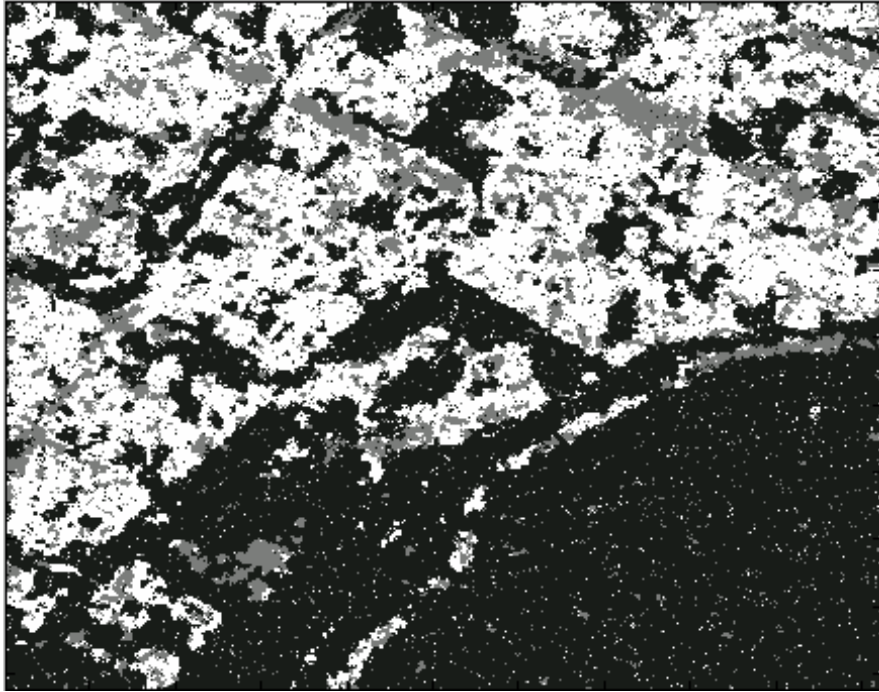


Figura 7.5 Resultado para 30 iteraciones del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad dándonos una mejor imagen

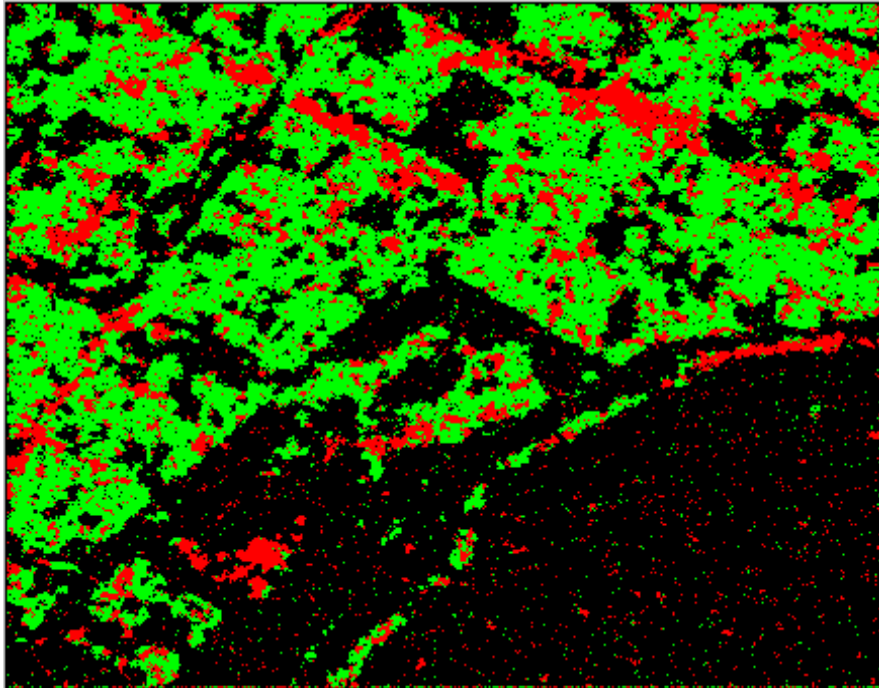


Figura 7.6 Resultado para 30 iteraciones (imagen en colores) falsos del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad (recociéndose) dándonos una mejor imagen

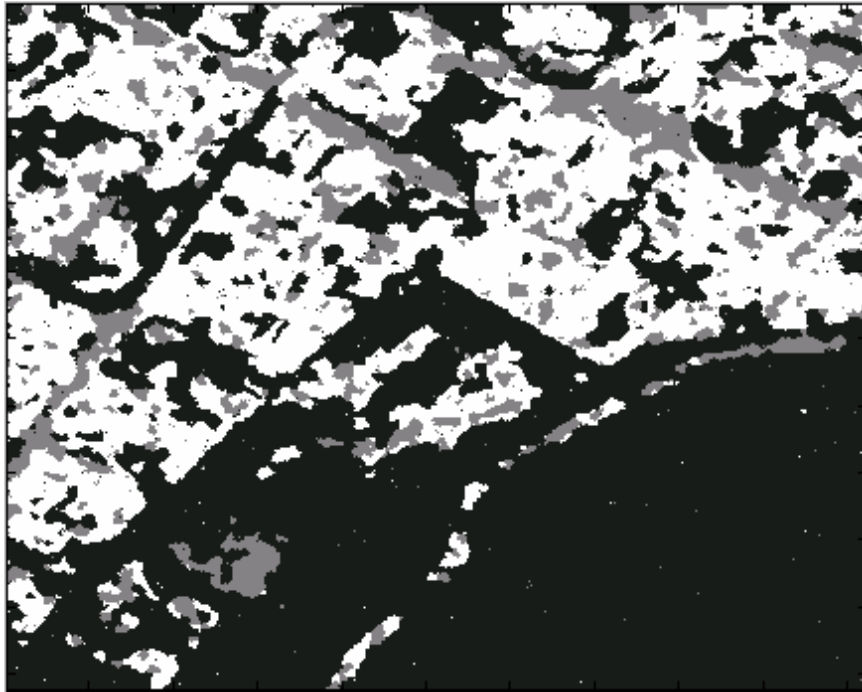


Figura 7.7 Resultado para 40 iteraciones del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han disminuido, dándonos una imagen mas clara segmentada en clases

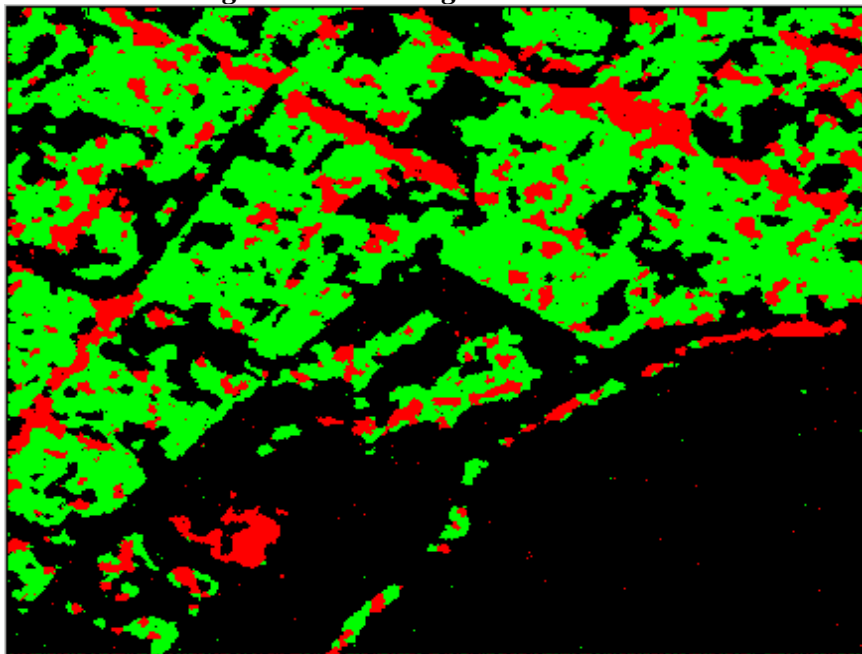


Figura 7.8 Resultado para 40 iteraciones (imagen en colores) falsos del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han disminuido, dándonos una imagen mas clara segmentada en clases

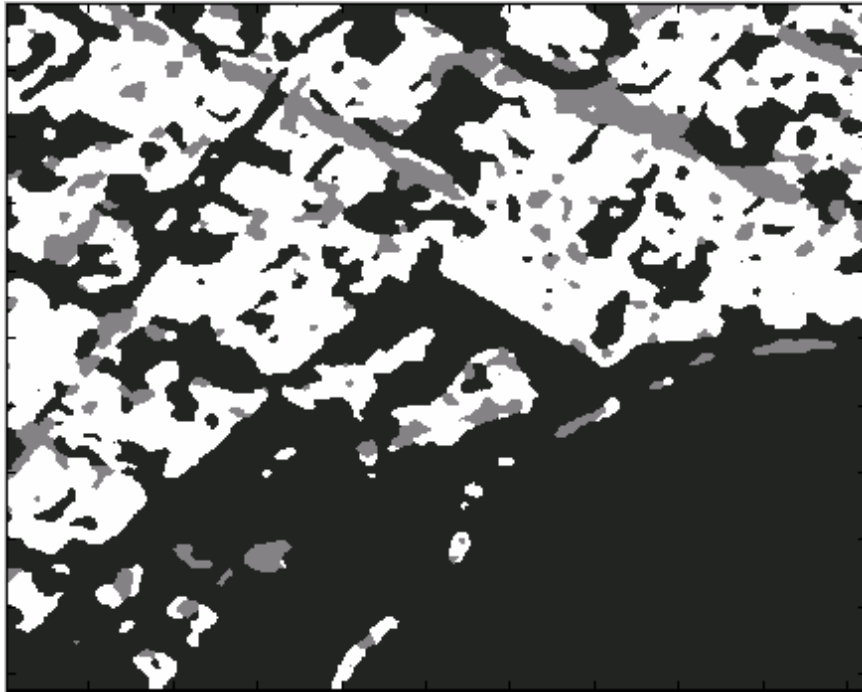


Figura 7.9 Resultado para 60 iteraciones del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases casi han desaparecido en un 100 %, visualizando una imagen segmentada en clases (recocida) pero no precisa

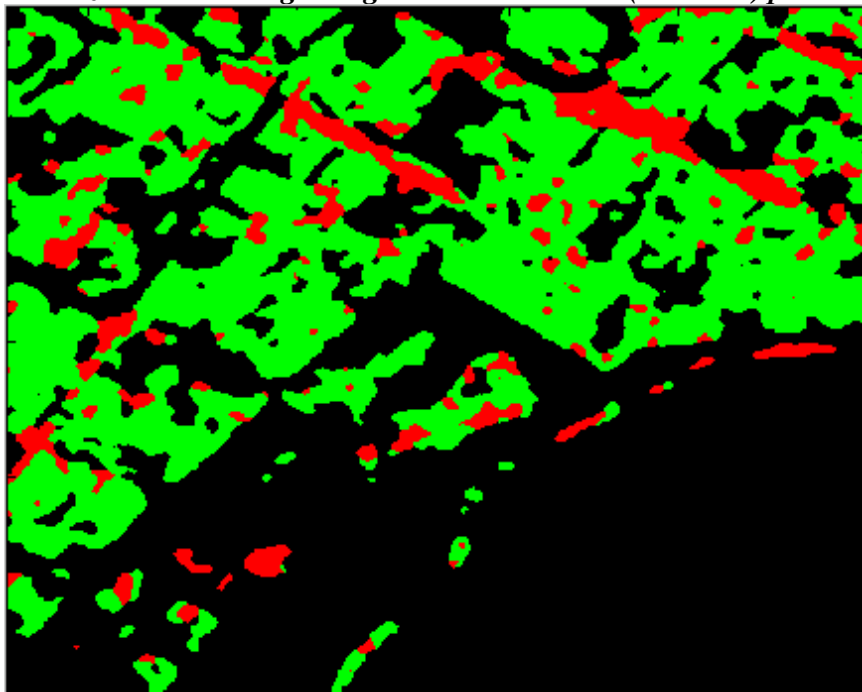


Figura 7.10 Resultado para 60 iteraciones imagen en colores falsos del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases casi han desaparecido en un 100 %, visualizando una imagen segmentada en clases (recocida) pero no precisa en lo que respecta a agrupamiento

