



Centro de Investigación en Matemáticas, A.C.

Segmentación de Imágenes de Satélite usando
Modelos Ocultos de Campos Markovianos

TESIS

que para obtener el grado de
Maestro en Ciencias

con especialidad en
Computación y Matemáticas Industriales

presenta

Ely Judith Rosina Gallo Ramírez

Director de Tesis

Dr. José Luis Marroquín Zaleta

Codirector de Tesis

M.C. Aurora Breceda

Agosto, 2002

Guanajuato, Gto. México.

Con todo mi amor, dedico esta modesta tesis a mis padres :

Jesús y Rosalina .

Agradecimientos

A Dios, por concluir esta linda etapa.

A mis padres, por regalarme lo mejor de sus vidas; con su ejemplo me han enseñado que a base de esfuerzo, dedicación y amor, es posible lograr lo que muchas veces parece inalcanzable.

A *Irasema, Lili, Iraís y Chuy* por las porras que me echaron; me han hecho sonreír en los momentos más difíciles y agradezco a Dios la hermosísima experiencia de haber crecido junto a ustedes.

A mis asesores; al Dr. Marroquín por su infinita paciencia y a la M.C. Breceda por su valiosa colaboración en el desarrollo de este trabajo.

A cada uno de mis profesores de maestría y especialmente al Dr. Johan Van Horebeek.

Al CONACYT, sin la beca no habría podido realizar mis estudios de maestría en el CIMAT.

A los profesores del Instituto Tecnológico de La Paz que despertaron en mí la inquietud por realizar estudios de maestría: al M.C. José Luis López López, al Ing. Jorge Bravo Espinoza y especialmente, al Dr. Gil Salgado González.

A mis amigos, con quienes he compartido momentos de angustias, carcajadas y llantos.

Y otro agradecimiento especial, para *Rafael*; por alegrar mi estancia en Guanajuato y compartir conmigo sus más grandes tesoros.

Índice General

Capítulo 1. Introducción	1
Capítulo 2. Percepción Remota	4
2.1 Introducción	4
2.2 Radiación electromagnética	5
2.3 El espectro electromagnético	6
2.4 Firmas espectrales	7
2.4.1 Características de la vegetación en el espectro óptico	8
2.4.2 Características del suelo en el espectro óptico	11
2.4.3 Características del agua en el espectro óptico	11
2.5 Interacciones de la atmósfera con la radiación electromagnética	12
2.5.1 Absorción atmosférica	12
2.5.2 Dispersión atmosférica	13
2.5.2 Emisión atmosférica	14
2.6 LANDSAT	14
2.6.1 La familia LANDSAT	14
2.6.2 Algunas características del sensor Thematic Mapper	15
2.7 El proceso de clasificación	15
2.7.1 Fase de entrenamiento	16
2.7.1.1 Método supervisado	18
2.7.1.2 Método no supervisado	19
2.7.2 Fase de asignación	21
Capítulo 3. Método propuesto	22
3.1 Herramientas	22
3.2 Algoritmo de minimización de la función de energía en el caso multibanda	23
3.3 Cómo inicializar el modelo de firmas espectrales	25
3.3.1 Modelo inicializado mediante muestras aleatorias	25
3.3.2 Modelo inicializado mediante segmentaciones monobanda	26
Capítulo 4. Pruebas con imágenes sintéticas	28
4.1 La imagen sintética	28
4.2 Calibración de parámetros del método de Descenso Newtoniano	31
4.3 Prueba de convergencia del modelo de firmas	33
4.4 Pruebas de segmentar la imagen sintética partiendo de un modelo estimado	35
4.4.1 Modelo inicializado mediante muestras aleatorias	35

4.4.2 Modelo inicializado mediante segmentaciones monobanda	38
Capítulo 5. Pruebas con imágenes del satélite LANDSAT TM	42
5.1 Segmentaciones de la zona de San José del Cabo	42
5.2 Segmentaciones en la reserva de la biósfera Sierra de La Laguna	48
5.2.1 Segmentaciones en la zona de la vertiente del Golfo de California.	48
5.2.2 Segmentaciones en la zona de la vertiente del Océano Pacífico.	53
5.3 Conclusiones y comentarios	57
Conclusiones generales y trabajo futuro	59
Conclusiones generales	59
Trabajo futuro	59
Apéndice A. Modelo Oculto de Campo de Medidas Aleatorio Markoviano para Segmentación de Imágenes	61
A.1 Modelos clásicos de Campos Aleatorios Markovianos	61
A.1.1 Algoritmos de Estimación	62
A.2 Modelos de Campos de Medidas Markovianos Ocultos	65
A.3 Algoritmo de estimación de la función de energía	67
Apéndice B. Demostración de que f_{no} es un CAM.	70
Apéndice C. Manual de usuario del Software MRMF	71
C.1 Pantalla principal	71
C.2 Dimensiones de la imagen	72
C.3 Lectura de bandas	73
C.4 Parámetros de segmentación multibanda	74
C.5 Inicialización del modelo de firmas espectrales	75
C.5.1 Modelo de firmas inicializado mediante segmentaciones monobanda	75
C.5.2 Modelo de firmas inicializado mediante muestras aleatorias	77
C.5.3 Modelo de firmas inicializado mediante archivo	78
C.6 Salir del sistema	78
Bibliografía	79

Índice de Figuras

1.1	Proceso de una imagen clasificada	2
2.1	Propagación del flujo energético	6
2.2	Espectro electromagnético	7
2.3	Firmas espectrales típicas de distintas cubiertas	8
2.4	Firma espectral de la vegetación sana	10
2.5	Ventanas atmosféricas	13
2.6	Relaciones entre clases informacionales y espectrales	17
3.1	Representación de muestras aleatorias	26
3.2	Regiones formadas por la combinación de segmentaciones monobandas	27
4.1	Campo de etiquetas de la imagen sintética	30
4.2	Bandas de la imagen sintética	30
4.3	Resultados de la prueba 4.2_a	32
4.4	Resultados de la prueba 4.2_b	32
4.5	Resultados de la prueba 4.3_b	34
4.6	Resultados de la prueba 4.4.1_a	36
4.7	Resultados de la prueba 4.4.1_b	36
4.8	Resultado de la prueba 4.4.1_c	37
4.9	Resultado de la prueba 4.4.1_d	37
4.10	Segmentaciones monobanda de la prueba 4.4.2_a	38
4.11	Segmentaciones multibanda de la prueba 4.4.2_a	38
4.12	Segmentaciones monobanda de la prueba 4.4.2_b	39
4.13	Segmentación multibanda de la prueba 4.4.2_b	39
4.14	Segmentaciones monobanda de la prueba 4.4.2_c	40
4.15	Primera segmentación multibanda de la prueba 4.4.2_c	40
4.16	Segunda segmentación multibanda de la prueba 4.4.2_c	40
4.17	Tercera y cuarta segmentación multibanda de la prueba 4.4.2_c	40
5.1	Zona de San José del Cabo, vista desde el satélite LANDSAT TM	43
5.2	Resultados de la prueba 5.1_a	43
5.3	Resultados de la prueba 5.1_b	44
5.4	Píxeles etiquetados por las clases 5 y 2	45
5.5	Comparación entre categorías	45
5.6	Resultados de la prueba 5.1_c	46
5.7	Imagen reclasificada 5.1_c	47
5.8	Ubicación de la reserva de la biósfera Sierra de La Laguna	48

5.9	Zona de prueba en la vertiente del Golfo de California, vista desde las bandas 4 y 5 del sensor	48
	Thematic Mapper	
5.10	Resultado de la prueba 5.2.1_a	49
5.11	Resultado de la prueba 5.2.1_b	50
5.12	Píxeles etiquetados por la clase 3	50
5.13	Imagen reclasificada 5.2.1_b	51
5.14	Resultado de la prueba 5.2.1_c	51
5.15	Imagen reclasificada 5.2.1_c	51
5.16	Selva amatorralada	52
5.17	Áreas de escasa vegetación y suelos	52
5.18	Selva cerrada	52
5.19	Selva abierta	52
5.20	Sombras	52
5.21	Arroyos	52
5.22	Zona de prueba en la vertiente del Océano Pacífico, vista desde las bandas 4 y 5 del sensor	53
	Thematic Mapper	
5.23	Resultado de la prueba 5.2.2_a	53
5.24	Resultado de la prueba 5.2.2_b	54
5.25	Imagen reclasificada 5.2.2_b	55
5.26	Resultado de la prueba 5.2.2_c	55
5.27	Imagen reclasificada 5.2.2_c	56
5.28	Selva amatorralada	56
5.29	Selva cerrada	56
5.30	Selva abierta	56
5.31	Áreas de escasa vegetación	56
5.32	Sombras	56
A.1	Modelo clásico de Campos Aleatorios Markovianos para segmentación de imágenes	61
A.2	Campo Aleatorio Markoviano oculto para segmentación de imágenes	65
C.1	Pantalla principal	71
C.2	Dimensiones de la imagen	72
C.3	Visualizar imagen	72
C.4	Mensaje para definir bandas	73
C.5	Abrir archivo	73
C.6	Parámetros de segmentación	73
C.7	Parámetros de segmentación multibanda	74
C.8	Opciones de inicializar el modelo de firmas	75
C.9	Segmentación monobanda	75

C.10	Selección de ruta de almacenamiento	75
C.11	Formato de archivo del modelo de firmas	76
C.12	Formato de archivo de cabecera	76
C.13	Número de muestras aleatorias	77
C.14	Presentación de estadísticas de muestras aleatorias	77

Índice de Tablas

2.1	Características del Thematic Mapper	2
4.1	Propagación de la energía electromagnética	6
4.2	Propagación de la energía electromagnética	6

Índice de Apéndices

Apéndice A	61
Apéndice B	70
Apéndice C	71

Prólogo

La elaboración de esta tesis surgió por dos razones: mi interés por participar en un proyecto del área de Visión Computacional y mis deseos para enfocar dicho proyecto a resolver un problema tratado en el estado de Baja California Sur. Dentro de este marco de consideraciones, encontré que una investigadora del Centro de Investigaciones Biológicas del Noroeste (CIBNOR), M.C. Aurora Breceda en su tesis doctoral pretende describir en diferentes escalas¹ el mosaico de vegetación de la selva baja caducifolia en la reserva de la biósfera Sierra de la Laguna, B.C.S.

Cabe señalar que el énfasis en la selva baja caducifolia se debe a la importancia de esta comunidad. En general, las selvas bajas caducifolias son consideradas como uno de los ecosistemas terrestres más amenazados [22], ya que se encuentran asociados a zonas con una elevada densidad de población [43]. En México, este tipo de vegetación cubre cerca del 10% del territorio nacional [47], siendo uno de los ecosistemas más extensos del país. En la Península de Baja California, la selva baja caducifolia se restringe a la porción meridional de la Península con una extensión aproximada de 200 000 hectáreas. La zona en donde está mejor representada abarca las laderas de los cañones de la Sierra de la Laguna. Dentro de esta serranía se encuentra la reserva de la biósfera “Sierra de la Laguna”, la cual da protección a una extensa porción de la selva baja caducifolia de la región. Esta comunidad representa el límite noroccidental de la distribución de la selva baja caducifolia en América y se caracteriza por su aislamiento geográfico y ecológico [1], asimismo representa una de las comunidades más secas de su tipo [3]. Sobre esta comunidad existen varios trabajos florísticos y estructurales; sin embargo, no se ha estudiado desde una perspectiva regional y no ha sido descrito el mosaico de vegetación de esta comunidad. Por tal motivo, esta tesis se ha enfocado en realizar la segmentación de imágenes del satélite LANDSAT TM de la zona en estudio.

La técnica de segmentación presentada en este trabajo es una nueva formulación Bayesiana, propuesta en el CIMAT por: Dr. J.L. Marroquín, Dr. Salvador Botello y el estudiante de doctorado Edgar Arce. Dicha técnica se basa en la idea de usar un modelo doblemente estocástico que al minimizar una función diferenciable permite encontrar el campo de etiquetas que particiona a la imagen y los parámetros óptimos de las regiones en esta imagen segmentada [34, 35].

Los resultados de segmentar las imágenes de satélite han sido evaluados por la M.C. Aurora Breceda, el experto que finalmente puede indicar si presentan una correspondencia con atributos de la comunidad a niveles locales.

¹ Desde la escala que nos proporciona una imagen de satélite con el sensor Landsat Thematic Mapper hasta la que se obtiene mediante el muestreo de cuadrantes.

Capítulo 1

Introducción

El problema de segmentación de imágenes puede ser visto como un problema de clasificación donde es necesario encontrar el campo de etiquetas que particiona a la imagen en regiones mutuamente excluyentes $\{R_1, \dots, R_K\}$, así como determinar el conjunto de parámetros que definen a dichas regiones. Siempre existe un atributo de interés para realizar esta tarea (nivel de gris, profundidad, color, movimiento, textura, etc.), en muchas ocasiones lo es el nivel de gris, por lo que los parámetros que definen las regiones de la imagen son generalmente la media y la varianza de dicho atributo.

Conociendo los parámetros de cada región (Φ_k) es posible calcular las probabilidades de que cierto pixel sea observado de alguna manera, dado que pertenezca a cada una de las regiones en cuestión (cálculo de verosimilitudes), pero el hecho de conocer las verosimilitudes para determinado pixel no implica que éste deba pertenecer a la región que le resultó con máxima verosimilitud, debido al ruido introducido en el proceso de adquisición de la imagen (por la luminosidad que se tenga en la escena, por el aparato que captura la imagen u otros factores). Veamos un ejemplo. Supongamos que tenemos un conjunto de naranjas y las vamos a clasificar en verdes y maduras. Para esto contamos con un fotosensor, las naranjas están sobre una banda y conforme avanza la banda, el fotosensor captura la intensidad de color de la naranja posicionada justo enfrente; más adelante está un brazo mecánico que separa las naranjas verdes de las maduras. Denotemos por f a la intensidad verdadera de la naranja y como g a la intensidad medida. Estos valores no necesariamente deberán ser iguales, precisamente por el ruido de observación mencionado anteriormente, entonces $g = f + n$, donde n representa el ruido. Ahora bien, podemos conocer la verosimilitud, $P(g/f)$ si especificamos cuál es la función de distribución del ruido, $P(g/f) = P(n = g - f)$; suponiendo que $n \sim N(0, \sigma^2)$ tenemos que $P(g/f) = P_n(g-f) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(g-f)^2}{2\sigma^2}}$. En este caso de clasificar

naranjas es necesario conocer dos valores de f , la intensidad media de las naranjas verdes y la intensidad media de las naranjas maduras; después de calcular las probabilidades de observar la intensidad medida dado que la naranja es verde/madura, el brazo mecánico va a colocar la naranja en el costal correspondiente a la clase de naranja que le resultó con máxima verosimilitud y bueno, es evidente que esta no es una forma muy confiable de clasificación (ni de naranjas, ni de pixeles en imágenes, ni de nada). Resultaría de gran ayuda saber si las intensidades de las naranjas varían dependiendo de la estación del año en que se midan o si existe otro factor que pudiéramos incluir como conocimiento *a priori* con el propósito de obtener una clasificación más cercana a la realidad.

En el caso de segmentar una imagen, es posible introducir al proceso, información sobre la interacción espacial de la misma en forma de función de distribución de probabilidad. Dicho conocimiento *a priori* generalmente refleja dependencias en la clasificación de pixeles vecinos y puede modelarse perfectamente haciendo uso de la teoría de Campos Aleatorios Markovianos¹. La figura 1.1 ilustra este párrafo.

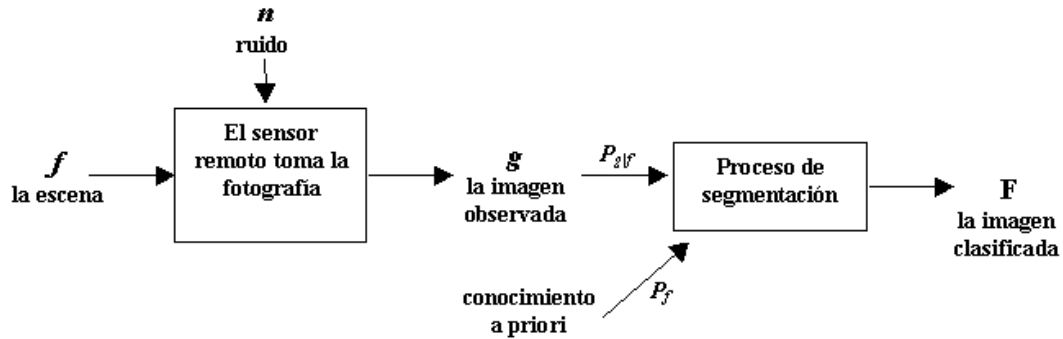


Figura 1.1 Proceso de una imagen segmentada

Cuando hacemos segmentación, deseamos obtener un campo de etiquetas para la escena f , pero no conocemos f , sólo conocemos g . Entonces podríamos generar muchos campos de etiquetas f y encontrar aquella donde la $P(f|g)$ sea máxima (máxima probabilidad a posteriori). Ahora bien, aplicando la regla de Bayes: $P(f|g) = \frac{P(g|f)P(f)}{P(g)}$, donde $P(g)$ es un valor constante y podemos decir que $P_{f|g} = \frac{1}{Z} P_{g|f} P_f$, donde

$P_{f|g}$ es la distribución a posteriori, $P_{g|f}$ es la verosimilitud y P_f es la información a priori en forma de función de distribución de probabilidad. Si suponemos que n es ruido Gaussiano y que en la imagen está distribuido de manera independiente sobre cada pixel, tendremos $P_{f|g} = \frac{1}{Z} e^{-U(f|g)}$ y entonces la tarea de maximizar esta probabilidad a posteriori, se convierte en minimizar la función $U(f|g)$ [32].

El problema de segmentación se hace especialmente difícil cuando hay que estimar al mismo tiempo los parámetros que caracterizan cada modelo Φ_k y las regiones válidas para esos modelos. Para resolver esto, el algoritmo EM (Expectation Maximization) [12] modela el conocimiento *a priori* con CAM y ejecuta un proceso de pasos iterativos. Este algoritmo es razonablemente eficaz, pero exige gran esfuerzo computacional. En esta tesis se hace uso de una nueva clase de modelos probabilísticos que permite encontrar la solución a complejos problemas de segmentación en términos de la minimización de una función de energía diferenciable, con lo que otros algoritmos eficientes pueden fallar. Estos modelos, fuertemente basados en la teoría de estimación Bayesiana, representan una mejora significativa sobre los

¹ CAM por sus iniciales, o en inglés MRF (Markov Random Fields).

modelos clásicos de CAM tanto en la precisión de las soluciones como en la complejidad computacional. Puede consultar los detalles del método en el apéndice A [34].

En el contexto de percepción remota [8], el problema de segmentación es meramente un problema de clasificación; de hecho, quienes se dedican al tratamiento de imágenes de percepción remota dicen *segmentar* cuando se refieren a *recortar* una imagen; para la mayor parte esos usuarios, la clasificación supone la fase culminante del tratamiento digital de imágenes. Los Números Digitales (ND²) adquiridos por el sensor remoto no tienen sentido en sí mismos hasta que son interpretables; es decir, hasta que son etiquetados en categorías que supongan un conocimiento del territorio.

El tratamiento digital se inicia con la *corrección* de la imagen. Cualquier imagen adquirida por un sensor remoto, ya sea aéreo o espacial, presenta una serie de alteraciones radiométricas y geométricas debidas a muy variados factores. Esto explica que la imagen detectada no coincida exactamente con el tono-posición-forma-tamaño de los objetos que incluye. En el caso más concreto de las imágenes espaciales, las deformaciones más frecuentes pueden agruparse de la siguiente manera:

- distorsiones originadas por la plataforma.
- distorsiones provocadas por la rotación terrestre.
- distorsiones provocadas por el sensor.
- distorsiones provocadas por la atmósfera.

Algunos de estos problemas se solventan rutinariamente en los centros encargados de la percepción o venta de las imágenes. Otros, sin embargo, persisten, haciendo necesarias una serie de técnicas de ajuste.

El siguiente paso consiste en definir las categorías en forma digital (se hace en una fase de entrenamiento supervisado y/o no supervisado) y luego se etiqueta cada pixel de la imagen en una de esas categorías (fase de asignación) para finalmente realizar la comprobación y verificación de resultados.

Las formas en que se llevan a cabo las fases de entrenamiento y asignación se plantean en el capítulo 2, donde también se encuentran algunos conceptos importantes para la interpretación de firmas espectrales y ciertas características del satélite LANDSAT TM. En el capítulo 3 veremos las aportaciones a esta tesis: la extensión del método presentado en [34] (apéndices A y B) al caso multibanda y dos propuestas para obtener un modelo de inicio. Luego, en el capítulo 4 veremos pruebas en imágenes sintéticas y en el capítulo 5 los resultados y conclusiones de realizar las segmentaciones en distintas zonas de las imágenes de satélite. Finalmente, el apéndice C es un pequeño manual de usuario, donde se explica brevemente cómo manejar el software desarrollado para realizar segmentaciones multibanda con el MRMF utilizando las propuestas del capítulo 3.

² El ND se conoce con distintos nombres en la bibliografía anglosajona: *Digital Number*, *Pixel Value*, *Brightness Value* o *Gray Level*.

Capítulo 2

Percepción Remota

2.1 Introducción

Una de las principales inquietudes que el hombre siempre ha tenido es la exploración del planeta tierra. Su atención ha sido atraída por una poderosa necesidad de saber qué se encuentra más allá de sus dominios. Al principio exploró el medio ambiente únicamente con sus sentidos, utilizando en primera instancia aquellos que no requerían de un contacto físico con el objeto estudiado, es decir, valiéndose únicamente de la vista y el oído, para posteriormente continuar con un análisis directo, principalmente por medio del tacto.

Así, la curiosidad del hombre y afán por conocer el territorio donde habita le condujeron primero a encaramarse en la copa de un árbol y posteriormente a subir a una colina o a una montaña para tener de esta manera una visión de conjunto o lo que es lo mismo, una visión sinóptica del paisaje. Aun en esta reseña esquemática de las actividades primitivas de exploración del hombre, pueden distinguirse los elementos básicos de la percepción remota: *la fuente de iluminación*, formada en este caso por el sol que emite luz o radiación solar; *el paisaje*, configurado por todos los objetos presentes en el territorio explorado por el hombre, tales como ríos, colinas, vegetación o rocas; *la escena*, aquella sección o superficie del paisaje donde se enfoca cierto interés, pudiendo ser la superficie de una roca o un valle completo; *el sensor remoto*, con el que se captura la luz proveniente de la escena y que, en este caso, permite obtener una representación visual de aquella región del paisaje que ha sido enfocada; *la plataforma*, que es el lugar donde se coloca o monta el sensor remoto para obtener una visión de conjunto de la escena; *el sistema de procesamiento*, compuesto por el dispositivo para procesar cualitativa o cuantitativamente los datos proporcionados por el sensor remoto acerca de la escena, los cuales, analizados e interpretados, producen información valiosa en relación al paisaje, la que es útil en la toma eventual de decisiones sobre el medio ambiente, y *el apoyo del campo*, que consiste en la inspección directa, en varios puntos selectos de la escena, de diferentes atributos de los objetos que se encuentran en el terreno, con el fin de evaluar los datos obtenidos a distancia [28].

La distancia con respecto a la escena, a la cual se coloca la plataforma, permite tener una visión sinóptica a diferentes escalas del paisaje, lo que se logra gracias a una sucesión de observaciones que van cubriendo poco a poco la región deseada. Éstas pueden hacerse, en general a diferentes horas del día, en diferentes épocas del año y, muy probablemente, por medio de diferentes filtros de color. En el proceso se genera una gran cantidad de datos, los que analizados adecuadamente por medio de criterios muy bien definidos, proporcionan valiosa información acerca de los recursos naturales terrestres y sientan las bases para su estudio racional.

Con el desarrollo moderno de las ciencias de la computación se ha dado un auge importante a la percepción remota, pues la disponibilidad de estas herramientas permite la evaluación cuantitativa de un gran volumen de datos. Así, en una simbiosis muy estrecha, las ciencias de la computación han impulsado el desarrollo de la percepción remota, y ésta, a su vez, ha generado nuevos sistemas de procesamiento digital con propósitos específicos.

En este capítulo veremos algunos factores que permiten una interpretación rigurosa de la imagen obtenida por el sensor remoto, algunas características del satélite LANDSAT TM y las formas en que los paquetes de software comerciales y Sistemas de Información Geográfica comúnmente definen firmas espectrales para la clasificación de una imagen y cómo realizan esta tarea.

2.2 Radiación electromagnética

Existen varias formas de adquirir información a partir de un sensor remoto: por reflexión, por emisión y por emisión-reflexión. La primera de ellas es la forma más importante de teledetección, pues se deriva directamente de la luz solar, principal fuente de energía de nuestro planeta. El sol ilumina la superficie terrestre, que refleja esa energía en función del tipo de cubierta presente en ella. Ese flujo reflejado se recoge por el sensor, que lo transmite posteriormente a las estaciones receptoras. Entre superficie y sensor se interpone la atmósfera que dispersa y absorbe parte de la señal original. De igual forma, la observación remota puede basarse en la energía emitida por las propias cubiertas, o bien, en la que podríamos enviar desde un sensor que fuera capaz tanto de generar su propio flujo energético como de recoger posteriormente su reflexión sobre la superficie terrestre.

En cualquiera de estos casos, el flujo energético entre la cubierta terrestre y el sensor constituye una forma de radiación electromagnética. Como es sabido, la energía térmica se transfiere de un lugar a otro por tres procesos: convección, conducción y radiación. De ellos, nos interesa el último, pues constituye la base del sistema de teledetección analizado en este trabajo.

Las propiedades de la radiación electromagnética se han explicado por dos teorías: aquella que la concibe como un haz ondulatorio (Huygens, Maxwell) y la otra que la considera como una sucesión de unidades discretas de energía llamados fotones, con masa igual a cero (Planck, Einstein) [8].

La energía electromagnética se propaga en el vacío a la velocidad de la luz, siguiendo un modelo armónico y continuo, conteniendo dos campos de fuerzas ortogonales entre sí: eléctrico y magnético; éstos ortogonales a la dirección de propagación del flujo energético (ver la figura 2.1) [53].

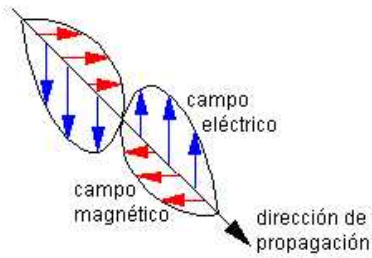


Fig. 2.1 Propagación del flujo energético.

Podemos describir este flujo así: $c=\lambda F$, donde c indica la velocidad de la luz (aproximadamente 300 000 km/seg en el vacío), λ es la longitud de onda y F la frecuencia. λ hace referencia a la distancia entre dos picos sucesivos de una onda mientras que la F indica el número de ciclos presenciados en una unidad de tiempo.

Gracias a la teoría cuántica, es posible calcular la cantidad de energía transportada por un fotón, siempre que se conozca su frecuencia o longitud de onda: $Q=hF$ o bien, $Q=hc/\lambda$ donde Q es la energía radiante de un fotón (medida en joules), y h la constante de Planck (6.6×10^{-34} Js). Lo que significa en definitiva, que a mayor longitud de onda (o menor frecuencia) el contenido energético será menor y viceversa. Esto implica que la radiación en longitudes de onda largas es más difícil de detectar que aquella centrada en longitudes cortas, de ahí que las primeras requieran más sofisticados medios de detección.

2.3 El espectro electromagnético

El espectro electromagnético suele organizarse en una serie de bandas de longitudes de onda o frecuencia en donde para cada banda, la radiación electromagnética manifiesta un comportamiento similar (Fig. 2.2). Comprende desde las longitudes de onda más cortas (rayos gamma, rayos X) hasta las kilométricas (telecomunicaciones). Las unidades de medida más comunes se relacionan con la longitud de onda. Generalmente, para las más cortas se utilizan micras ($\mu m=10^{-6}$ metros), mientras las más largas se miden en centímetros o metros.