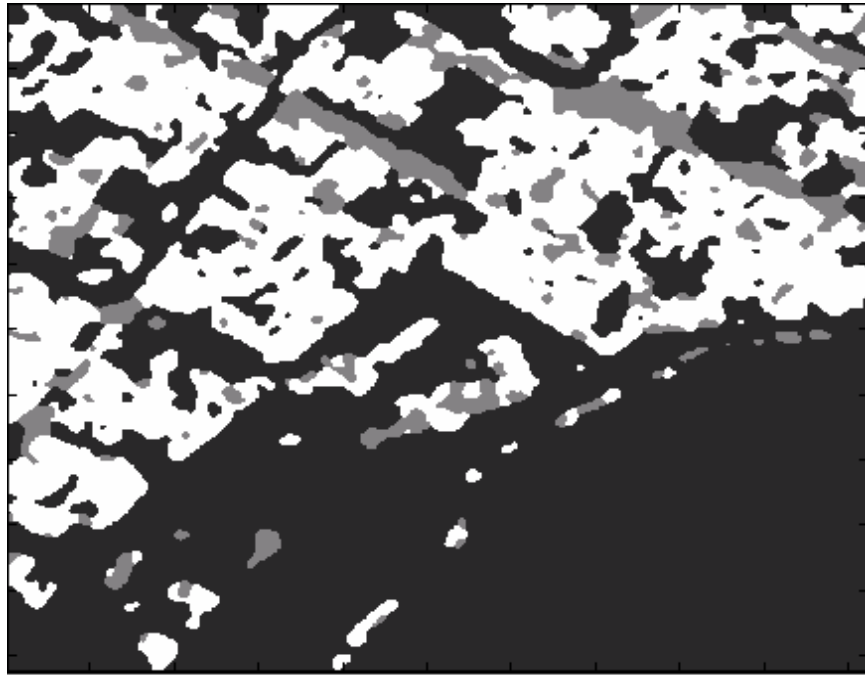


Figura 7.11 Resultado para 80 iteraciones del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han desaparecido (imagen recocida) visualizando una imagen segmentada en clases (urbano, vegetación y calles) con un agrupamiento más exacto e identificable

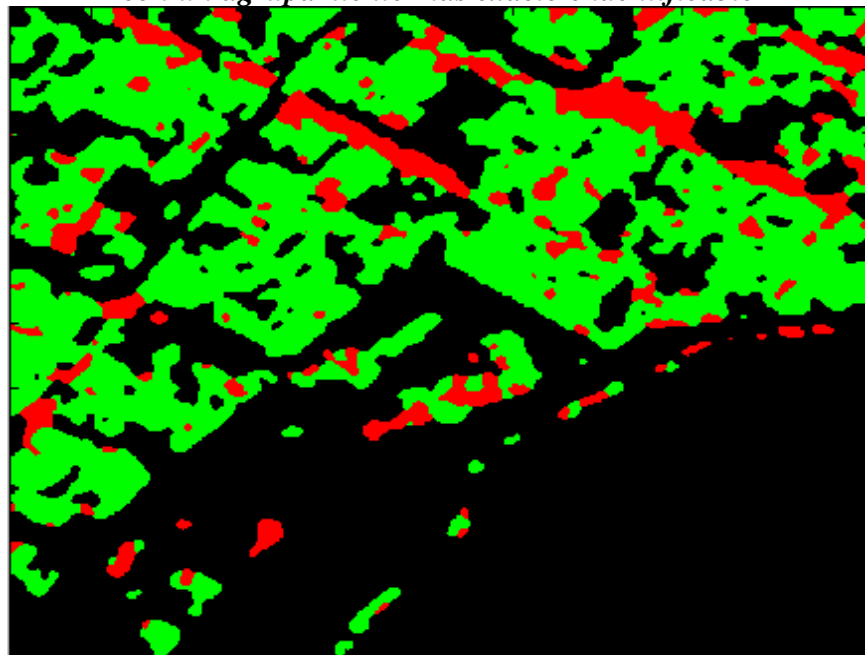


Figura 7.12 Resultado para 80 iteraciones (imagen en colores) falsos del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han desaparecido (imagen recocida) visualizando una imagen segmentada en clases (urbano, vegetación y calles) con un agrupamiento más exacto e identificable

7.2 Resultados para la imagen DF1_pr1 (20,30,40,60 y 80 iteraciones)

Las imágenes mostradas a continuación son el resultado de la fusión de datos a partir de 3 de sus descriptores de textura, los cuales fueron obtenidos de la imagen original bajo un esquema de fusión de datos por matriz de co-ocurrencia. Estos descriptores de textura son el de homogeneidad, el de autocorrelación y el de entropía, los cuales fueron seleccionados por sus características particulares de definición de medias y dependiendo de cada caso. A continuación se muestran los 11 descriptores de textura de Haralick obtenidos por matriz de co-ocurrencia para la 2^o imagen del distrito federal (DF1_pr1) y posteriormente los 3 que se seleccionaron de estos 11 figura 7.13. Los cuales se obtuvieron reduciendo a 32 niveles de gris, con un desplazamiento horizontal de 2, y para un tamaño de la ventana de 5x5.



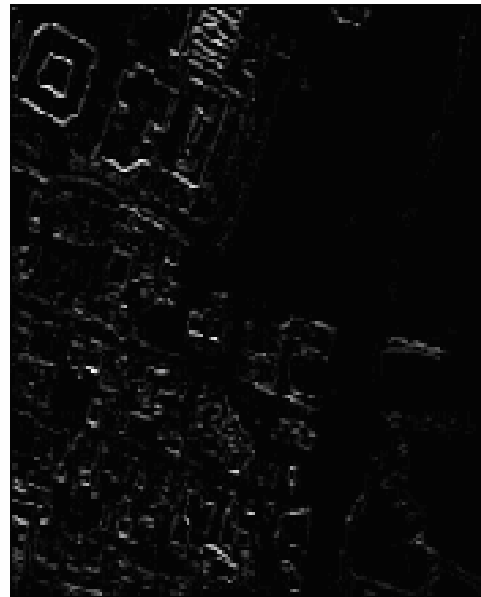
Figura 7.13 Imagen del Distrito Federal (DF1_pr1)

Descriptores de textura obtenidos por matriz de co-ocurrencia 11 en total de estos se seleccionaron 3 para el proceso de fisión de datos. Los parámetros de entrada del programa se muestran a continuación.

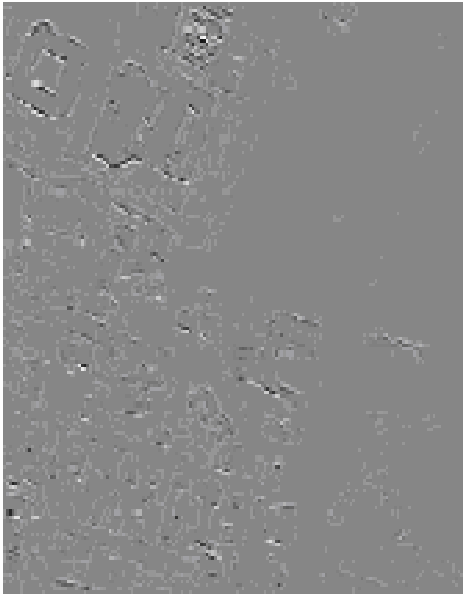
Imagen	No. De niveles de gris reducidos	Desplazamiento Horizontal	Tamaño de la ventana
DF1_pr1	32	2	5x5