Desde el punto de vista de la teledetección, conviene destacar una serie de bandas espectrales, que son las más frecuentemente empleadas con la tecnología actual. Su denominación y amplitud varían según distintos autores, aunque la terminología más común es la siguiente:

- Espectro visible (0.4 a 0.7 μm). Se denomina así por tratarse de la única radiación electromagnética que pueden percibir nuestros ojos, coincidiendo con las longitudes de onda en donde es máxima la radiación solar. Suelen distinguirse tres bandas elementales, que se denominan azul (0.4 a 0.5 μm), verde (0.5 a 0.6 μm) y rojo (0.6 a 0.7 μm), en razón de los colores elementales asociados a esas longitudes de onda.

- Infrarrojo próximo (0.7 a 1.3 μm). A veces se denomina también infrarrojo reflejado y fotográfico, puesto que puede detectarse a partir de films dotados de emulsiones especiales. Resulta de especial importancia por su capacidad para discriminar masas vegetales y concentraciones de humedad.
- Infrarrojo medio (1.3 a 8 μm), en donde se entremezclan los procesos de reflexión de la luz solar y de emisión de la superficie terrestre.
- Infrarrojo lejano o térmico (8 a 14 μm), que incluye la porción emisiva del espectro terrestre.
- Microondas (a partir de 1mm), con gran interés por ser un tipo de energía bastante transparente a la cubierta nubosa.

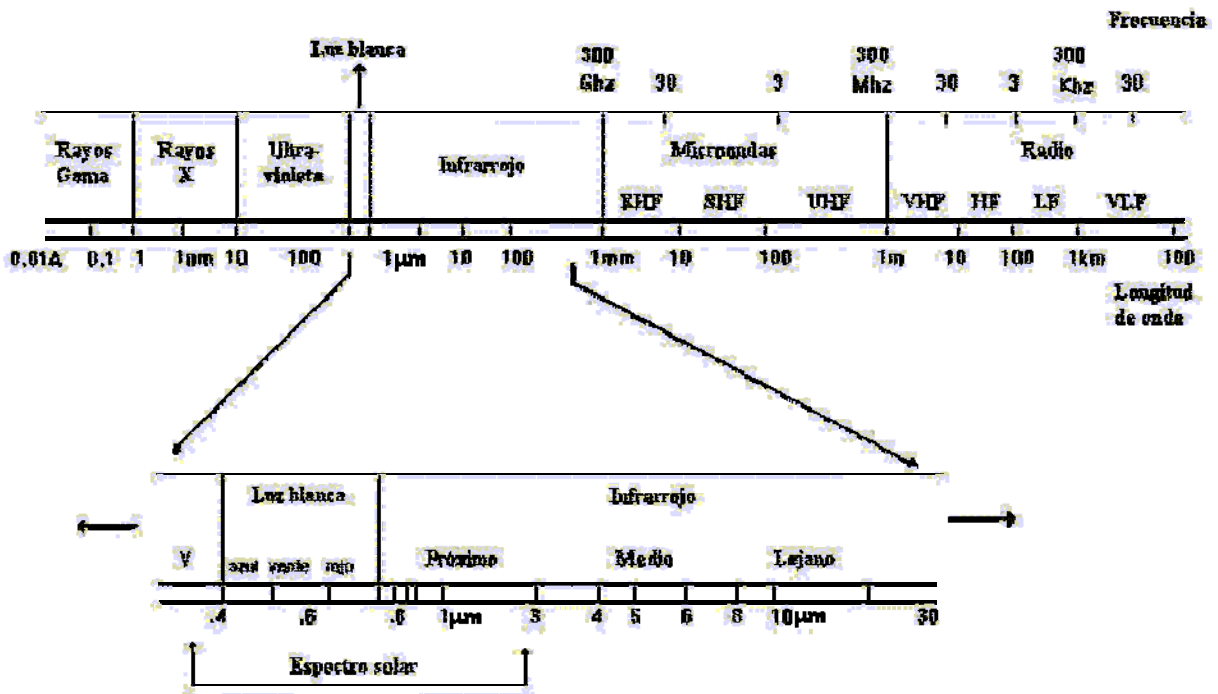


Figura 2.2 Espectro electromagnético

Enseguida estudiaremos el comportamiento espectral de las principales cubiertas terrestres en cada una de estas bandas del espectro.

2.4 Firmas espectrales

Tengamos en cuenta que el flujo incidente sobre una superficie siempre es reflejado, transmitido o absorbido y que la proporción del flujo incidente que es reflejado, absorbido y transmitido depende de las características de dicha superficie.

Por cuanto esa proporción varía en distintas bandas del espectro, nos resulta muy interesante conocer el comportamiento de dicha cubierta en diversas longitudes de onda, de cara a realizar una más atinada discriminación, puesto que sólo cuando existan divergencias espectrales entre dos superficies podrán éstas separarse adecuadamente.

En el caso del espectro visible, ese comportamiento disimilar de los objetos a distintas longitudes de onda se manifiesta en lo que llamamos color: un objeto es azul si refleja intensamente la energía en esa banda del espectro y poco en el resto (dicho de otro modo, si absorbe o transmite poca energía incidente en esa banda), mientras será verde si su reflectividad se centra en esa banda y es baja en otras.

A partir de medidas de laboratorio, se han obtenido unas curvas de reflectividad espectral para las principales cubiertas terrestres (Fig. 2.3). Como puede observarse, algunas tienden a presentar una respuesta uniforme en distintas longitudes de onda, mientras otras ofrecen un comportamiento mucho más selectivo.

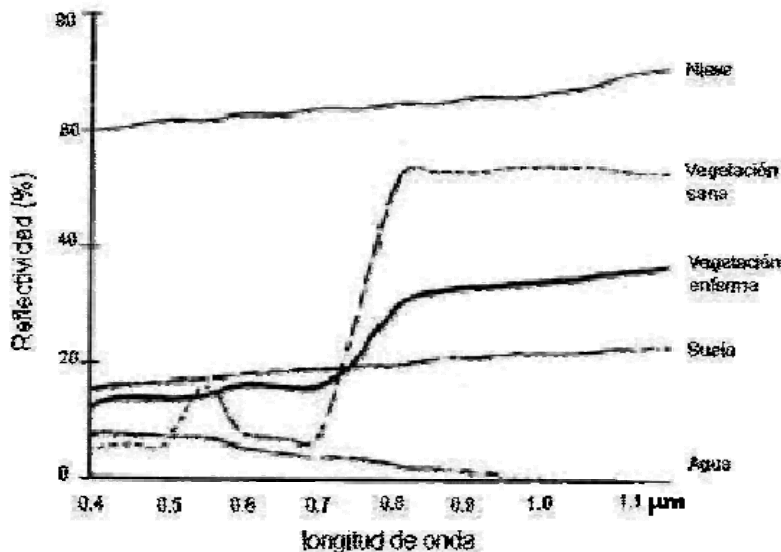


Figura 2.3 Firmas espectrales típicas de distintas cubiertas

La nieve presenta una reflectividad alta y constante, pues refleja la mayor parte de la energía incidente a distintas longitudes de onda. Por el contrario, el agua absorbe la mayor parte de la energía que recibe, tanto más nos situamos en longitudes de onda mayores. Por su parte, la vegetación presenta un comportamiento muy cromático, con bajos valores de reflectividad en el espectro visible y más elevados en el infrarrojo cercano.

El comportamiento típico de la vegetación vigorosa muestra una reducida reflectividad en las bandas visibles, con un máximo relativo en la porción verde del espectro (en torno a 0.55 µm). Por el contrario, en el infrarrojo cercano presenta una elevada reflectividad, reduciéndose paulatinamente hacia el infrarrojo medio.

2.4.1 Características de la vegetación en el espectro óptico

El comportamiento típico de la vegetación vigorosa muestra una reducida reflectividad en las bandas visibles, con un máximo relativo en la porción verde del espectro (en torno a 0.55 µm). Por el contrario, en el infrarrojo cercano presenta una elevada reflectividad, reduciéndose paulatinamente hacia el infrarrojo medio.

Estas características espectrales se relacionan primordialmente con la acción de los pigmentos fotosintéticos y del agua que almacenan las hojas. En concreto, la baja reflectividad en la porción visible del espectro se debe al efecto absorbente de los pigmentos de la hoja, principalmente las clorofilas, xantofilas y carotenos [14]. Todos ellos absorben en la banda del espectro situada en torno a los 0.445 μm , mientras la clorofila presenta una segunda banda de absorción en torno a los 0.645 μm . Entre ambas porciones del espectro, aparece una banda intermedia, alrededor de los 0.55 μm , en donde el efecto absorbente es menor. Por esta causa aparece un pico relativo de reflectividad que coincide con la banda verde del espectro visible, y causa el color con el que nuestros ojos perciben la vegetación vigorosa.

Cuando se aproxima la caída otoñal de las hojas, la clorofila ejerce una menor influencia, lo que explica su mayor reflectividad en la banda roja y, en definitiva, su color amarillento (verde + rojo). En algunas especies resulta destacada la acción de otro pigmento, la antocianina, buen reflector de la porción roja del espectro, que causa ese color en épocas de senescencia.

En cuanto a la elevada reflectividad en el infrarrojo cercano, parece deberse a la estructura celular interna de la hoja [19]. En concreto, la capa esponjosa del mesófilo, con sus cavidades de aire internas, ejerce un papel protagonista, al difundir y dispersar la mayor parte de la radiación incidente en esta banda del espectro. Por ello, la hoja sana ofrece una alta reflectividad en el infrarrojo cercano (entre 0.7 y 1.3 μm), en claro contraste con la baja reflectividad que ofrece en el espectro visible, especialmente con la banda roja. Puesto que la estructura de la hoja es muy variada según las especies, esta banda también resulta idónea para discriminar entre plantas, incluso entre aquellas que no podrían separarse en el espectro visible.

A partir de 1.4 μm el efecto absorbente del agua es muy claro, por lo que la reflectividad de la vegetación sana se reduce drásticamente en el infrarrojo medio. Las medidas de laboratorio muestran una notable diferencia en esta región del espectro entre las hojas secas y las infiltradas con agua [26,10]. Entre las bandas donde la absorción del agua es más clara, se sitúan dos picos relativos de reflectividad, en torno a 1.6 μm y 2.2 μm . Queda patente pues, que la observación de esta región del espectro es de gran interés para conocer el estado de vigor de la hoja, en función de su contenido de humedad (ver figura 2.4).

Obviamente el sensor no mide hojas aisladas, sino agrupaciones de hojas formando masas de vegetación, por lo que intervienen otros elementos adicionales que complican el esquema previo. Algunos de ellos son: el ángulo de elevación solar, directamente relacionado con la proporción de sombras que detecta el sensor; las características del suelo que sustenta la vegetación; la geometría de la cubierta vegetal; su estado fenológico, etc.

Pese a la complejidad final, es evidente que estas ideas nos sirven para centrar el estudio científico de la vegetación a partir de sensores espaciales. De ellas, resulta evidente que el contraste más nítido en la reflectividad espectral de la vegetación sana se produce entre las bandas visibles, especialmente el rojo (en torno a $0.645 \mu\text{m}$) y el infrarrojo cercano (0.7 a $1.3 \mu\text{m}$). De ahí que podemos enunciar como principio genérico, que cuanto mayor sea el contraste entre ambas bandas mayor será el vigor de la vegetación, y más clara su discriminación frente a otros tipos de cubierta. Queda implícito que cualquier fuente de estrés en la vegetación se mostrará en un comportamiento espectral más o menos alejado del anteriormente expuesto. La hoja senescente o enferma tiende a perder actividad clorofílica y en consecuencia, a ofrecer una menor absorptividad en las bandas azul y roja del espectro visible. El aumento consecuente de la reflectividad en estas bandas elimina el máximo relativo antes situado en el verde, por lo que la hoja tiende a mostrar un color amarillento. Por el contrario, en el infrarrojo cercano se produce una reducción de la reflectividad, como consecuencia de un deterioro en la estructura celular de la hoja. La curva espectral por tanto, se hace más plana, menos cromática [44,26].

Este comportamiento espectral teórico de las cubiertas vegetales ha sido la base para obtener una serie de índices de vegetación, que se basan entre otros factores, en el contraste entre las bandas roja e infrarrojo cercano del espectro. Se construyen a partir de la combinación lineal entre estas dos bandas, cuando disponemos de una imagen multispectral. Muy variados en su desarrollo, todos ellos tienden a presentar de forma más nítida las características de la vegetación, facilitando su aislamiento de otras cubiertas y la detección de su estado vital [39,21].

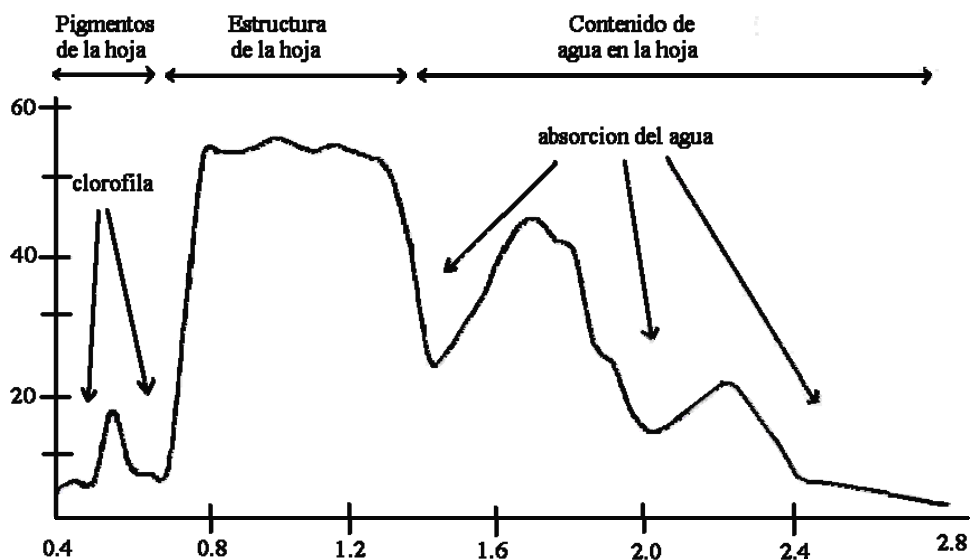


Figura 2.4 Firma espectral de la vegetación sana

2.4.2 Características del suelo en el espectro óptico

Como consecuencia del efecto de pantalla provocado por la cubierta vegetal, no resulta sencillo obtener información del sustrato geológico a partir de sensores espaciales. Sin embargo, pueden obtenerse algunos datos indirectos sobre las características del suelo/litología gracias a las anomalías que pueden detectarse en esa cubierta vegetal [41,30].