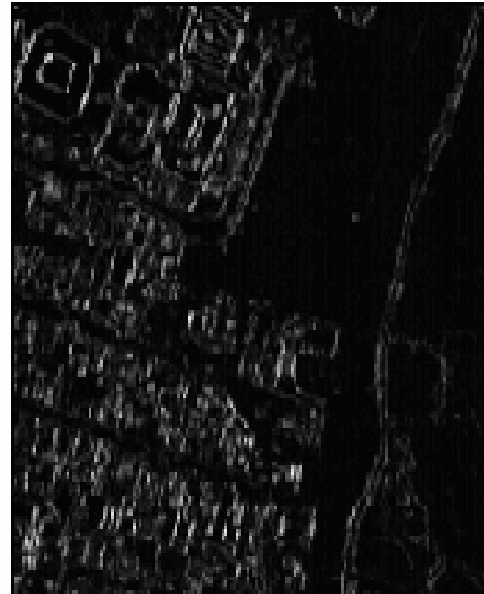
Autocorrelacion



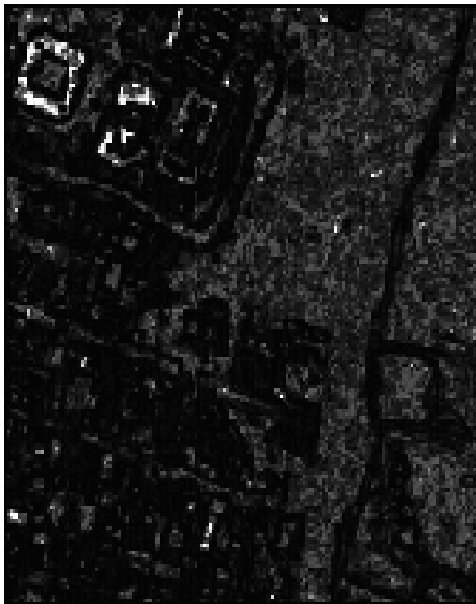
Cluster prominence



Cluster shade



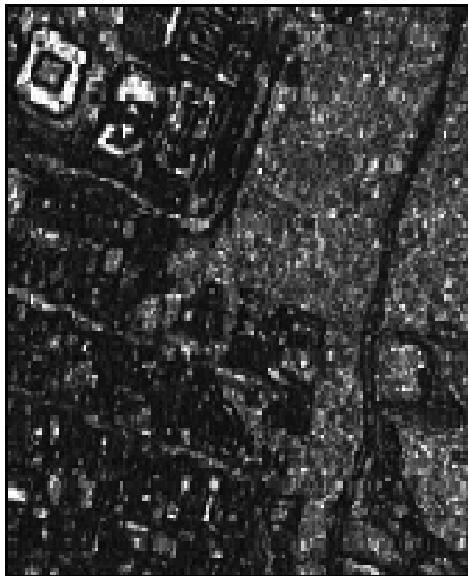
Contraste



Correlacion



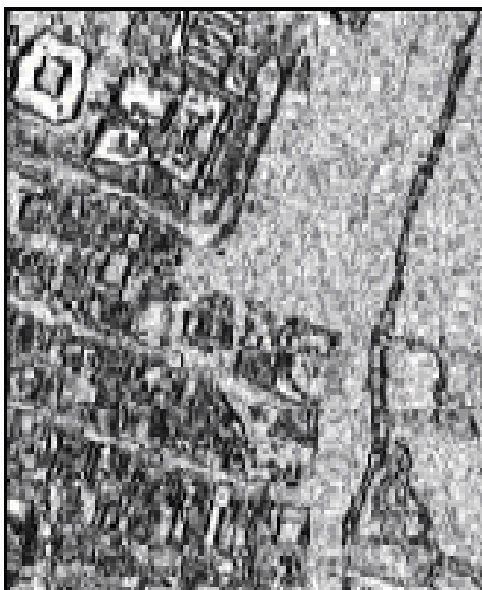
Dissimilarity



Energía



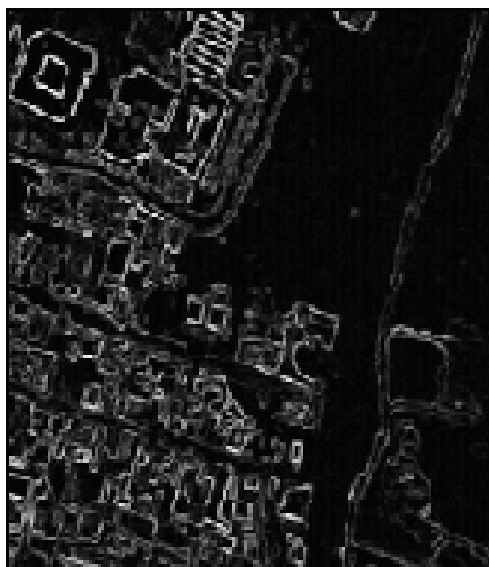
Entropía



Homogeneity



Probabilidad maxima



Varianza

De los cuales los que se seleccionaron fueron los correspondientes a autocorrelación, entropía y homogeneidad, mostrados a continuación.

Homogeneidad



Entropía



Autocorrelación

Para la fusión de datos se tomó en cuenta el factor de confiabilidad dependiendo de la clase que se aprecia mejor en cada imagen de textura con respecto a las demás. Estos factores (λ) se multiplicarán por cada clase a identificar en cada imagen de textura se presentan a continuación.

Para la presente imagen del distrito DF1_pr1 (Figura7.13) sólo se identificaron 3 clases, correspondientes a vegetación, urbano y calles.

Imagen de textura autocorrelación

Clase	Valor del factor de confiabilidad (λ)
Vegetación	0
Urbano	0.5
Calles	1

Imagen de textura homogeneidad

Clase	Valor del factor de confiabilidad (λ)
Vegetación	1
Urbano	0.5
Calles	0.5

Imagen de textura entropía

Clase	Valor del factor de confiabilidad (λ)
Vegetación	1
Urbano	0.5
Calles	1

Los resultados de la fusión de datos se muestran a continuación. Estos resultados son productos del algoritmo de recocido simulado para fusión de datos para diferente número de iteraciones, partiendo de una imagen aleatoria y haciendo un barrido de toda la imagen en un análisis en 8-conexidad, con una temperatura de $T = 2.5$. En estos resultados se muestra la imagen en niveles de gris junto con la de colores falsos, de tal manera que se puedan identificar con mayor facilidad las clases (calles, urbano y vegetación). En las imágenes de las figuras 7.14 a la 7.24 se muestra la secuencia que se va tomando al recocer la imagen a partir de la imagen aleatoria, se muestra que gradualmente a medida que se incrementa el número de iteraciones (20,30,40,60 y 80) la aleatoriedad se va disminuyendo dándonos una imagen final segmentada en clases producto de la fusión de datos

Imagen aleatoria

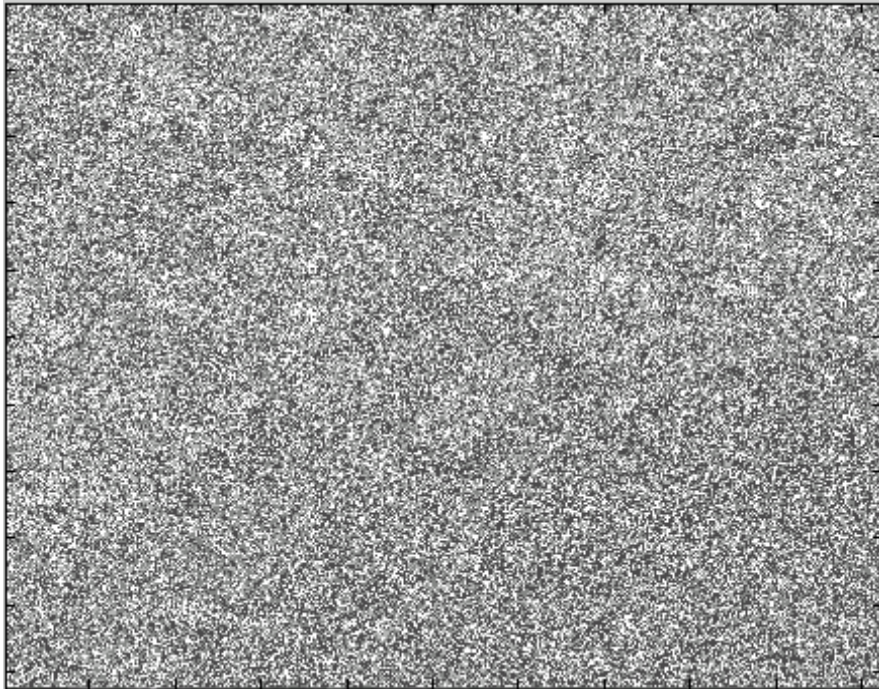


Figura 7.14 Imagen aleatoria

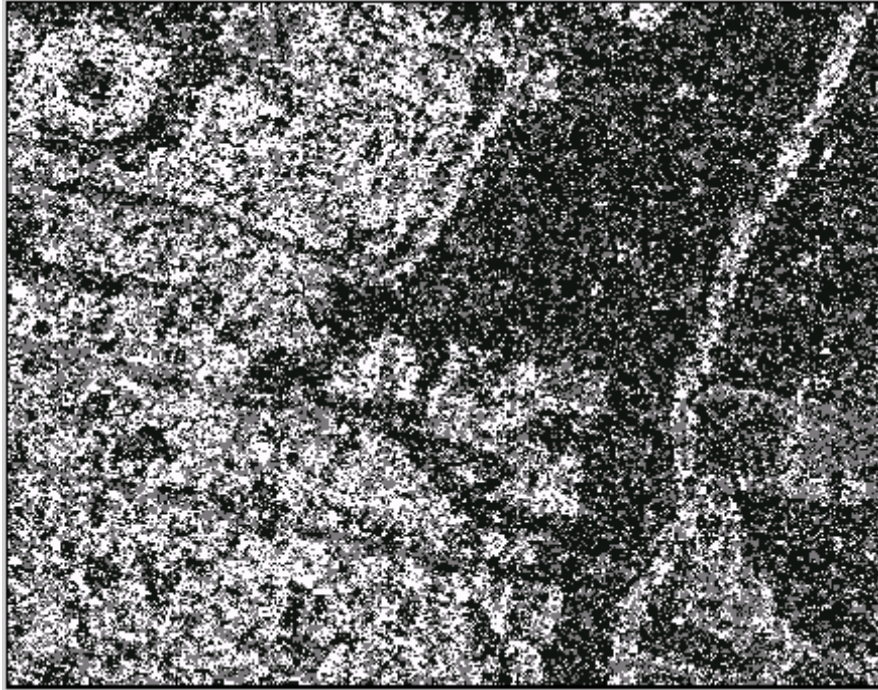


Figura 7.15 Resultado para 20 iteraciones del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad o recociendo la imagen

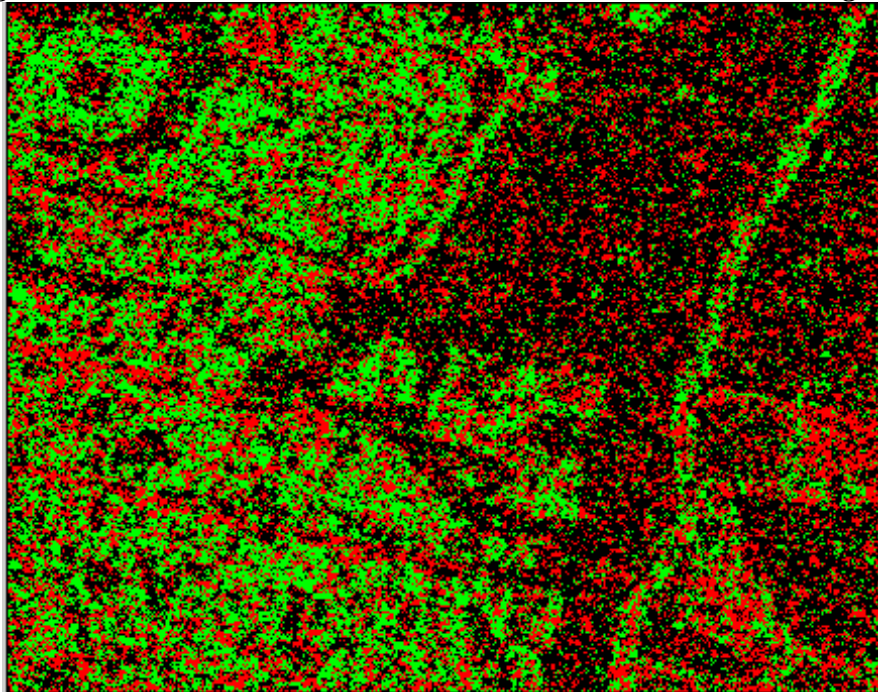


Figura 7.16 Resultado para 20 iteraciones (imagen en colores) falsos del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad o recociendo la imagen de la inicialización aleatoria

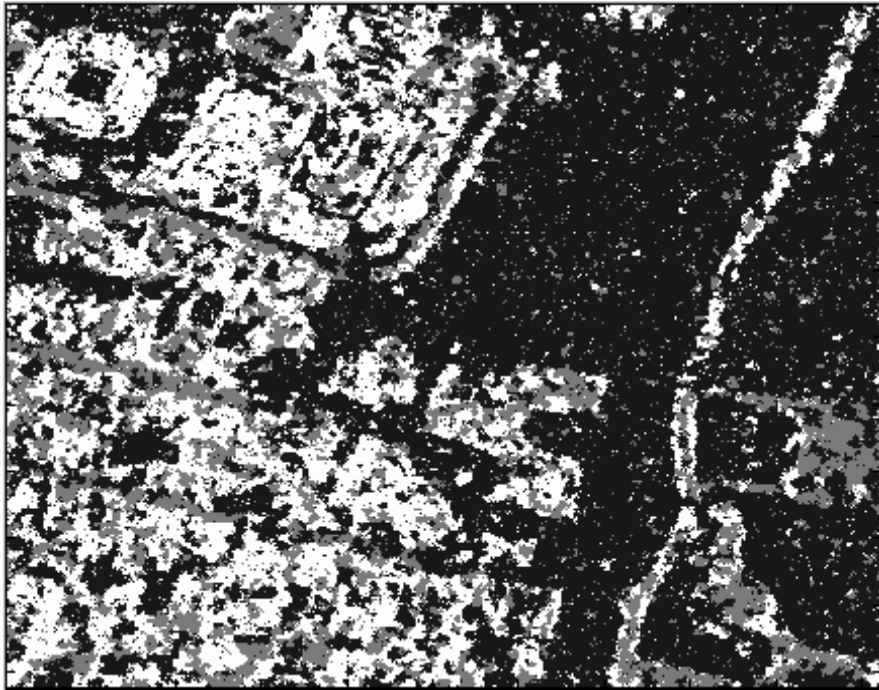


Figura 7.17 Resultado para 30 iteraciones del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad dándonos una mejor imagen