En lo que atañe a los suelos desnudos, su comportamiento espectral es mucho más uniforme que el de vegetación, mostrando una curva espectral bastante plana y de carácter ascendente. Los principales factores que intervienen en este caso son la composición química del suelo, su textura, estructura y contenido de humedad.

La composición química es la causa del color dominante con el que percibimos el suelo. Los de origen calcáreo tienden al color blanco, indicando una alta reflectividad en todas las bandas visibles. Por su parte, los suelos arcillosos ofrecen una mayor reflectividad en el rojo, como consecuencia de su alto contenido en óxido de hierro.

Respecto a las características físicas del suelo, puede afirmarse en términos generales, que la reflectividad espectral resulta tanto mayor cuando se trata de suelos más groseros, apelmazados, secos y sin materia orgánica. El contenido de humedad es uno de los elementos destacados importantes en la reflectividad de las longitudes de onda más largas, infrarrojo cercano y medio, como consecuencia de la alta absorción del agua en estas bandas. De hecho, un suelo de textura gruesa puede presentar una reflectividad menor que los finos cuando el contenido de humedad es bajo [42].

Como es bien sabido, textura, estructura y contenido de humedad están altamente relacionados. Por ejemplo, un suelo arcilloso tiende a tener un alto contenido de humedad, fuerte estructura y fina textura, lo que supone una baja reflectividad. En contraste, un suelo arenoso tiende a tener una débil estructura y bajo contenido de humedad, por lo que presentará una reflectividad más elevada [10]. A ello convendría añadir el efecto de la materia orgánica en el suelo, que tiende a oscurecerle y en consecuencia, a reducir su reflectividad.

2.4.3 Características del agua en el espectro óptico

Las superficies acuáticas absorben o transmiten la mayor parte de la radiación visible que reciben, siendo mayor su absorción cuanto mayor sea la longitud de onda. La curva espectral, por tanto, es similar a la del suelo, aunque en sentido contrario. La mayor reflectividad del agua clara se produce en el azul,

reduciéndose paulatinamente hacia el infrarrojo cercano, donde ya es prácticamente nula. Por esta razón, la frontera tierra-agua es muy nítida en esta banda.

La variabilidad del agua es mejor detectable en las longitudes de onda más cortas (azul y verde), que hemos de relacionar con su profundidad, contenido de materiales en suspensión (clorofila, arcillas y nutrientes), y rugosidad de la superficie.

La profundidad del agua influye directamente en el aporte de reflectividad derivado de los materiales de los fondos. En aguas poco profundas la reflectividad aumenta, por cuanto se produce un aporte de las características espectrales de los fondos. En consecuencia, la absorptividad será tanto mayor cuanto más profunda sea la capa de agua.

2.5 Interacciones de la atmósfera con la radiación electromagnética

Ahora, hay que tomar en cuenta que la atmósfera juega un papel importante cuando se trata de percepción remota puesto que se interpone entre el sensor y la superficie terrestre, interfiriendo de diversas formas con el flujo radiante. Como es sabido, la atmósfera se compone de gases (bióxido de carbono, oxígeno, ozono, nitrógeno y argón, principalmente), vapor de agua y aerosoles. Entre estos componentes, el bióxido de carbono, el ozono y el vapor de agua son los principales responsables de la interacción con la energía electromagnética.

Su efecto es triple: *absorción* de la energía en determinadas bandas del espectro, lo que limita la observación espacial a aquellas bandas en donde esta absorción es menos intensa o nula; *dispersión* de determinados flujos, en función de su longitud de onda; y *emisión*, como todo cuerpo caliente, que tiene su mayor importancia en el infrarrojo térmico. Estos procesos introducen modificaciones, en ocasiones muy severas en la radiación originalmente propagada entre la cubierta y el sensor.

2.5.1 Absorción atmosférica

La atmósfera se comporta como un filtro selectivo a distintas longitudes de onda, de tal manera que en algunas bandas del espectro elimina prácticamente cualquier posibilidad de observación remota. Los causantes de esta absorción son:

- Oxígeno atómico (O_2). Filtra las radiaciones ultravioleta por debajo de $0.1 \mu m$, así como pequeños sectores del infrarrojo térmico y las microondas.
- Ozono (O_3). Elimina la energía ultravioleta inferior a $0.3 \mu m$, así como un sector de las microondas (en torno a $27mm$).

- Vapor de agua. Fuerte absorción en torno a $6 \mu\text{m}$ y otras menores entre 0.6 y $2 \mu\text{m}$.
- Bióxido de carbono (CO_2). Absorbe en el infrarrojo térmico ($15 \mu\text{m}$), con importantes efectos en el infrarrojo medio, entre 2.5 y $4.5 \mu\text{m}$.

Como consecuencia de esta absorción, la observación espacial se reduce a determinadas bandas del espectro, conocidas como ventanas atmosféricas, en donde la transmisividad de la atmósfera es suficientemente alta (ver Fig. 2.5). Las principales ventanas son las siguientes: espectro visible e infrarrojo cercano, situada entre 0.3 y $1.35 \mu\text{m}$; varias en el infrarrojo medio, de 1.5 a $1.8 \mu\text{m}$, 2.0 a $2.4 \mu\text{m}$, 2.9 a $4.2 \mu\text{m}$ y 4.5 a $5.5 \mu\text{m}$; infrarrojo térmico, entre 8 y $14 \mu\text{m}$; y microondas, por encima de 20mm , en donde la atmósfera es prácticamente transparente.

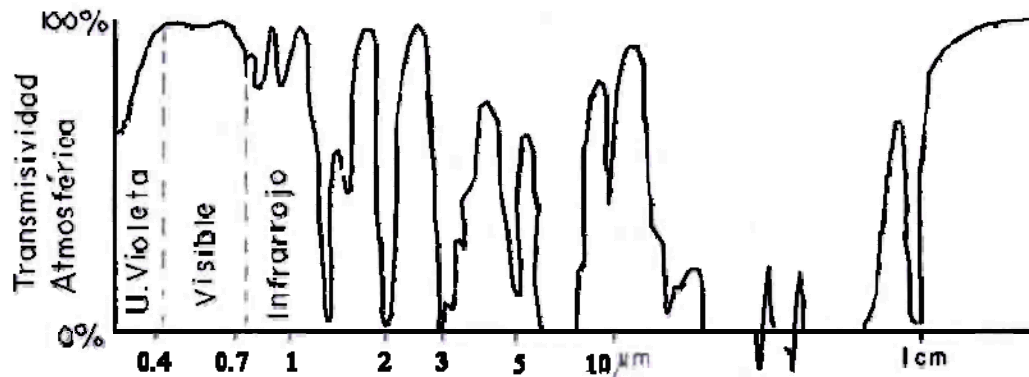


Figura 2.5 Ventanas atmosféricas

Estas ventanas atmosféricas son idóneas para realizar procesos de teledetección, por lo que el diseño de sensores espaciales tiende a ajustarse a estas bandas, evitando interferencias extrañas al fenómeno que pretende observarse (en la sección 2.6.2 veremos que las bandas del satélite LANDSAT TM cumplen esta característica). Cuando se pretende, por el contrario, observar la atmósfera en lugar de la superficie terrestre, los sectores espectrales más convenientes son precisamente aquellos en donde la absorción atmosférica es alta. Por tal razón, los satélites meteorológicos incorporan bandas en estas regiones del espectro.

2.5.2 Dispersión atmosférica

Es causada por la interacción entre la radiación electromagnética y los gases y partículas atmosféricas en suspensión. La reflexión consiguiente a ese choque supone un aporte adicional a la radiancia proveniente de la superficie terrestre (luz atmosférica). Puesto que las partículas atmosféricas son muy variables en el tiempo y en el espacio, resulta muy complejo cuantificar su influencia final en la imagen adquirida por el sensor. Por lo que la corrección atmosférica se basa en relaciones entre elementos de la propia imagen. Esta

estimación puede ser burda en algunos casos, y dificulta en última instancia, el uso de las imágenes adquiridas por sensores espaciales como medida cuantitativa de parámetros biofísicos.

2.5.3 Emisión atmosférica

El efecto de emisión atmosférica resulta fundamental dentro del infrarrojo térmico, si se pretenden obtener mediciones de temperatura a partir de imágenes espaciales. Al igual que cualquier cuerpo por encima del cero absoluto, la atmósfera emite su propia energía calorífica, por lo que este parámetro debe considerarse para estimar la emitancia espectral originalmente proveniente del suelo.

2.6 LANDSAT

2.6.1 La familia Landsat

La serie de satélites de la familia Landsat tiene su origen en 1972 cuando se puso en órbita el primero, denominado ERTS (Earth Resource Technology Satellite). A partir del segundo lanzamiento que tuvo lugar en 1975, se les rebautizó con el nombre de LANDSAT. Landsat puede producir las imágenes de todas las partes de la superficie terrestre sin pedir permisos de los gobiernos y ofrece las imágenes a todos los que se interesen en ellas a precios uniformes. Landsat suministró la primera base de datos de la tierra completa con resoluciones espaciales y espectrales adecuadas para varias aplicaciones. Los datos de Landsat están disponibles en forma digital. Los satélites de Landsat de la primera generación: Landsat 1, 2 y 3 fueron colocados en sus órbitas en los años 1972, 1975 y 1978. La segunda generación de Landsat se constituye de dos satélites: Landsat 4 y 5, lanzados en 1982 y 1984. La principal aportación de estos satélites es la incorporación de un nuevo sensor denominado Thematic Mapper (TM), especialmente diseñado para la cartografía temática. Este sensor mejora su antecesor MultiSpectral Scanner (MSS) principalmente en los siguientes aspectos: pasa de 79 a 30 metros su resolución espacial, incrementa la resolución espectral de 4 a 7 bandas, y la resolución radiométrica pasa de 6 a 8 bits. Las mejoras que ofrece el sensor TM facilitan una mayor precisión para la cartografía temática, ampliando el rango de aplicaciones operativas tanto terrestre como marinas. En 1993 se intentó poner en órbita al satélite Landsat 6 pero el intento falló. El último satélite de esta familia, el Landsat 7, fue lanzado el 10 de abril de 1999 y se caracteriza por una combinación de cobertura sinóptica, por alta resolución espacial con bandas correspondientes a la observación visible del infrarrojo y por una banda adicional pancromática de 15 m de resolución y un rango espectral cubriendo el verde al infrarrojo cercano.

2.6.2 Algunas características del sensor Thematic Mapper

El 'thematic mapper' detecta las bandas espectrales de 1 a 7. A continuación mostramos las características de cada una de las bandas.

Tabla 2.1 Características del Thematic Mapper

Banda	Longitud de onda (en μm)	Región del espectro	Algunas Características
1	0.45 - 0.52	Azul	Penetración máxima por agua, adecuada para la distinción entre suelo y vegetación
2	0.52 - 0.60	Verde	Delinea la reflectancia de la vegetación.
3	0.63 - 0.69	Rojo	Delinea una banda de absorción de clorofila, importante para distinguir tipos de vegetación.
4	0.76 - 0.90	Infrarrojo cercano	Adecuada para determinar el contenido en biomasa y para el mapeo de líneas de ribera.
5	1.55 - 1.75	Infrarrojo medio	Indica el contenido de agua en suelos y vegetación; penetra por nubes delgadas, presenta distintos tipos de vegetación en contrastes diferentes.
6	10.4 - 12.5	Infrarrojo térmico	Imágenes nocturnas son adecuadas para un mapeo térmico y para estimar el contenido de agua en suelos.
7	2.08 - 2.35	Infrarrojo medio	Coincide con la banda de absorción causada por iones hidróxilos en minerales.

Por su resolución espacial gruesa (120 m) la banda 6 se utiliza rara vez, pero es adecuada para un levantamiento térmico. Las demás bandas visibles y el IR se pueden combinar y asignar los colores azul, verde y rojo a las bandas de una combinación distinta con el propósito de formar una imagen en falso color. En total existen 120 combinaciones posibles de colores, del que prácticamente solamente se utilizan una pequeña cantidad para la mayoría de las aplicaciones. La combinación óptima de bandas es determinada por el terreno, el clima y el objetivo de la interpretación. Para el caso que nos interesa no haremos combinaciones de bandas, utilizaremos las bandas 1, 2, 3, 4, 5 y 7 tal como están y tampoco usaremos la 6 porque de ninguna manera trataremos con la temperatura del territorio al que corresponde la imagen.

2.7 El proceso de clasificación

Como se mencionó en el capítulo introductorio, esta tarea se lleva a cabo después de darle un tratamiento a la imagen y consiste en definir primeramente las categorías (a esta parte se le llama fase de entrenamiento) y luego en etiquetar cada pixel de la imagen en una de esas categorías (fase de asignación).

Tradicionalmente se han dividido los métodos de clasificación en dos grupos: supervisado y no supervisado, de acuerdo a la forma en que son obtenidas las estadísticas de entrenamiento. El método supervisado parte de un conocimiento previo del terreno, a partir del cual se seleccionan muestras para cada una de las categorías. Por su parte, el método no supervisado procede a una búsqueda automática de grupos de valores homogéneos dentro de la imagen. En este caso, queda al usuario la labor de encontrar correspondencias entre esos grupos y sus categorías de interés. Estos métodos son abordados con más detalle en los siguientes apartados y en general, veremos las formas en que los softwares comerciales hacen todo esto.

2.7.1 Fase de entrenamiento

Como se ha mencionado, diversos factores introducen una cierta dispersión en torno al comportamiento espectral medio de cada cubierta. Esto implica que las distintas categorías no se definan por un sólo ND, sino por una serie de ND más o menos próximos entre sí. Por ello, la fase de entrenamiento resulta ciertamente compleja: se trata de definir con rigor cada una de las categorías que pretenden discriminarse, teniendo en cuenta su propia variabilidad en la zona de estudio. Esto se logra seleccionando una muestra de píxeles de la imagen que represente adecuadamente a la categoría de interés. A partir de esos píxeles pueden calcularse los ND medios y la variabilidad numérica de cada categoría en todas las bandas que intervienen en la clasificación. El objetivo es identificar debidamente las distintas categorías, reduciendo a lo imprescindible los trabajos de campo. Al igual que en otras aplicaciones de muestreo, las estimaciones posteriores se basan sobre la muestra seleccionada, por lo que una incorrecta selección de ésta conduciría inexorablemente a pobres resultados en la clasificación posterior. En definitiva, la fase de entrenamiento constituye la columna vertebral de la clasificación numérica.

Antes de mencionar las formas de hacer el entrenamiento, conviene introducir una reflexión sobre el tipo de categorías que pueden intervenir en una clasificación temática. Suelen distinguirse, dos tipos de clases: informacionales y espectrales [27].

Las primeras son las que constituyen la leyenda de trabajo que pretende deducir el intérprete (tipos de ocupación de suelo, por ejemplo). Las segundas corresponden a los grupos de valores espectrales homogéneos dentro de la imagen, en función de ofrecer una reflectividad similar para las bandas consideradas y en la fecha concreta de la imagen. Las categorías informacionales se establecen por el usuario, mientras que las espectrales se deducen de los ND presentes en la imagen.

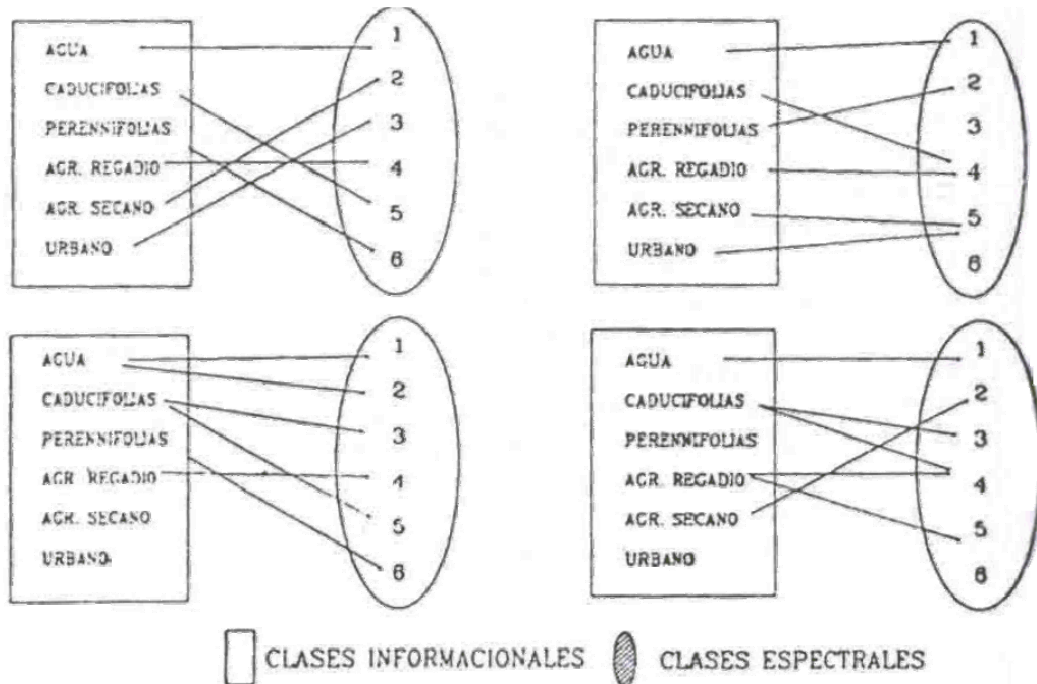


Figura 2.6 Relaciones entre clases informacionales y espectrales

Idealmente habría que producirse una correspondencia biunívoca entre clase de cobertura y clase espectral (Fig. 2.6, arriba izquierda); desgraciadamente este caso es muy poco frecuente, produciéndose, normalmente alguna de las siguientes situaciones:

- Que una categoría de cubierta esté expresada en varias clases espectrales (Fig. 2.6, abajo izquierda). Por ejemplo, el agua podría aparecer en varios grupos, de acuerdo a su profundidad o turbidez. En estos casos, basta perfeccionar el muestreo, de tal forma que se defina correctamente la dispersión espectral de cada clase. Si los grupos espectrales a que corresponden son muy variados puede dividirse una categoría informacional en varias subclases, para luego fundirlas (tras la clasificación) en una sola especie temática. Esto evita contar con clases muy heterogéneas espectralmente, y por tanto fácilmente sujetas a confusión con otras vecinas.
- Otro caso sería que dos o más categorías informacionales compartan una sola clase espectral (Fig. 2.6, arriba derecha). Lo que significaría que algunas categorías no son discriminables a partir de la imagen con que se cuenta. En este caso, la opción más razonable sería optar por una clave más general, evitando descender a un nivel de desagregación no abordable para los datos con que se cuentan. Por ejemplo, podría optarse por una clase denominada urbana, en lugar de distinguir distintas densidades de edificación. En caso de que las confusiones se den entre categorías no asimilables a una clave más genérica, habría que analizar una imagen perteneciente a otra fecha del año en donde esa confusión sea inevitable. Lógicamente también pueden intentarse otros medios, como refinar el método de muestreo o añadir alguna variable auxiliar.

- Por último, puede ocurrir que varias clases informacionales compartan clases espectrales (Fig. 2.6, abajo derecha). En este caso convendría replantear la estrategia de trabajo.

Ahora parece claro que el método supervisado pretende definir clases informacionales, mientras que el no supervisado tiende a identificar las clases espectrales presentes en la imagen.

2.7.1.1 Método supervisado

Parte de un cierto conocimiento de la zona de estudio, adquirido por la experiencia previa o por trabajos de campo. Esta familiaridad con la zona de estudio permite al intérprete delimitar sobre la imagen áreas piloto que considere representativas de las categorías que componen la leyenda. Estas áreas son los llamados campos de entrenamiento. La mayor parte de los paquetes de tratamiento digital permiten localizar interactivamente los campos de entrenamiento con ayuda del ratón eléctrico. El usuario indica en el monitor los vértices que los definen, teniendo en mente que también deben tratarse de áreas homogéneas en la imagen. Los trabajos de campo y otros documentos (como la fotografía aérea o cartografía convencional) pueden ser de gran ayuda para una localización más precisa. Acabada la delimitación de un campo de entrenamiento, pueden seleccionarse otros para la misma categoría, o bien culminar con él la definición de esa clase. En general, se sugiere seleccionar varios campos de entrenamiento por categoría, a fin de reflejar adecuadamente su variabilidad en la zona de estudio.

Finalizada la selección de las áreas de entrenamiento, la computadora calcula las estadísticas elementales de cada categoría a partir de los ND de los píxeles incluidos en los campos de entrenamiento asignados a cada clase. Lógicamente, este cálculo se aplica a todas las bandas que intervienen en la clasificación. Y definitivamente se asume que las medidas extraídas definen convenientemente a las clases. De aquí que sea crucial la correcta selección de los campos de entrenamiento. De otra forma, la clasificación estaría mal desde el inicio, al pretender discriminar categorías con escaso significado espectral.

Este proceso es una forma peculiar de muestreo espacial. Por ello, conviene tener en cuenta algunos criterios habituales en este tipo de técnicas para elegir el tamaño y la distribución más idónea de la muestra, así como para realizar estimaciones a partir de ella.

Respecto al tamaño, algunos autores recomiendan seleccionar un número de píxeles comprendido entre $10b$ y $100b$ por categorías, donde b indica el número de bandas que integran el análisis [23, 38]. El muestreo convencional considera que las muestras seleccionadas son aleatorias e independientes. Dicha asunción no es válida en este caso, lo que implica una fuerte correlación entre píxeles vecinos de un mismo campo de entrenamiento. Esta asociación supone que la información extraída de un número de píxeles contiguos es

menor que si fuera extraída de otros dispersos [9]. Por lo que resulta más conveniente elegir varios campos de pequeño tamaño que uno solo de mayor dimensión, ya que tendería a infravalorarse la variabilidad de esa categoría. También se puede optar por seleccionar píxeles aleatoriamente dentro de campos de mayor tamaño, en lugar de admitirlos a todos [6,7].

En cuanto a la distribución, conviene considerar las propias características de la imagen, intentando abarcar las variaciones espaciales que se presentan en cada categoría: orientación, pendiente, densidad, vigor, contenido de humedad, tipo de suelo, etc. Varios autores han determinado que la variabilidad de una cubierta es directamente proporcional a la resolución del sensor [11]. Es decir, cuanto mayor sea el detalle espacial de la imagen, mayor será también su sensibilidad para detectar las variaciones internas de una categoría (ocultas bajo un píxel de mayor tamaño). Un incremento en la resolución espacial puede complicar la clasificación digital [31] y el efecto se presenta más agudo en los tipos de cubierta más heterogéneos.

2.7.1.2 Método no supervisado

Este método se encarga de definir las clases espectrales presentes en la imagen, sin implicar conocimiento alguno del área de estudio, por lo que la intervención humana se centra más en la interpretación que en la consecución de los resultados.

En esta estrategia, se asume que los ND de la imagen forman una serie de agrupaciones o conglomerados (*clusters*). Estos grupos equivaldrían a píxeles con un comportamiento espectral homogéneo, que deberían definir clases temáticas de interés. Desgraciadamente estas categorías espectrales no siempre pueden equipararse a las clases informacionales que el usuario pretende deducir, por lo que resta a éste interpretar el significado temático de esas categorías espectrales.

El método para definir los agrupamientos espectrales se basa en un proceso de 3 pasos [24,18] :

- Selección de las variables que intervienen en el análisis, es decir, las bandas espectrales involucradas en la clasificación, ya sean originales o fruto de alguna transformación.
- Selección de un criterio para medir la similitud o distancia entre píxeles. Se han propuesto diversos criterios, el más utilizado es el de la distancia euclidiana, aunque en la literatura especializada se citan otros criterios, como sería el caso de la distancia media o la de Mahalanobis [46].
- Selección de un criterio para agrupar los píxeles similares. Las opciones para hacer esto también son numerosas, pero uno de los algoritmos más extendidos es el denominado ISODATA [13]. En esencia, realiza lo siguiente:

- a) Se señalan una serie de centros de clase, de acuerdo al número y forma indicados por el usuario.
- b) Se asignan todos los píxeles de la imagen al centro de la clase más próxima.
- c) Se calculan de nuevo los centros de clase, considerando los valores de los píxeles incorporados en la fase anterior.
- d) Se vuelven a asignar todos los píxeles de la imagen al centro más cercano, la cual no tiene por qué coincidir con la asignación realizada en (b).
- e) Vuelven a calcularse los centros de clase, teniendo en cuenta las nuevas incorporaciones y se regresa al (d).

El proceso se repite iterativamente hasta que el centro de clase no se desplace significativamente, lo que indica un buen ajuste a los grupos presentes en la imagen. Antes de cada iteración se verifican algunos parámetros de control definidos por el usuario: ¿se alcanzó el número máximo de grupos? (para interrumpir el agrupamiento), ¿algún grupo tiene un número de píxeles inferior al mínimo señalado? (para suprimirlo), ¿algunos grupos están separados entre sí a una distancia inferior a la mínima señalada? (para fusionarlos), ¿algún grupo cuenta con una varianza superior a la máxima señalada? (para que se particione). Una vez solventadas estas cuestiones, se procede a comenzar un nuevo bucle o bien, se finaliza el desplazamiento entre grupos.

En ocasiones resulta artificioso señalar algunos parámetros de control significativos, puesto que el usuario no conoce de partida el número de grupos espectrales presentes en la imagen, ni el valor idóneo de dispersión interna o de distancia entre grupos. Así que el análisis no supervisado se convierte en un proceso interactivo, donde el usuario establece unos valores arbitrarios que se van modificando a la vista de los resultados.

Tanto el método supervisado como el no supervisado cuentan con ventajas y desventajas [20, 25]. El primero resulta muchas veces subjetivo, pues el usuario establece a priori unas categorías sin valorar las características espectrales de la imagen. El segundo, no garantiza que las clases deducidas tengan significado para el usuario del producto, además de ser de difícil verificación. Con objeto de aminorar los inconvenientes de ambos métodos, se proponen diversas alternativas que combinan ambos métodos.

Independientemente del método empleado para definir la fase de entrenamiento, conviene reflexionar sobre la discriminabilidad de las categorías seleccionadas. En otras palabras, es preciso evaluar la viabilidad de que esas categorías puedan clasificarse sin grave riesgo de error. Si se comprueba que dos o más son muy similares habrá una gran probabilidad de confundirlas, por lo que se aconseja lo siguiente: confirmar que las estadísticas de entrenamiento hayan sido correctamente deducidas, adoptar una leyenda más general, con nuevas categorías que supongan una mezcla de las que ofrecen mayor riesgo de confusión, y recabar información auxiliar o imágenes de la misma zona con otras fechas.

2.7.2 Fase de asignación

Una vez culminada la selección y estudio de las categorías que intervendrán en la clasificación, se inicia la fase de asignación. Esto se realiza en función de los ND de los píxeles en cada una de las bandas que intervienen en el proceso. Fruto de esta fase es una nueva imagen, cuyos ND expresan la categoría en la cual se ha incluido cada uno de los píxeles de la imagen original.

Los criterios más comunes para establecer la asignación son: mínima distancia, por el cual el píxel se asigna a la clase más cercana; paralelepípedos, que permite señalar al usuario unos umbrales de dispersión asociados a cada clase; y máxima probabilidad (o máxima verosimilitud), en donde el píxel se asigna a aquella clase con la que posee mayor probabilidad de pertenencia.