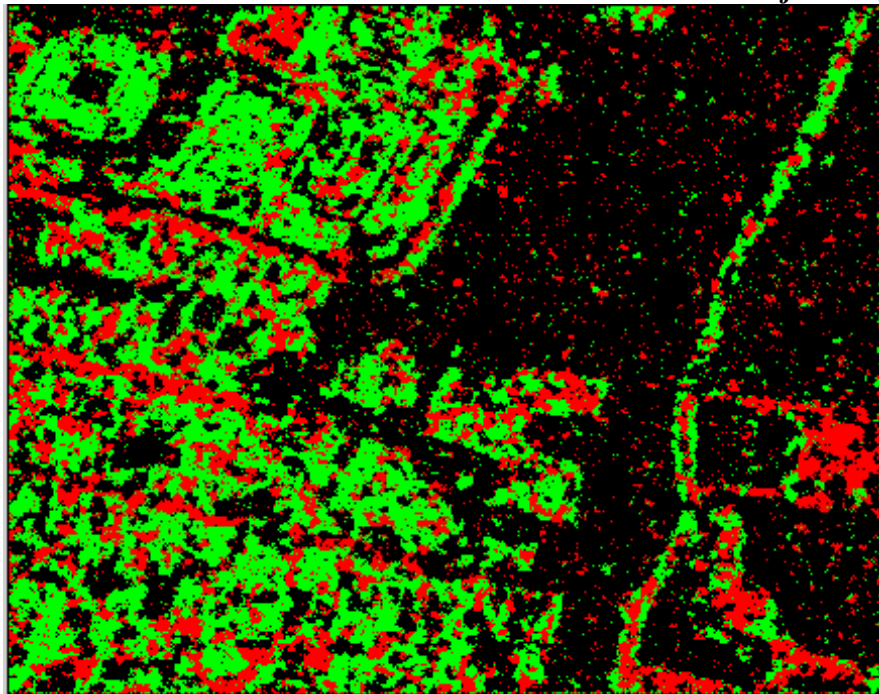


Figura 7.18 Resultado para 30 iteraciones (imagen en colores) falsos del recocido simulado, se muestra que gradualmente va saliendo de la aleatoriedad (recociéndose) dándonos una mejor imagen

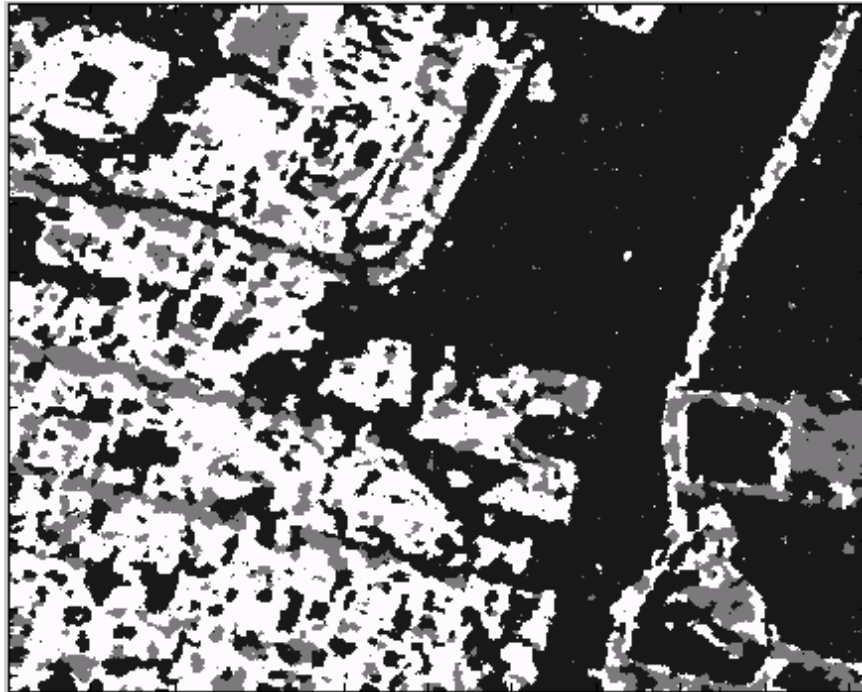


Figura 7.19 Resultado para 40 iteraciones del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han disminuido, dándonos una imagen mas clara segmentada en clases

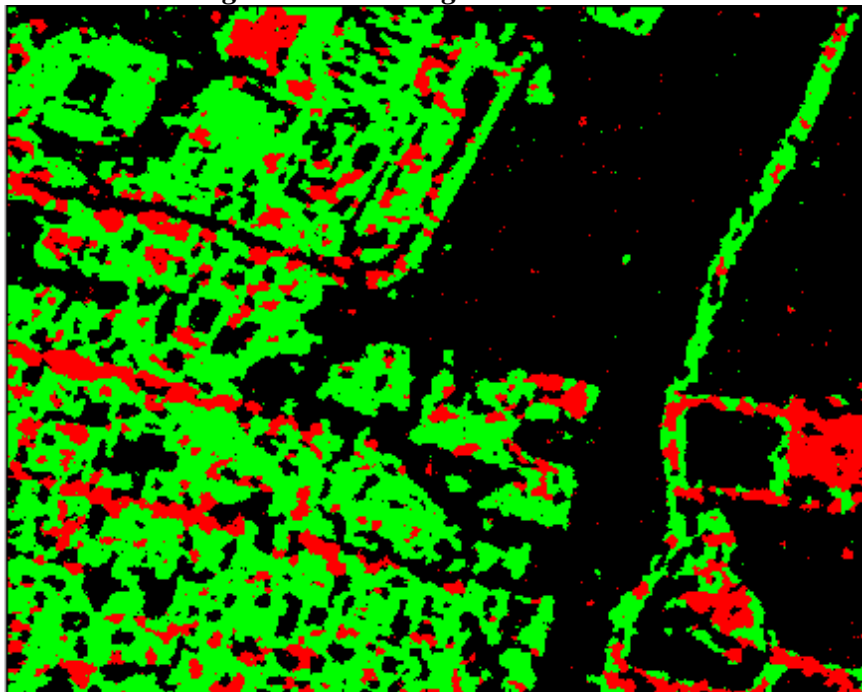


Figura 7.20 Resultado para 40 iteraciones (imagen en colores) falsos del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han disminuido, dándonos una imagen mas clara segmentada en clases

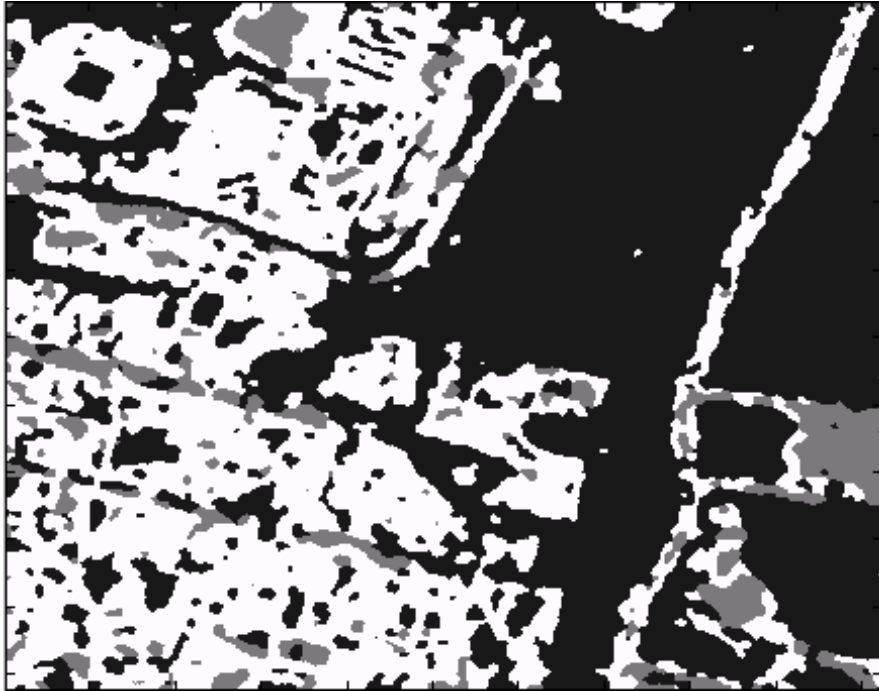


Figura 7.21 Resultado para 60 iteraciones del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases casi han desaparecido en un 100%, visualizando una imagen segmentada en clases pero no precisa

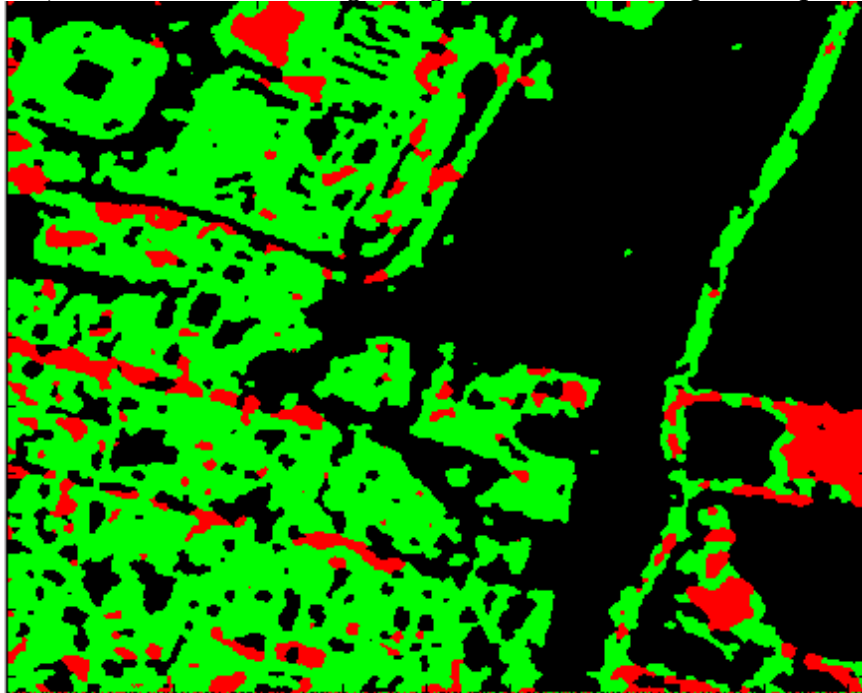


Figura 7.22 Resultado para 60 iteraciones imagen en colores falsos del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases casi han desaparecido en un 100 %, visualizando una imagen segmentada en clases (recocida) pero no precisa en lo que respecta a agrupamiento



Figura 7.23 Resultado para 80 iteraciones del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han desaparecido (imagen recocida) visualizando una imagen segmentada en clases (urbano, vegetación y calles) con un agrupamiento más exacto e identificable

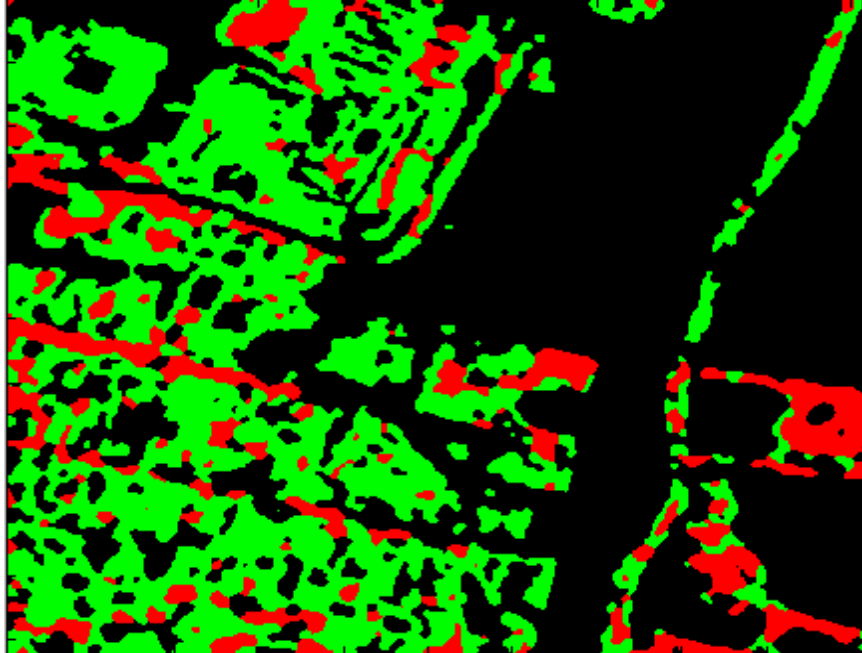


Figura 7.24 Resultado para 80 iteraciones (imagen en colores) falsos del algoritmo de recocido simulado, en esta imagen los puntos producto de la aleatoriedad de clases han desaparecido (imagen recocida) visualizando una imagen segmentada en clases (urbano, vegetación y calles) con un agrupamiento más exacto e identificable

VIII. Conclusiones

El principal objetivo del procesamiento digital de imágenes es manipular y analizar los datos de imágenes capturados por sensores remotos, de tal forma que sea revelada “información” que no pueda ser reconocida inmediatamente en la forma original.

Actualmente existen una gran variedad de técnicas para la segmentación de imágenes, siendo algunas de propósito general y otras de propósito específico. El algoritmo que se presenta en este trabajo permite resolver problemas de estimación en el sentido del MAP (máximo a posteriori).

La teoría de los campos aleatorios de Markov (CAM) permite describir con facilidad: los objetivos de una escena (es decir, el proceso de etiquetamiento de regiones) y el proceso de formación de las imágenes (a través de la estadística de datos). Por el contrario, la principal desventaja de la segmentación markoviana radica en la pérdida de resolución espacial.

Los algoritmos de segmentación generalmente están basados en dos criterios importantes a considerar: es la homogeneidad de la región y la discontinuidad entre regiones disjuntas adyacentes. Generalmente, la imagen segmentada resultante depende de los valores umbral predeterminados y el algoritmo falla a menudo al juntar regiones que deben estar separadas o al separar regiones que no necesitaban ser separadas, porque la información sobre la uniformidad de una región y la discontinuidad entre regiones diferentes no está bien incorporada a los algoritmos. De ahí la importancia de una adecuada integración de la información a los algoritmos a utilizar.

Como el modelo CAM tiene una potente capacidad para integrar varias informaciones visuales, ha sido aplicado a problemas tales como: restauración y segmentación de

imágenes. Debido al gran número de píxeles en los que el CAM debe ser definido, requiere mucho tiempo de computación (costo computacional en la relación tiempo / resultados) para conseguir etiquetas óptimas, lo que hace difícil aplicar el modelo CAM.

El papel de una función de energía en visión basada en minimización es doble: como la medida cuantitativa de la calidad global de la solución y como una guía para la búsqueda de una solución óptima. Dado que cada punto en una imagen puede tomar diferentes valores, es importante usar métodos que ayuden a encontrar una energía mínima para obtener una probabilidad máxima de las posibles soluciones; es decir, para encontrar el valor o valores más probables para cada punto en la imagen.

La teoría de los campos aleatorios de Markov (CAM) proporciona una manera conveniente y constante de modelar entidades dependientes del contexto, tales como píxeles de la imagen y otras características especiales correlacionadas.

La mayoría de los problemas pueden ser considerados como problemas de etiquetado, usando restricciones, debido al conocimiento a priori y a las observaciones. En este caso, la solución óptima es definida por el etiquetado MAP y se calcula minimizando la energía a posteriori.

Para describir enteramente el proceso de adquisición de la imagen se define una probabilidad a posteriori usando la regla de Bayes, a partir de un modelo a priori y un modelo de probabilidad:

El modelo de probabilidad está relacionado con 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 CAM nos proporcionan herramientas para codificar restricciones contextuales en la probabilidad a priori

(probabilidad actual y probabilidad elección). Esta es la principal razón para el uso del modelo CAM.

Los campos de Markov, además de ser un conjunto de métodos para el análisis y la representación de imágenes, están ligados a un método específico de optimización: el recocido simulado. Este asegura la convergencia hacia un óptimo global del problema. En la práctica se utiliza un algoritmo de Metrópolis o modelo de Potts a temperatura variable con el fin de minimizar la energía $U(X/Y)$.

Para poder hacer uso de la analogía del sistema termodinámico en la resolución de problemas matemáticos de optimización (recocido simulado) debemos primero proporcionar los elementos siguientes:

- 1.- Una descripción de las posibles configuraciones del sistema, es decir, una forma de representar una solución del problema de minimización (maximización); generalmente ésto incluye alguna configuración de los parámetros que representan una solución.
- 2.- Un generador de cambios aleatorios a una configuración dada (clase elección); estos cambios proporcionan, típicamente, una nueva solución en el vecindario de la solución actual (clase actual), por ejemplo el cambio en uno de los parámetros.
- 3.- Una función objetivo $E(X)$ cuya minimización es el objetivo del procedimiento.
- 4.- Un parámetro de control T y un plan de “annealing”, el cual indica como debe bajarse de una temperatura a otra; es decir, después de cuantos cambios aleatorios a la configuración es que T es reducida y de cuanto.

El objetivo del presente trabajo de tesis fue desarrollar, a través de métodos de análisis de texturas y técnicas de fusión y fisión de datos, un esquema estadístico para la

segmentación de fotografía aérea de alta resolución a efectos de integrar la información parcial proveniente de diversas fuentes. La teoría de los campos aleatorios de Markov facilitan la integración de datos parciales en un entorno de análisis contextual.

Se propuso una metodología llamada fisión y fusión para segmentación de imágenes, la cual está compuesta en 2 etapas. El paso de fisión consiste en obtener estimadores texturales, llamados descriptores de textura, obtenidos por el método de matriz de co-ocurrencia de donde se seleccionarán las clases que se distingan lo mejor posible. Por otro lado, el paso de fisión toma las correspondientes funciones de energía de cada imagen respecto a cada una de sus clases, las multiplica por un factor de confiabilidad que nos dice la calidad de la clase con respecto a las demás en los captors. Posteriormente se suma las respectivas funciones de cada captor respecto a cada media de clase multiplicada por su factor de confiabilidad y divirlas entre una normalización de sus factores. Una vez que se realizó lo anterior se prosigue con el modelo de Potts en un análisis en 8 conexidad, comenzando con una inicialización aleatoria en el rango del número de clases y terminando con la imagen segmentada.

Por lo tanto, el propósito del esquema de fisión-fusión es el de solucionar el problema de que en la imagen original los modos están mezclados de alguna manera tal que hacen difícil la obtención de un buen resultado, mientras que en los captors no están mezclados por lo que la forma del histograma es mejor. Con la fusión de datos lo que se pretende también es juntar informaciones parciales para obtener un resultado más completo. Por ende con la fusión de datos se llegó a una aproximación en la solución al problema de los modos mezclados en la imagen original.

Una de las principales ventajas de este esquema de fisión-fusión es su sencillez, que permite la obtención de datos de informaciones parciales provenientes de la imagen original para obtener un mejor resultado. Por otro lado una desventaja de este esquema de fisión-fusión es que en las imágenes de textura (captors) la información no siempre es

muy buena, de allí que se utilice el factor de confiabilidad para la discriminación de cada media de clase, restándole calidad al proceso de segmentación.

Cabe agregar que una de las limitantes de las imágenes en blanco y negro es precisamente por los tonos de gris tan parecidos que hacen que sea más difícil la identificación de clases en las imágenes.

El esquema de fisión y fusión de datos del presente trabajo de tesis llega a un resultado óptimo y cumple con los objetivos planteados en ésta, ya que al fusionar sólo las 3 imágenes de textura (homogeneidad, auto correlación y entropía) en las cuales se seleccionaron las clases de urbano, calles y vegetación calificadas por un factor de confiabilidad hace que el resultado sea más exacto al pasar por el proceso de fusión de datos que el solo trabajar con la imagen original seleccionando las clases en ésta ya que en la imagen original los modos se encuentran mezclados lo cual se reduce al emplear las imágenes de textura. El resultado de la fusión de datos llega a una imagen segmentada en donde se aprecian solo las clases seleccionadas previamente.

En el resultado de fusión de las 3 imágenes de textura provenientes de la imagen DF1_izq Figura 8.1 la mayoría de las clases se muestran en el lugar en donde en la imagen original se aprecia lo urbano la vegetación y las calles, mientras que algunas otras se encuentran un poco dispersas en la imagen o en lugares en donde no deberían aparecer calificando al esquema de fisión y fusión como de bueno pero no exacto.

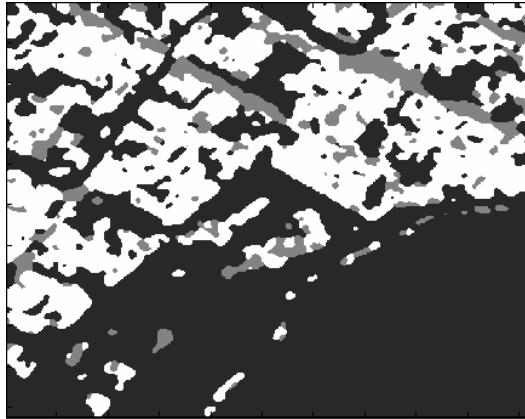


Figura 8.1 Resultado de la fusión de datos para las 3 imágenes de textura provenientes de la imagen DF1_izq.

De la misma manera el resultado de fusión de las 3 imágenes de textura provenientes de la imagen DF1_pr1 Figura 8.2 la mayoría de las clases se muestran en el lugar en donde en la imagen originales aprecia lo urbano la vegetación y las calles, mientras que algunas otras se encuentran un poco dispersas en la imagen o en lugares en donde no deberían aparecer invadiendo zonas pertenecientes a otras clases, calificando de nueva cuenta al esquema como de bueno pero no exacto.

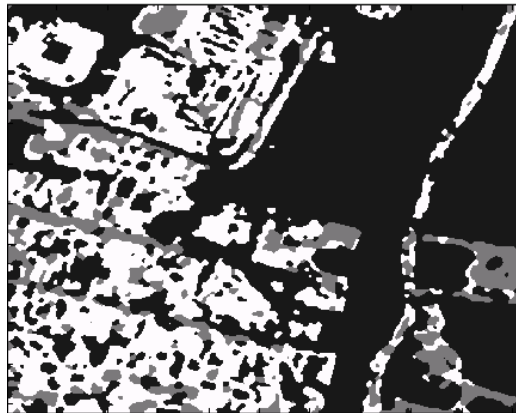


Figura 8.2 Resultado de la fusión de datos para las 3 imágenes de textura provenientes de la imagen DF1_pr1.