Estos criterios de clasificación permiten obtener resultados relativamente rápidos, aunque muchas veces no son muy confiables, por ejemplo, podemos encontrar la región de una misma clase salpicada por píxeles asignados a otras categorías, perdiéndose su unidad geográfica y dificultando la interpretación de los resultados. Este efecto se debe a una doble causa: por un lado, a que la mayor parte de los criterios de asignación no tienen en cuenta la pertenencia de píxeles vecinos y por otro, a la propia complejidad del paisaje.

Para solventar este problema, suele aplicarse una suavización, el objetivo es reasignar aquellos píxeles aislados o poco significativos, resultando una distribución espacial más nítida de las distintas categorías. Para esto, se procede a comparar la clase asignada a cada píxel con las de los píxeles vecinos. Si la clase es similar a las del entorno, se mantiene la asignación original, en otro caso se modifica. Normalmente se realiza la comparación a partir de una ventana móvil de 3x3 píxeles y se distinguen dos tipos de filtro:

- Filtro modal. El píxel situado en el centro de la matriz se asigna a la categoría más frecuente entre los vecinos, siempre que no pertenezca ya a ella.
- Filtro de mayoría. El píxel central se asigna a la categoría que cuente con la mitad más uno de los píxeles vecinos. En caso de que ninguna categoría rebese este valor, el píxel central mantiene su asignación original. [8]

El siguiente capítulo aborda un modelo que además de contar con otras características importantes, considera las categorías de píxeles vecinos, por lo que no es necesario realizar ningún tipo de filtrado a los resultados de la clasificación.

Capítulo 3

Método Propuesto

A diferencia del proceso de clasificación comúnmente utilizado en percepción remota, el modelo que se presenta en [34, 35] (el MRMF, mostrado en los apéndices A y B) tiene como objetivo estimar de manera simultánea las estadísticas elementales de cada categoría y la segmentación de la imagen. En este capítulo veremos en forma breve las herramientas de dicho modelo, su extensión al caso multibanda y sobretodo, cómo definir las firmas espectrales de inicialización del método.

3.1 Herramientas

El método utiliza la teoría de Campos Aleatorios Markovianos para introducir en forma de función de distribución de probabilidad, información sobre la coherencia espacial de la imagen, al proceso de segmentación. Interpretando esto desde un punto de vista físico, podemos imaginar que tenemos un sistema de partículas interactuando entre sí, donde el efecto que produce la energía del sistema sobre cierta partícula se reduce al efecto que le produce su vecindad. Entonces, eligiendo ciertas funciones de energía potencial y definiendo la vecindad, se puede exigir que cada partícula sea un tanto similar a sus vecinas. Ahora veamos cómo lo hace MRMF (consulte los detalles en la sección A.2).

El método genera primero un campo aleatorio markoviano p de vectores, donde el vector correspondiente a cada pixel $p(r)$ es interpretado como una medida de probabilidad discreta sobre el espacio de estados de etiquetas (el vector tiene K entradas, donde K indica la cantidad de categorías/etiquetas en cuestión). Luego genera el campo de etiquetas f , de modo que $f(r)$ es una muestra independiente de la distribución $p(r)$; f no es un CAM, aunque p sí lo sea, así que el modelo f difiere de los clásicos (mencionados en el primer capítulo, ver apéndice B). Sin embargo, las funciones potenciales para el campo p se usan para obtener coherencia espacial en las regiones $\{R_1, \dots, R_K\}$. En este caso, los potenciales fueron definidos en la ecuación (11) del apéndice A como $V_{\langle r, s \rangle}(p(r), p(s)) = \lambda |p(r) - p(s)|^2 = \lambda \sum_{k=1}^K (p_k(r) - p_k(s))^2$ donde $\langle r, s \rangle$ son pixeles vecinos y λ es el parámetro positivo que indica qué tanto deben parecerse $p(r)$ y $p(s)$. Como el campo p no es directamente observable, suele llamársele modelo de CAM oculto.

Después de aplicar la regla de Bayes y hacer algunos cálculos vemos que para encontrar el estimador óptimo de f hay que realizar el siguiente procedimiento de 2 pasos:

1. Obtener los estimadores MAP p^* , θ^* , es decir, minimizar la función de energía:

$$U(p, \theta) = -\sum_{r \in L} \log(v(r) \cdot p(r)) + \sum_C V_C(p) - \log P_\theta(\theta)$$

donde p es el campo de medidas aleatorio markoviano oculto y θ es el arreglo de K vectores con NB entradas que define las firmas espectrales (NB representa la cantidad de bandas en la imagen, por lo que $\theta_{k,b}$ contiene la media de la clase k en la banda b).

2. Hallar la moda de cada medida discreta $p^*(r)$ en forma desacoplada $f^*(r) = \arg \max_k p_k^*(r)$.

La carga computacional recae en el primer paso, pero como la función es diferenciable, se puede minimizar muy eficientemente utilizando una técnica de descenso Newtoniano (ver la sección A.3).

3.2 Algoritmo de minimización de la función de energía en el caso multibanda

Veamos primero algunos detalles de la función de energía:

- En el caso multibanda, la verosimilitud para cada pixel se calcula de la siguiente manera:

$$v_k(r) = \frac{1}{(2\pi)^{NB/2} \sqrt{|S_k|}} \exp\left(-\frac{1}{2}(gr - \theta_k)^T S_k^{-1} (gr - \theta_k)\right) \quad \text{donde } NB \text{ indica el número de bandas de la imagen,}$$

$|S_k|$ es el determinante de la matriz de covarianzas de la clase k , S_k^{-1} es la inversa de la matriz de covarianzas y finalmente, gr y θ_k son vectores de NB entradas, gr contiene los niveles de gris del pixel r observados en cada una de las bandas y θ_k contiene los niveles de gris promedio de la clase k en cada banda. Considerando a S_k como una matriz diagonal constante (independiente de la clase) se puede reescribir la verosimilitud como $v_k(r) = \frac{1}{Z} \exp(-\beta(gr - \theta_k)^T (gr - \theta_k))$, donde Z es una constante de normalización y podemos interpretar a β como un parámetro que depende de la varianza del ruido en las imágenes;

- Los potenciales son los indicados en el segundo párrafo de la sección 3.1; y
- $P_\theta(\theta)$ equivale a cualquier constante, ya que no existen restricciones sobre el campo θ .

La técnica de minimización de la función de energía utilizada en MRMF es descenso Newtoniano, un método de descenso de gradiente con inercia, en el que es posible moverse hacia el mínimo con pasos de tamaño fijo y acelerar la convergencia calibrando el coeficiente de fricción. Sólo que en cada iteración, $p_k(r)$ debe proyectarse de nuevo al simplejo de K vértices S_K , $S_K = \{u \in \mathfrak{R}^K : \sum_{k=1}^K u_k = 1, u_k \geq 0, k = 1, \dots, K\}$.

Entonces, el algoritmo que permite encontrar el estimador óptimo de f es el siguiente:

1. inicializar modelo de firmas espectrales (las θ_k)
2. calcular verosimilitudes con dichas firmas (las $v_k(r)$)
3. inicializar vectores de medidas ($p_k(r)$)

4. definir parámetros de descenso Newtoniano para minimizar la función de energía con respecto al modelo de firmas espectrales (tamaño de paso $h1$ y coeficiente de fricción $\alpha1$)
5. definir parámetros de descenso Newtoniano para minimizar la función de energía con respecto al modelo p (tamaño de paso $h2$ y coeficiente de fricción $\alpha2$)
6. hasta que converja el método hacer DescensoNewtoniano (aproximadamente unas 200 iteraciones)
7. realizar fase de asignación

En la sección 3.3 mostramos dos propuestas para inicializar las firmas espectrales (paso 1); la inicialización de los vectores de medidas se refiere a asignar a todas las $p_k(r)$ el mismo valor, de modo que $\sum_k p_k(r) = 1$.

Los parámetros definidos en los pasos 4 y 5 deben ser calibrados de manera que permitan obtener buenos resultados y cabe mencionar que no hay razón para que $h1$ deba ser igual a $h2$ o $\alpha1$ igual a $\alpha2$. Una manera de hacer esto es generando una imagen sintética y echar a andar el MRMF donde el modelo inicial de firmas espectrales sea precisamente el modelo con el que se creó la imagen sintética. Así, habrá que ajustar los parámetros hasta que se obtenga la clasificación que se desea. Una vez calibrados los parámetros, podrá usarlos para realizar la segmentación de otras imágenes y obtener sus modelos de firmas.

El paso 6 se refiere a hacer lo siguiente:

$$\theta_{k,b}^{(t+1)} = \frac{2}{\alpha1 * h1 + 1} \theta_{k,b}^{(t)} + \frac{\alpha1 * h1 - 1}{\alpha1 * h1 + 1} \theta_{k,b}^{(t-1)} - \frac{h1^2}{\alpha1 * h1 + 1} \nabla_{\theta_{k,b}} U(p^{(t)}, \theta^{(t)}) \quad \forall \theta_{k,b},$$

$$\tilde{p} = \frac{2}{\alpha2 * h2 + 1} p^{(t)} + \frac{\alpha2 * h2 - 1}{\alpha2 * h2 + 1} p^{(t-1)} - \frac{h2^2}{\alpha2 * h2 + 1} \nabla_p U(p^{(t)}, \theta^{(t)})$$

donde $\theta_{k,b}$ indica la media de la clase k en la banda b y el vector \tilde{p} debe ser proyectado al simplejo de K vértices S_K como se muestra enseguida. Finalmente, la fase de asignación (paso 7) consiste en etiquetar cada $f(r)$ con la k que corresponde a la máxima $p_k(r)$.

Algoritmo de proyección:

1. Definir un vector de K entradas $x = \tilde{p}(r)$

2. Mientras $x \notin S_K$ hacer:

- a. $x_k = x_k - \frac{\sum_{i=1}^K x_i - 1}{K}$, donde $1 \leq k \leq K$

- b. inicializar un vector de K entradas $modifica$, donde $modifica_k = true$, $1 \leq k \leq K$

c. desde $k=1$ hasta $k=K$ hacer

si $x_k < 0$

{ $x_k = 0$

$modif\acute{c}a_k = false$

$x_j = x_j - \frac{\sum_{i=1}^K x_i - 1}{K}$ donde $1 \leq j \leq K$ y $modif\acute{c}a_j = true$

}

3.3 Cómo inicializar el modelo de firmas espectrales

En esta sección se presentan dos propuestas para inicializar el modelo de firmas espectrales, la primera es una técnica semiautomática, muy parecida a la que se realiza en percepción remota clásica; y la otra es completamente automática, utiliza el MRMF para hacer segmentaciones monobanda y parte de ellas para generar el modelo inicial. Veamos los detalles.

3.3.1 Modelo inicializado mediante muestras aleatorias

Este método consiste en generar muestras representativas de las K clases que se desean identificar en la imagen y calcular sus estadísticas elementales en cada banda, sólo que en este caso no es el usuario del sistema quien elige las zonas de entrenamiento, sino que son generadas de manera aleatoria.

Primero, el usuario define el tamaño de la ventana que capturará las muestras. Luego, para cada clase k se genera una muestra aleatoria sobre la imagen y se despliegan en pantalla los histogramas, medias y desviaciones estándares de la muestra en las NB bandas.

Visualizando los histogramas, el usuario puede identificar si la zona entrenada es homogénea o no; además, con las desviaciones estándares puede darse una idea de la dispersión de los datos y entonces decidir si es conveniente usar dichos parámetros en el modelo inicial de la segmentación multibanda. En caso de que no desee considerarlos, se generará otra muestra aleatoria para la misma clase k , y así, hasta concluir con la última clase.

La siguiente pantalla se presenta para ejemplificar el procedimiento.

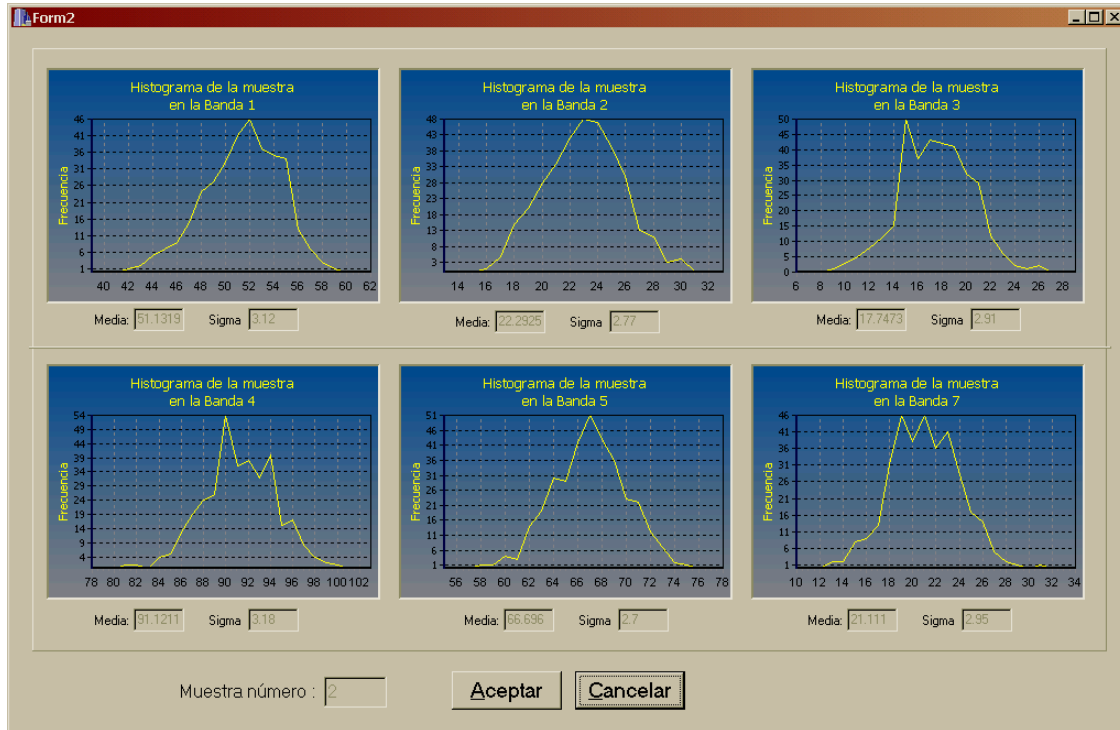


Figura 3.1 Representación de muestras aleatorias

Esta forma de clasificación semiautomática permite obtener buenos resultados. Lo malo es que en ocasiones hay que esperar un poco a que el sistema encuentre una zona homogénea.