IX. Referencias

- [1] Chuvieco, Emilio. “**Fundamentos y teledetección espacial**” 3° edición , ed. RIALP S.A. Madrid 1996.
- [2] Morales Avila David. “**Aproximación estadística para la segmentación de fotografía aérea de alta resolución**”. Tesis para obtener el título de Ingeniero en computación, UNAM Facultad de Ingeniería, 2004.
- [3] Descombes Xavier. “**A fisió n and fusion Markovian approach for multi-channel segmentation**”. Département images, Télécom Paris, 46 rue Barrault, France.
- [4] Fukanga , Keinosuke “**Introduction to statistical pattern recognition**” 2° Edición, ed. Academic press, EUA. 1990.
- [5] Maravall Darío, “**Reconocimiento de formas y visión artificial**”, Addison-Wesley Iberoamericana, Ra-ma España 1994.
- [6] Méndez, R. “**Transformada polinomial y campos aleatorios de Markov: una aplicación a la detección de líneas de costa en imágenes de radar de apertura sintética**”. Tesis para obtener al título de Ingeniero en Telecomunicaciones, UNAM .México 1997.
- [7] Mujol, R “**Aplicación de los modelos de campos de Markov en visión artificial**” Revista electrónica de visión por computador, Departamento de ciencia de la computación e inteligencia artificial.
- [8] Clausi David A. **Rapid determination of co-occurrence textura features** IEEE transactions on Acoustics, Speech and Signal Processing; vol. 38 no.9, 1990; pp 1595-1606

Direcciones electrónicas

- [9] Mapas de la ciudad de México

www.guiaroji.com.mx

Referencias

www.inegi.gob.mx

[10] Satélite Ikonos

www.spaceimageurope.com/products_layout.htm
www.aeroterra.com/HTMs/ikonos.htm

[11] Satélite Landsat

www.aeroterra.com/HTMs/landsat7ETM.htm
www.cct.rncan.gc.ca/ccrs/tekrd/satsens/sats/landsat.html

[12] Satélite Spot

www.spotimage.fr/home/system/introsat/orbit/welcome.htm
www.ccrs.rncan.gc.ca/ccrs/tekrd/satsens/sats/spot.html

[13] Teledetección

www.uco.es/docin/invest/ayudas/apoyo_invest/scit/teledeteccion.htm
www.ija.csic.es/tele/rsweb/sensor.htm

[14] Satélite Envisat

http://www.esa.int/esaEO/SEMWYN2VQUD_index_0_m.html
<http://www.aerospace-technology.com/projects/envisat/>

[15] Satélites ERS1 y 2

<http://www.gisdevelopment.net/aars/acrs/2000/ts14/sari0004.shtml>

[16] Satélite Aster

<http://asterweb.jpl.nasa.gov/>
<http://edcdaac.usgs.gov/aster/asteroverview.asp>

[17] Radar de apertura sintética (SAR)

<http://www.sandia.gov/RADAR/sar.html>
http://www.tdx.cesca.es/TESIS_UPC/AVAILABLE/TDX-0905102-135541/2_3_Satelites_con_SAR.pdf

APÉNDICE A

Teorema de Bayes

El teorema de Bayes admite un manejo subjetivo de la probabilidad, esto permite emitir juicios de probabilidad sobre la validez de una hipótesis H . La probabilidad a priori de una hipótesis $P(H)$, es la que se fija o sugiere antes de tener en cuenta los datos (por eso es a priori)

Una vez realizada una observación o experimento, se incorpora la evidencia que aportan los datos al análisis a través del teorema y se obtiene una probabilidad a a. Una vez observados los datos, el teorema “devuelve” también una distribución que ha de interpretarse como la nueva percepción, ahora “actualizada” por los datos. El teorema se expresa de la siguiente forma:

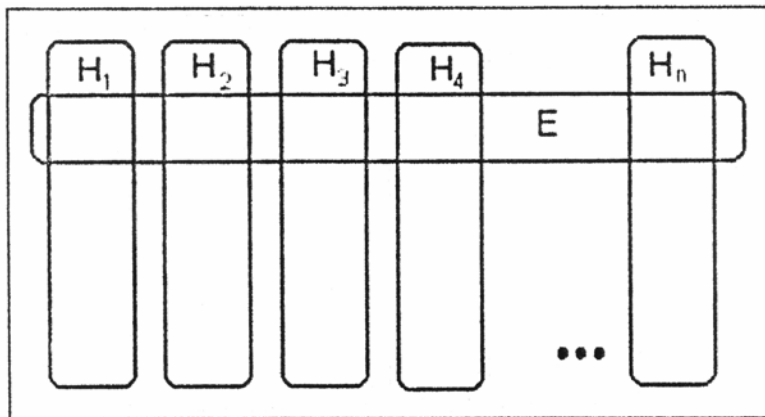


Figura A.1 Representación de los conjuntos $H_i (i=1, \dots, n)$ y E . H_i son mutuamente excluyentes y E no es vacío.

La figura representa H_i como eventos mutuamente excluyentes y cuya unión es el espacio muestral Q . E es un evento en Q con una probabilidad no cero de ocurrencia; además, la probabilidad de obtener E es:

$$P(E) = E \cap H_1 + E \cap H_2 + \dots + E \cap H_n = \bigcup_{i=1}^{i=n} E \cap H_i$$

Ecuación A.1

Considerando la estimación de probabilidades conjuntas $P(AB)$ cuando los eventos son condicionales, se tiene la siguiente equivalencia:

$$P(H_i | E) = P(H_i)P(E | H_i) = P(E)P(H_i | E)$$

Ecuación A.2

Por lo tanto:

$$P(H_i | E) = \frac{P(H_i)P(E | H_i)}{P(E)}$$

Ecuación A.3

Se puede reformular la parte final de la ecuación A.1 como:

$$P(E) = P\left(\bigcup_{i=1}^{i=n} E \cap H_i\right) = \sum_{i=1}^{i=n} P(E \cap H_i) = \sum_i^n P(E | H_i)P(H_i)$$

Ecuación A.4

Lo anterior es válido porque los $E H_i$ son mutuamente excluyentes. El término del denominador en la ecuación A.3 se puede sustituir por su valor en 4.4 y se tiene:

$$P(H_i | E) = \frac{P(E | H_i)P(H_i)}{\sum_i^n P(E | H_i)P(H_i)}$$

Ecuación A.5

Esta fórmula es conocida como: **El teorema de Bayes**

Dado que:

$$\sum_i^n P(E | H_i)P(H_i) = P(E)$$

Ecuación A.6

Es la probabilidad total del evento | píxel.

Tomando en cuenta la ecuación esta última expresión puede describirse como:

$$P(H_i | E) = \frac{P(E | H_i)P(H_i)}{P(E)}$$

Ecuación A.7

Esta fórmula es conocida como: **La regla de Bayes**

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

P(H_i | E) : Probabilidad de que un vector de características E pertenezca a la clase H_i. Esta representa a la probabilidad a posteriori

P(E | H_i) : Probabilidad de que dada H_i, el valor de la variable aleatoria sea precisamente E. Dicho en otras palabras, es la función densidad de probabilidad de la clase H_i considerada como una variable aleatoria. Esta representa a la probabilidad conjunta o condicional de E dado H_i. Esta se evalúa mediante las funciones de energía U(E | H_i)

P(H_i) : Probabilidad a priori de que se presente un elemento de la clase H_i, esta se evalúa mediante el modelo de Potts.

P(E) : Probabilidad a priori de que se presente un objeto a clasificar con un vector de características igual a E (considerando un vector numérico concreto o constante de normalización).

Este último elemento puede despreciarse, ya que presenta el mismo valor para un conjunto de n clases, H₁, H₂, ..., H_n, compitiendo por el vector E a clasificar. Se trata, pues de un factor de escala que se elimina por no incorporar información discriminante. Dándonos por resultado la aproximación de la regla de Bayes.

$$P\langle H_i | E \rangle \doteq P\langle E | H_i \rangle P(H_i)$$

Ecuación A.8

Clasificador Bayesiano Simple

Estimar la probabilidad: $P(E|H)$ es complejo, pero se simplifica si se considera que los atributos son independientes dada la hipótesis:

$$P(E_1, E_2, \dots, E_N | H) = P(E_1 | H) P(E_2 | H) \dots P(E_N | H)$$

Ecuación A.9

Por lo que la probabilidad de la hipótesis dada la evidencia puede estimarse como:

$$P\langle H | E_1, E_2, \dots, E_N \rangle = \frac{P(H)P\langle E_1 | H \rangle P\langle E_2 | H \rangle \dots P\langle E_N | H \rangle}{P(E)}$$

Ecuación A.10

Esto se conoce como el Clasificador Bayesiano Simple. Como se mencionó anteriormente, no es necesario calcular el denominador:

$$P(H | E_1, E_2, \dots, E_N) \sim P(H) P(E_1 | H) P(E_2 | H) \dots P(E_N | H)$$

Ecuación A.11

- $P(H)$ se conoce como la probabilidad a priori
- $P(E_i | H)$ es la probabilidad de los atributos dada la hipótesis (verosimilitud), y $P(H | E_1, E_2, \dots, E_n)$ es la probabilidad a posteriori.

El modelo de probabilidad está relacionado con 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 CAM nos proporcionan

Apéndice

herramientas para codificar restricciones contextuales en la probabilidad a priori. Esta es la principal razón para el uso del modelo CAM.

A través de aplicar los campos aleatorios de Markov a la aproximación de la regla de Bayes se llega a la siguiente expresión:

$$U(Y|X) \doteq U(X|Y) + U(Y)$$

Ecuación A.12

Esta expresión se trata de minimizar con recocido simulado

APÉNDICE B

Construcción de la función de energía

La función de energía se definió como:

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

Ecuación A.13

Se considera entonces que $V_c(f)$ es el potencial del clique, y que $U(f)$ es la suma de todas las funciones potenciales de los cliques, $V_c(f)$ para todas los cliques posibles, donde el valor $V_c(f)$ depende de la configuración local del clique c , considerándose la distribución Gaussiana como un caso especial de la familia de distribuciones de Gibbs.

La función de energía permite la clasificación de cada píxel de la imagen original, esto es, decide la pertenencia de un píxel a una clase específica definida dentro de la imagen de la segmentación. La función de energía utilizada para la segmentación Markoviana se especifica gráficamente en la Figura A.1. A continuación se describe el procedimiento de construcción de las funciones de energía $U(X/w_i)$.

El nivel de gris de cada píxel se proyecta sobre la función a la que pertenezca el píxel analizado, es decir, la clase w_i . El número de funciones está determinado por el número de clases definidas previamente.

Para la construcción de las funciones de energía se definen las siguientes variables:

y: Es la energía condicional.

x: Es el píxel analizado.

ma: Es la media de la clase analizada

S: Es la diferencia de las medias de clase. Dependiendo del caso, $S = m(\text{sig}) - m(\text{act})$ ó

$S = m(\text{act}) - m(\text{ant})$.

m(ant) =es la media de clase anterior a la actual

Apéndice

$m(\text{act})$ =es la media de la clase actual

$m(\text{sig})$ =es la media de la clase siguiente a la actual.

Pendiente positiva (clase actual)	Pendiente negativa (clase siguiente)
Función (A)	Función (B)
$y = \frac{2}{S}x - \frac{2}{S}m_a - \frac{1}{2}$	$y = -\frac{2}{S}x + \frac{2}{S}m_a + \frac{3}{2}$

- Si el nivel de gris del píxel analizado (x) está en el rango: $0 < x < m_o + (S/4)$, la energía condicional para la clase w_o (clase inicial) es de 0. Para las demás clases la energía será de 1.
- Si el nivel de gris del píxel analizado (x) está en el rango: $x > m_n$, la energía condicional para la clase w_n (última clase) es de 0. Para las demás clases la energía será de 1.
- Si el nivel de gris del píxel analizado (x) no está en alguno de los casos anteriores se realiza la diferencia entre las medias de las clases de la siguiente manera: $S = m_s - m_a$, donde m_s es la media de la clase siguiente y m_a es la media de la clase actual. El resultado de esta diferencia se divide entre 4, es decir, $S/4$.
- Si el nivel de gris del píxel analizado (x) está en el rango: $m_a < x < m_a + (S/4)$, la energía condicional para la clase w_a (clase actual) es de 0. Para las demás clases la energía será de 1.
- Si el nivel de gris del píxel analizado (x) está en el rango: $m_a + (S/4) < x < m_a + (3/4)S$ y pertenece a la clase inicial w_o , la energía condicional para la clase w_o (clase inicial) se determina a través de la función (B), para ello se utilizará la diferencia de medias de la clase actual menos la clase anterior, es decir: $S = m(\text{act}) - m(\text{ant})$.

Para cualquier otra clase, si el píxel analizado (x) está en el rango $m_{ant} + (S/4) < x < m(\text{ant}) + (3/4)S$ se utiliza la función (B) y la diferencia de medias $S = m(\text{act}) - m(\text{ant})$. Si el

Apéndice

píxel analizado (x) está en el rango $m(\text{act}) + (S/4) < x < m(\text{act}) + (3/4)S$ se utiliza la función (A) y la diferencia de medias $S = m(\text{sig}) - m(\text{act})$. Las graficas se muestran en la figura A.2

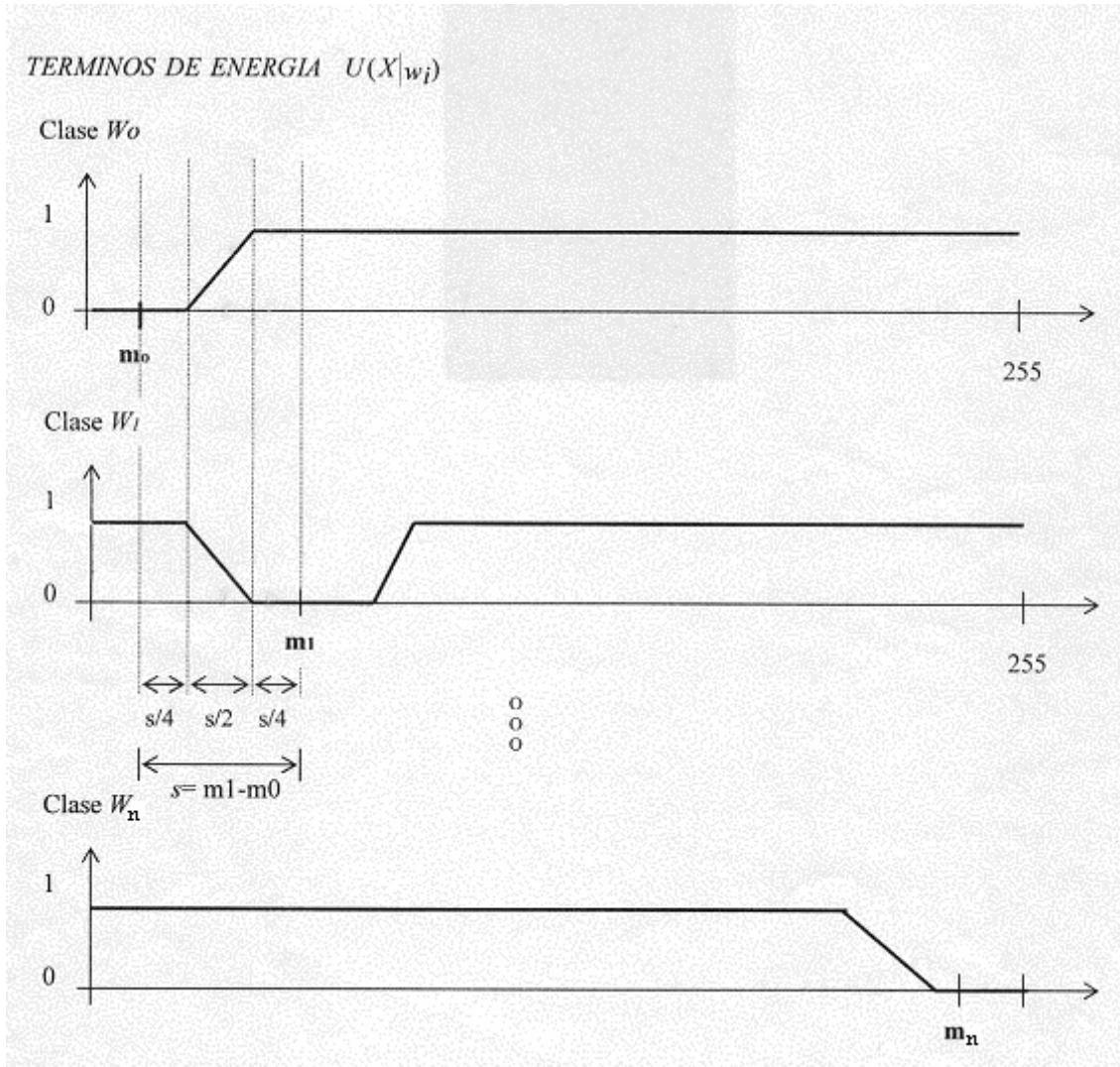


Figura A.2 Funciones de energía para cada media de clase

La energía condicional obtenida anteriormente se suma con la probabilidad de la clase en cuestión esto se aplica en el modelo de Potts. Esto permite calcular la variación de la energía local asociada al site r (píxel analizado).

Apéndice

Los métodos bayesianos asociados 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 asegura la convergencia hacia el mínimo global y evita la convergencia a los mínimos locales.

APÉNDICE C

Algoritmo de Recocido Simulado

Sea X la imagen original (en niveles de gris) e Y la imagen segmentada (por ejemplo, en cuatro clases), la simulación toma como datos iniciales a X y a la imagen de clases aleatoria originada en la inicialización $Y(0)$. La figura A.3 muestra la esquematización del algoritmo de simulación, para el caso de n iteraciones:

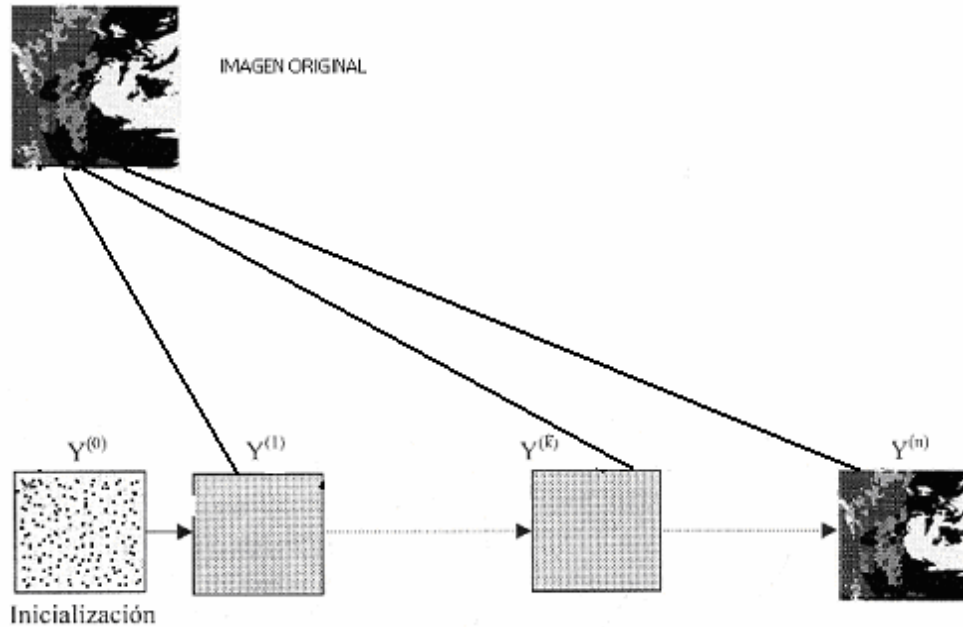


Figura A.3 Obtención de la imagen fusionada a partir de una inicialización aleatoria

Para el caso de la energía $U(Y)$ se considera el modelo de Potts, el cual se aplica en vecindades en 8-conexidad.

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 $[\xi]$).
- Escoger una temperatura inicial $T(0)$ suficientemente elevada (por ejemplo, $T(0)=2.5$).

2. Hacer un contador con un número específico de iteraciones $N = 1 \dots n$

2.2 Barrer secuencialmente la imagen

En cada etapa (n)

- Escoger un site: $r=s(n)$, su valor es $Y_r(n)$ y su configuración de vecindad es $V_r(n)$
- Efectuar el sorteo de una variable aleatoria (clase “elección”) con distribución uniforme en el rango del número de clases $[\xi]$ (solo se varia elección y se compara contra actual).

2.1 Criterio de metrópolis y modelo de Potts

- Se utiliza el algoritmo de metrópolis de temperatura variable con el fin de minimizar la energía $U(\mathbf{X}|\mathbf{Y})$ ó $U(\mathbf{X}|\mathbf{W})$.

- Para el caso de la energía $U(\mathbf{Y})$ ó $U(\mathbf{W}_n)$ se considera el modelo de Potts (visto en el capítulo V), el cual se aplica en vecindades en 8-conexidad.

a	a	a
a	x	a
a	a	a

Clases	Probabilidades
“actual”	<u>Probabilidad actual = 0</u>
“elección”	<u>Probabilidad elección = 0</u>

Análisis en 8-conexidad

- Si “actual” = “x”: $\text{prob_actual} = \text{prob_actual} - \beta[\text{actual}]$
- Si “elección” = “x”: $\text{prob_elección} = \text{prob_elección} - \beta[\text{elección}]$

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

- $\text{prob_actual} = \text{prob_actual} + \text{potencial}[\text{actual}][i][j]$
- $\text{prob_elección} = \text{prob_elección} + \text{potencial}[\text{elección}][i][j]$

donde $\text{potencial}[i][j]$ representa la energía condicional $U(X/Y)$ e $[i][j]$ son las coordenadas del píxel analizado.