3.3.2 Modelo inicializado mediante segmentaciones monobanda

Este método es completamente automático y consta de dos etapas, en la primera de ellas se utiliza el MRMF para realizar la segmentación de cada una de las bandas de manera independiente; para esto, se define el número de clases $K_{monobanda}$ que se desean identificar en esta fase.

Como modelo de inicio para dichas segmentaciones monobanda, se calcula el rango dinámico de la banda que se desea segmentar, se divide entre la cantidad de clases $K_{monobanda}$ y así se obtienen las medias iniciales.

Al concluir con las NB segmentaciones, se inicia la siguiente etapa, ésta consiste en realizar la segmentación multibanda tal y como se indica en el algoritmo de la sección 3.2. Pero para esto, el modelo de firmas espectrales se obtiene de la siguiente manera:

Primero se genera una imagen que combina las NB segmentaciones monobanda, de modo que esta nueva imagen contiene todas las particiones identificadas en cada segmentación monobanda. La Fig. 3.2 muestra un ejemplo, donde cada región de esta nueva imagen es identificada por una de las $K_{monobanda}^{NB}$ combinaciones posibles.

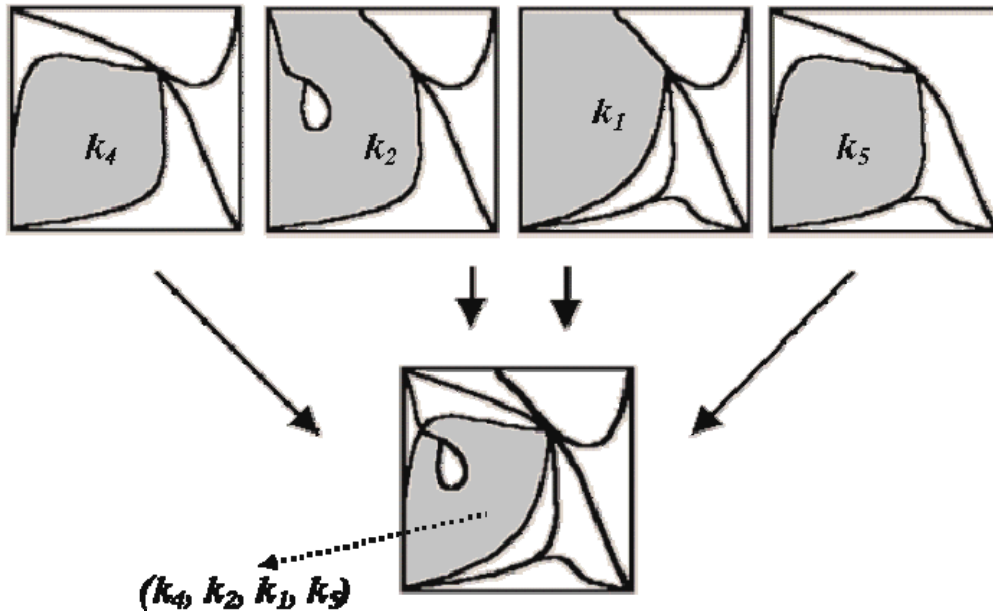


Figura 3.2 Regiones formadas por la combinación de segmentaciones monobanda

Una vez identificadas las combinaciones que forman las regiones de la nueva imagen, se toman las medias (por banda) de las $K_{multibanda}$ regiones de mayor soporte; donde $K_{multibanda}$ indica el número de clases que se desean detectar en la segmentación multibanda. Por lo que ya tenemos otro modelo para iniciar la clasificación multibanda.

Cabe mencionar la importancia de definir el número de clases para hacer la segmentación monobanda; ya que entre más grande sea este valor, mucho mayor será la cantidad de regiones posibles a considerar para modelar los parámetros de la segmentación multibanda; definitivamente NO hay razón para elegir muchas clases en esa primera etapa, más adelante veremos algunos ejemplos que sustentan este hecho.

Capítulo 4

Pruebas con imágenes sintéticas

Para realizar segmentación de imágenes hay que considerar que existen dos formas de modelar parámetros: con modelos fijos o con modelos variables. Veamos un ejemplo para diferenciarlos.

Cuando se tratan imágenes del cerebro (de resonancia magnética) se tienen las clases: *materia blanca*, *materia gris* y *líquido cefalorraquídeo*. En el caso particular de la *materia gris*, existen pixeles donde ésta se ve de un tono más claro que en otros sitios; lo que lleva a definir un modelo de parámetros que identifique a los pixeles de ambos casos como pertenecientes a la clase *materia gris*. Éste es un caso donde el modelo de parámetros de la clase *materia gris* es variable. Por otro lado, existen imágenes donde las clases se visualizan relativamente homogéneas independientemente de la posición donde se encuentren, por lo que es suficiente definir para ellas un modelo de parámetros fijos.

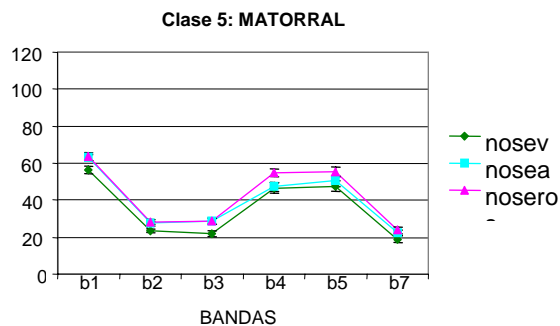
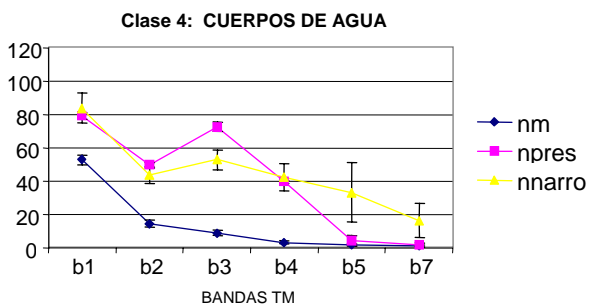
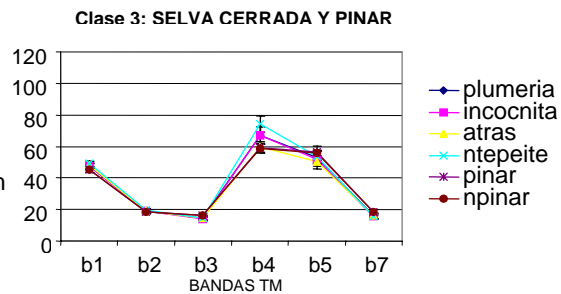
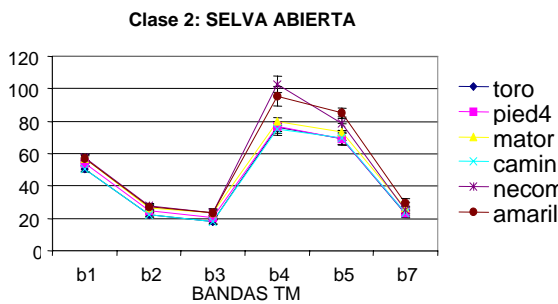
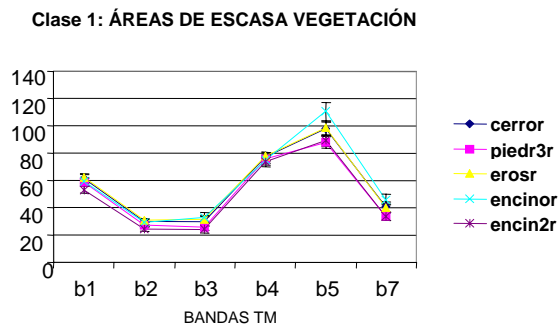
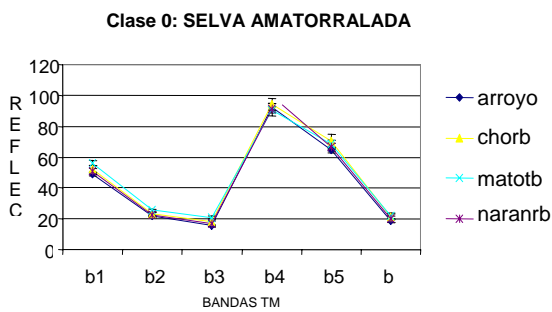
En este capítulo veremos resultados de segmentar imágenes sintéticas de 6 bandas con el método de MRMF, utilizando modelos de parámetros fijos.

4.1 La imagen sintética

Con el propósito de calibrar los parámetros del método MRMF y de valorar su eficiencia con imágenes multibanda, se planeó realizar una serie de pruebas con imágenes sintéticas.

Para esto, había que definir el modelo de parámetros para las clases, por lo que se consideraron las categorías que la M.C. Aurora Breceda identificó en su trabajo de tesis doctoral [4]. Cuando ella llevó a cabo la clasificación de la imagen satelital correspondiente a la zona de la reserva de la biósfera Sierra de La Laguna, encontró 36 firmas espectrales distintas y con ellas formó 10 grupos/categorías/clases, de modo que cada categoría agrupara firmas espectrales parecidas, ya que dichas firmas corresponden a una misma leyenda.

Ahora bien, el modelo de firmas espectrales que se definió finalmente para la imagen sintética, corresponde a los promedios de 6 de las 10 categorías que la M.C. Aurora Breceda menciona en su documento de tesis. A continuación se muestran dichas categorías y más adelante, los promedios:

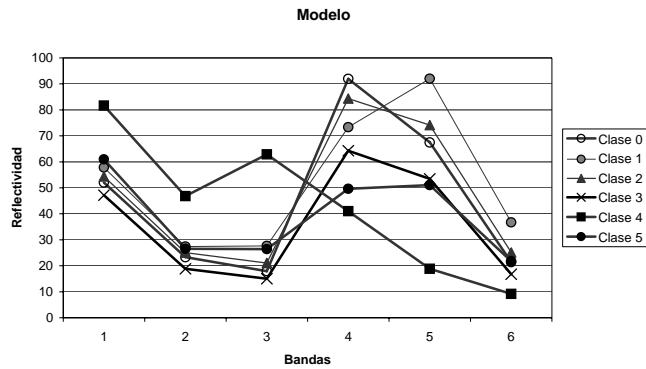


El modelo que se muestra en la tabla 4.1 indica los valores medios de los pixeles de cada clase en cada banda.

Tabla 4.1 Modelo de firmas espectrales de la imagen sintética

	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 6
Clase 0	52.04	23.3	17.91	91.97	67.45	21.35
Clase 1	57.83	27.37	27.62	73.33	92.03	36.65
Clase 2	54.42	25.04	21.04	84.4	74.19	25.09
Clase 3	47.15	18.86	15	64.33	53.45	16.71
Clase 4	81.69	46.75	62.88	41	18.85	9.2
Clase 5	60.98	26.5	26.41	49.62	51.03	21.74

Ahora veamos la gráfica del modelo:



A partir de dicho modelo, se definió un campo de etiquetas ideal, mostrado en la Fig. 4.1; donde cada tono de gris corresponde a una clase en particular y el número dentro de cada región en la imagen indica el número de la clase a la que pertenece la región.

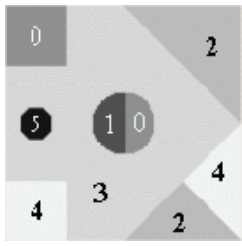


Fig. 4.1 Imagen Sintética

A lo largo de este capítulo veremos distintos resultados de segmentar esta imagen sintética; en ocasiones vamos a comparar las firmas espectrales obtenidas en las pruebas contra las originales. En esos casos, sólo veremos las gráficas de los modelos, sin indicar a qué clases pertenecen las firmas espectrales suponiendo que no será necesario, ya que cada clase marcará sus valores medios de la misma manera en que se mostraron en la gráfica anterior. Es decir, la clase correspondiente a la región 0 siempre se va a graficar con un círculo vacío; la de la región 2 siempre se va a graficar con un triángulo relleno, la de la región 4 con un cuadrado relleno, etcétera.

Ahora veamos las bandas de esta imagen:

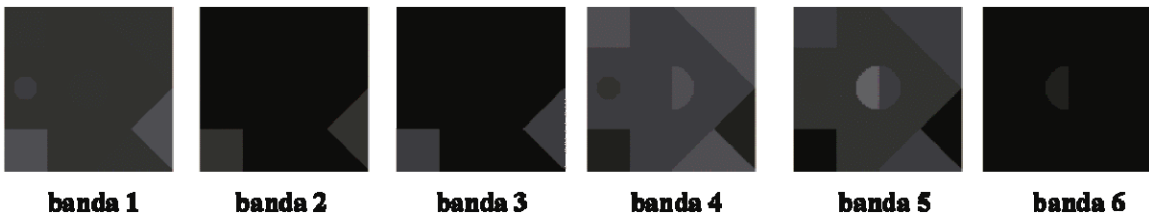


Figura 4.2 Bandas de la imagen sintética

El hecho de agregar ruido gaussiano de manera independiente a cada una de las bandas de la imagen, genera variaciones en el rango dinámico (ya que cada pixel en cada banda es modificado en un rango de $\pm 4\sigma$ aproximadamente); la tabla 4.2 muestra el efecto de agregar distintos niveles de ruido blanco a la imagen sintética:

Tabla 4.2 Rango dinámico de la imagen sintética con distintos niveles de ruido blanco.