- Si $\text{prob_elección} \leq \text{prob_actual}$: $Y[i][j] = \text{“elección”}$

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 de $[0,1]$
- Si $\xi_2 < p$ entonces $Y[i][j] = \text{“elección”}$
- Si no $Y[i][j] = \text{“actual”}$

3. Regresar al punto 2.1 hasta terminar un barrido

4. Decrementar la temperatura

$$T(n) = T(n-1) * 0.95$$

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

APÉNDICE D

Programa que simula el algoritmo de recocido simulado creado en Matlab.

```
%Este programa procesa el algoritmo de Recocido Simulado
%para imagenes de extension tiff,gif,bmp y raw
%Para usarlo solo de el nombre de la imagen
%oprime enter y despues su extension.
%Solo recuerde tener a la imagen en el mismo
%subdirectorio que el programa.
clear; % Borra todas las variables en memoria
close all; % Cierra todas las ventanas activas
tic
fprintf(' Implementacion del Algoritmo de Recocido Simulado: ');
nombre=input('\n\n De el nombre de la imagen: ','s');
extension=input('\n\n extension(bmp,tif,gif,raw): ','s');
k=strcmp(extension,'tif');
k2=strcmp(extension,'raw');
if k2==1
%Algoritmo para leer imagenes con extension raw
ren=input(' Numero de Renglones: ');
% Indicar el número de renglones de la imagen
col=input(' Numero de Columnas: ');
% Indicar el número de columnas de la imagen
ima=strcat(nombre, '.',extension)
x=fopen(ima,'r');%Se abre la imagen cuyo nombre y extension son dados por "nombre" y "extension"
y=fread(x,[ren,col]);
% Lee la imagen secuencialmente y la asigna a la variable y
x=y';
% La imagen raw debe rotarse para su correcta visualización
clear y ;
% Se borran las variables que ya no se van a utilizar
map=gray(256);%Este es el mapa de color
else
[x,map]=imread(nombre,extension);% Si es otra extension se abre de forma clasica
if k==1
    map=gray(256);
else
    map=map;
end
end
```

Apéndice

```
xa=double(x)+1;
[ren,col]=size(xa);
figure(1)
colormap(map)
image(xa)
title('IMAGEN ORIGINAL');
r=input('\n\nOprima enter','s');
mapa=[0 0 0 ; 1 0 0 ; 0 1 0 ; 0 0 1 ; 1 1 0 ; 1 0 1 ; 0 1 1 ; 1 1 1];

fprintf('\n\n Algoritmo de Recocido Simulado: ');
% Despliega en pantalla
numclas=input('\n\nDe el numero de Clases para hacer la segmentacion: ');
% Pide el número de clases en que se segmentará la imagen
%(se recomienda segmentar entre 3 y 8 clases)
fprintf('\n\nEscoga los puntos para obtener las medias: ');

dame1=ginput(numclas);
dame2=round(dame1);
al=dame2(1:numclas,1);
an=dame2(1:numclas,2);
dame=diag(xa(al,an));
grismin=min(dame);
grismax=max(dame);
h=input('\n\nOprima enter');
```

Apéndice

```
% Variable en la que se almacenan los centros de clase
medias=zeros(1,numclas);
% Variable para almacenar las medias de clase
% (al final de cada iteración)
% Tamano de la ventana y obtencion de la media

for i=1:numclas
    d=(xa(an(i)-2:an(i)+2,al(i)-2:al(i)+2));
    a=mean(d);
    b=sum(a,2)/5;
    medias(i)=round(b);
end

% Definicion de las variables a utilizar

pXW1=zeros(ren,col);
pXW2=zeros(ren,col);
pXW3=zeros(ren,col);
pXW4=zeros(ren,col);
pXW5=zeros(ren,col);
pXW6=zeros(ren,col);
prob_actual=0;
prob_eleccion=0;
prob=0;
prob=0;
sk=zeros(1,numclas-1);

for w=1:numclas-1
    sk(w)=medias(w+1)-medias(w);
end

% Creacion de las funciones de energia de cada media de clase
```

Apéndice

```
for w=1:numclas

    a=w;
    b=1;
    for i=1:ren
        for j=1:col
            if a-b==0
                if xa(i,j)>0 & xa(i,j) < medias(w)+sk(w)/4
                    pXW1(i,j)=0;

                    elseif xa(i,j)>=medias(w)+sk(w)/4 & xa(i,j)<medias(w)+3*sk(w)/4
                    pXW1(i,j)=(2/sk(w))*(xa(i,j))-(2/sk(w))*medias(w)-1/2;

                    elseif xa(i,j)>=medias(w)+3*sk(w)/4 & xa(i,j)<=256
                    pXW1(i,j)=1;

                    else
                        c=0;
                    end
                end

            if a~=1 & a~=numclas

                if xa(i,j)>0 & xa(i,j)< medias(w-1)+sk(w-1)/4
                    if w==2
                        pXW2(i,j)=1;
                    elseif w==3
                        pXW3(i,j)=1;
                    elseif w==4
                        pXW4(i,j)=1;
                    elseif w==5
                        pXW5(i,j)=1;
                    else w==6
                        pXW6(i,j)=1;
                    end

                    elseif xa(i,j)>=medias(w-1)+sk(w-1)/4 & xa(i,j)< medias(w)-sk(w-1)/4
                    if w==2
                        pXW2(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
                    elseif w==3
                        pXW3(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
                    elseif w==4
                        pXW4(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
                    elseif w==5
                        pXW5(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
                    else w==6
                        pXW6(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
                    end
                end
            end
        end
    end
end
```

```
elseif xa(i,j)>=medias(w)-sk(w-1)/4 & xa(i,j)<medias(w)+sk(w)/4
  if w==2
    pXW2(i,j)=0;
  elseif w==3
    pXW3(i,j)=0;
  elseif w==4
    pXW4(i,j)=0;
  elseif w==5
    pXW5(i,j)=0;
  else w==6
    pXW6(i,j)=0;
  end

elseif xa(i,j)>=medias(w)+sk(w)/4 & xa(i,j)<medias(w)+3*sk(w)/4
  if w==2
    pXW2(i,j)=(2/sk(w))*(xa(i,j))-(2/sk(w))*medias(w)-1/2;
  elseif w==3
    pXW3(i,j)=(2/sk(w))*(xa(i,j))-(2/sk(w))*medias(w)-1/2;
  elseif w==4
    pXW4(i,j)=(2/sk(w))*(xa(i,j))-(2/sk(w))*medias(w)-1/2;
  elseif w==5
    pXW5(i,j)=(2/sk(w))*(xa(i,j))-(2/sk(w))*medias(w)-1/2;
  else w==6
    pXW6(i,j)=(2/sk(w))*(xa(i,j))-(2/sk(w))*medias(w)-1/2;
  end

elseif xa(i,j)>=medias(w)+3*sk(w)/4 & xa(i,j)<=256
  if w==2
    pXW2(i,j)=1;
  elseif w==3
    pXW3(i,j)=1;
  elseif w==4
    pXW4(i,j)=1;
  elseif w==5
    pXW5(i,j)=1;
  else w==6
    pXW6(i,j)=1;
  end

else
  d=0;
end
end
```

```
if a-numclas==0

    if xa(i,j)>0 & xa(i,j)< medias(w-1)+sk(w-1)/4
        if w==2
            pXW2(i,j)=1;
        elseif w==3
            pXW3(i,j)=1;
        elseif w==4
            pXW4(i,j)=1;
        elseif w==5
            pXW5(i,j)=1;
        else w==6
            pXW6(i,j)=1;
        end

    elseif xa(i,j)>=medias(w-1)+sk(w-1)/4 & xa(i,j)<medias(w)-sk(w-1)/4
        if w==2
            pXW2(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
        elseif w==3
            pXW3(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
        elseif w==4
            pXW4(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
        elseif w==5
            pXW5(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
        else w==6
            pXW6(i,j)=-(2/sk(w-1))*(xa(i,j))+(2/sk(w-1))*medias(w-1)+3/2;
        end

    elseif xa(i,j)>=medias(w)- sk(w-1)/4 & xa(i,j)<=256
        if w==2
            pXW2(i,j)=0;
        elseif w==3
            pXW3(i,j)=0;
        elseif w==4
            pXW4(i,j)=0;
        elseif w==5
            pXW5(i,j)=0;
        else w==6
            pXW6(i,j)=0;
        end

    else
        e=0;
    end
end

end
end
a=w;
end
```

Apéndice

```
% Criterio de Metropolis y modelo de Potts

% Definicion de la matriz aleatoria en el rango del numero de clases que nos dara la inicializacion
y=unidrnd(numclas,512,512);

%temperatura inicial
t=2.5;

for n=1:20

    for i=2:ren-1
        for j=2:col-1
            prob_actual=0;
            prob_eleccion=0;
            actual=y(i,j);
            x1=y(i-1,j-1);    x2=y(i-1,j);    x3=y(i-1,j+1);
            x4=y(i,j-1);      x5=y(i,j+1);
            x6=y(i+1,j-1);    x7=y(i+1,j);    x8=y(i+1,j+1);
            eleccion=ceil(random('unif',0,numclas));
            %eleccion=floor(rand(1));
            %eleccion=rand(1);
            %prob_actual=prob_actual-beta;
            bet=0.35;
            if actual==x1
                b1=1;
            else
                b1=0;
            end

            if actual==x2
                b2=1;

            else
                b2=0;
            end
            if actual==x3
                b3=1;
            else
                b3=0;
            end
            if actual==x4
                b4=1;
            else
                b4=0;
            end
        end
    end
end
```



```
end
if actual==x5
    b5=1;
else
    b5=0;
end
if actual==x6
    b6=1;
else
    b6=0;
end
if actual==x7
    b7=1;
else
    b7=0;
end
if actual==x8
    b8=1;
else
    b8=0;
end

%prob_eleccion=prob_eleccion-betal;

bet=0.35;

if eleccion==x1
    v1=1;
else
    v1=0;
end
if eleccion==x2
    v2=1;
else
    v2=0;
end
if eleccion==x3
    v3=1;
else
    v3=0;
end
if eleccion==x4
    v4=1;
else
    v4=0;
end
if eleccion==x5
    v5=1;
else
    v5=0;
end
if eleccion==x6
    v6=1;
else
    v6=0;
end
if eleccion==x7
    v7=1;
else
    v7=0;
end
if eleccion==x8
    v8=1;
else
    v8=0;
end
%Es el fin de potts

beta=-bet*(b1+b2+b3+b4+b5+b6+b7+b8);
betal=-bet*(v1+v2+v3+v4+v5+v6+v7+v8);
```

Apéndice

```
    if actual==1
    prob_actual=prob_actual + pXW1(i,j);
    elseif actual==2
    prob_actual=prob_actual + pXW2(i,j);
    elseif actual==3
    prob_actual=prob_actual + pXW3(i,j);
    elseif actual==4
    prob_actual=prob_actual + pXW4(i,j);
    elseif actual==5
    prob_actual=prob_actual + pXW5(i,j);
    else actual==6
    prob_actual=prob_actual + pXW6(i,j);
    end

    if eleccion==1
    prob_eleccion=prob_eleccion + pXW1(i,j);
    elseif eleccion==2
    prob_eleccion=prob_eleccion + pXW2(i,j);
    elseif eleccion==3
    prob_eleccion=prob_eleccion + pXW3(i,j);
    elseif eleccion==4
    prob_eleccion=prob_eleccion + pXW4(i,j);
    elseif eleccion==5
    prob_eleccion=prob_eleccion + pXW5(i,j);
    else eleccion==6
    prob_eleccion=prob_eleccion + pXW6(i,j);
    end

    if(prob_eleccion <= prob_actual)
        y(i,j)=eleccion;

    else
        p=exp(-((prob_eleccion-prob_actual)/t));

        %teta=floor(random('unif',0,1));
        teta=rand(1);

        if(teta<p)
            y(i,j)=eleccion;

        else
            y(i,j)=actual;
        end

    end

end

end

t=t*.95;
l=n;
end

% Se define la variable para almacenar la imagen en niveles de gris
zal=xa;
% Ciclo para asignar a cada pixel el nivel de gris correspondiente a su clase
% En vez de asignar al pixel su número de clase, se le asigna el nivel de gris correspondiente

for w=1:numclas

for i=2:ren-1
    for j=2:col-1
        if y(i,j)==w
            zal(i,j)=medias(w);
        else
            e=0;
        end
    end
end
end

end
```

Apéndice

```
% Impresion de resultados

figure(2);
colormap(map);
image(zal);
mapa=[0 0 0 ; 1 0 0 ; 0 1 0 ; 0 0 1 ; 1 1 0 ; 1 0 1 ; 0 1 1 ; 1 1 1];
figure(3)
colormap(mapa)
image(y)
tiempo=toc/60
```