	$\sigma = 0$ (sin ruido)		$\sigma = 5$		$\sigma = 7$		$\sigma = 8$		$\sigma = 10$	
	min	max	min	max	min	max	min	max	min	max
banda 1	47	81	27	96	18	101	16	108	5	112
banda 2	18	46	0	65	0	73	0	72	0	79
banda 3	15	62	0	78	0	93	0	88	0	96
banda 4	41	91	20	107	21	113	7	117	10	129
banda 5	18	92	2	107	0	109	0	116	0	120
banda 6	9	36	0	53	0	52	0	60	0	77

4.2 Calibración de parámetros del método de Descenso Newtoniano

En esta parte, el objetivo es calibrar los parámetros del método de Descenso Newtoniano (ver sección 3.2). Para ello, se realiza primero la segmentación de la imagen sintética sin ruido, inicializando el modelo de firmas espectrales con los datos de la tabla 4.1 y la tarea consiste en observar los resultados de la segmentación y ajustar a los parámetros hasta que se obtenga la segmentación perfecta. Definitivamente esta tarea es la debilidad del método pero es primordial hacerla con cuidado.

Una vez que se obtienen los parámetros del método de Descenso Newtoniano para la imagen sin ruido, hay que probarlos para una imagen con un poco de ruido y si para esa imagen también se obtiene una buena clasificación, puede probarlos para otra imagen con un poco de más ruido.

Después de efectuar varias pruebas, pudimos observar que se obtiene una clasificación perfecta (incluso para imágenes con ruido de $\sigma=7$) con los valores que se indican a continuación:

$h1$ (tamaño de paso para minimizar con respecto al modelo de firmas) = 0.015

$\alpha1$ (coeficiente de fricción para minimizar con respecto al al modelo de firmas) = 7.8

$h2$ (tamaño de paso para minimizar con respecto al campo de medidas) = 0.08

$\alpha2$ (coeficiente de fricción para minimizar con respecto campo de medidas) = 5

Estos parámetros ya no hay que moverlos. Para las siguientes pruebas sólo hay que definir los parámetros λ (parámetro de coherencia espacial³) y β (coeficiente en la exponencial del cálculo de la verosimilitud⁴) según el caso.

³ indica qué tanto deben parecerse los pixeles vecinos

⁴ depende de la varianza del ruido en las imágenes

Pruebas :

- a) La figura 4.3 muestra los resultados de segmentar con $\lambda = 1$ y $\beta = 0.035$ las imágenes sintéticas con ruido de $\sigma = 0$ a la izquierda (sin ruido), $\sigma = 5$ en el centro y $\sigma = 7$ a la derecha.



Figura 4.3 Resultados de la prueba 4.2_a

Podemos notar que se obtuvieron muy buenos resultados considerando el hecho de que añadir ruido gaussiano de $\sigma = 7$ en cada banda implica modificar el nivel de gris de cada pixel en un rango de ± 28 .

- b) Clasificación de la imagen con ruido gaussiano de $\sigma = 10$.

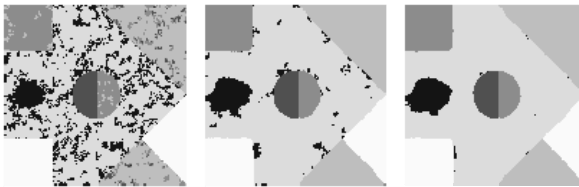


Figura 4.4 Resultados de la prueba 4.2_b

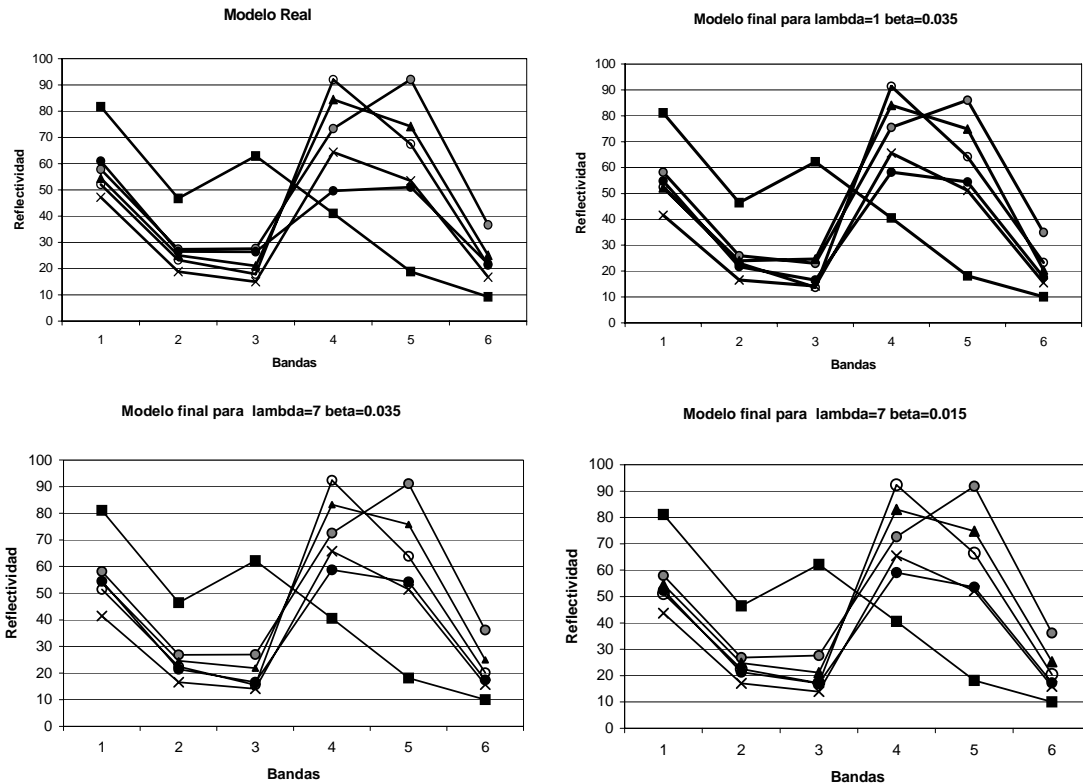
Parámetros:

Izquierda $\lambda=1, \beta=0.035$

Centro $\lambda=7, \beta=0.035$

Derecha $\lambda=7, \beta=0.015$

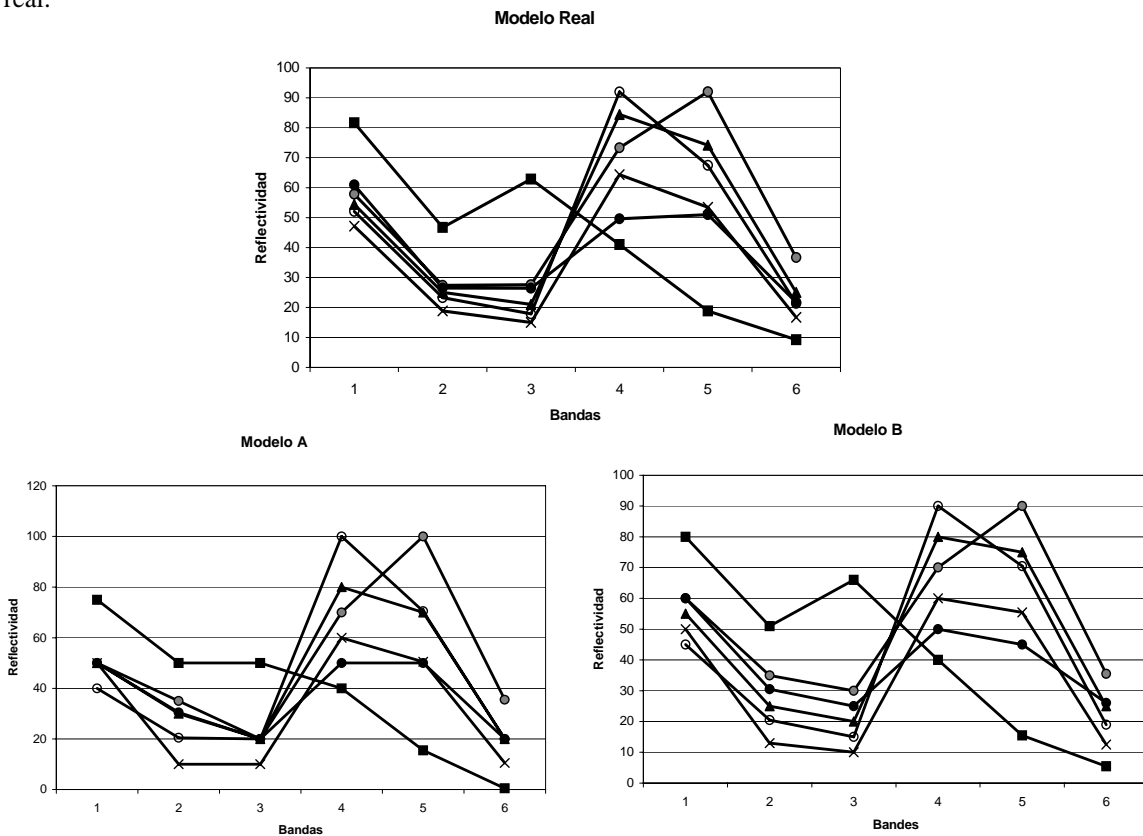
Para cada resultado de la fig. 4.4, comparemos el modelo de firmas espectrales al iniciar la segmentación (modelo real) contra el que se obtuvo al concluirla:



Antes de concluir cualquier cosa, debemos considerar que el nivel de ruido es grande, lo que provoca que el modelo se desajuste; ahora bien, en la prueba con $\lambda=1$ no estamos dándole mucha importancia a la coherencia espacial, por lo que obtenemos mucha granularidad en la clasificación. Notamos, por ejemplo, que la clase correspondiente a la región 5 (marcada en la gráfica con círculos rellenos de negro) se acercó a la de la región 3 (marcada con cruces) en todas las bandas; la de la región 2 (triángulos rellenos) está cerca de la clase de la región 0 (círculos vacíos) en las bandas 1, 2 y 6 y de la clase de la región 1 (círculos rellenos de gris) en las bandas 2 y 3; la de la región 4 (con cuadrados) no tiene con cuál confundirse y por tanto, esta zona de la imagen no tiene granularidad. Al aumentar el valor de λ reducimos la granularidad y logramos desacoplar un poco algunas firmas (la 2 ya no está tan cerca ni de la 0 ni de la 1). Finalmente, al disminuir el valor de β se obtiene todavía menos granularidad (recuerde que este parámetro depende de la varianza del ruido e incluso puede decrementarse más para obtener mejores resultados).

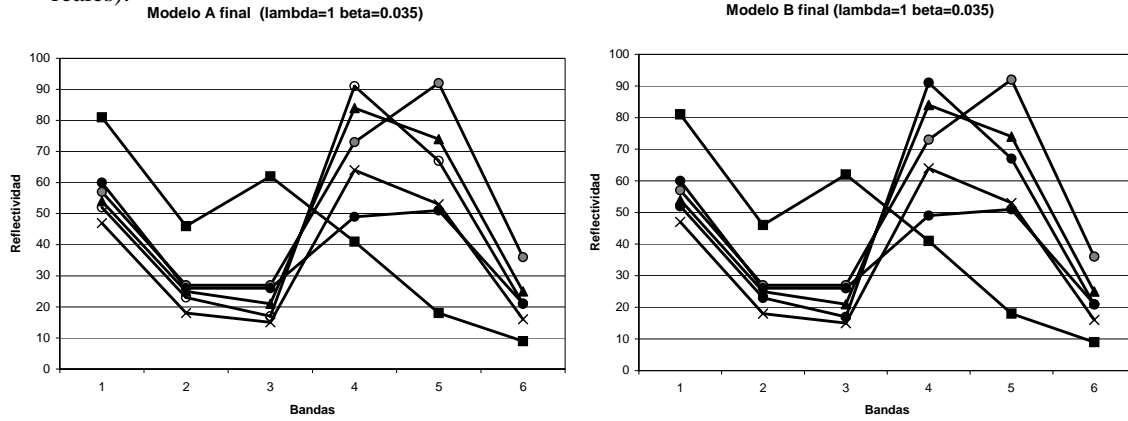
4.3 Prueba de convergencia del modelo de firmas

En esta sección veremos algunas pruebas de segmentación donde se muestra que el método es capaz de converger a las firmas espectrales originales en aquellos casos cuando se inicia con un modelo distinto del real.



Pruebas :

- a) Al segmentar la imagen sintética sin ruido partiendo de cualquiera de los modelos anteriores, la clasificación es perfecta y finalmente se obtienen las siguientes firmas (mucho muy cercanas a las reales):



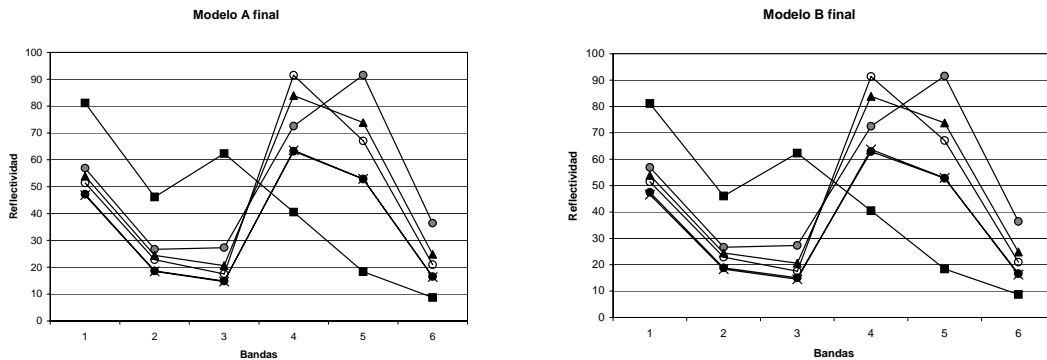
- b) Al segmentar la imagen sintética con ruido gaussiano de $\sigma = 5$ se obtuvieron los resultados de la figura 4.5. La imagen del lado izquierdo es la segmentación iniciando con el modelo A y la imagen derecha es el resultado de iniciar con el modelo B.



Parámetros: $\lambda=7$, $\beta=0.015$ en ambos casos.

Figura 4.5 Resultados de la prueba 4.3_b

En las siguientes gráficas podemos ver que casi todas las firmas se recuperaron correctamente, sólo la clase correspondiente a la región 5 (marcada en la gráfica con círculos rellenos de negro) se encimó a la clase de la región 3 (marcada con cruces).



Conclusiones :

En las secciones 4.2 y 4.3 notamos que una vez calibrados los parámetros del método de Descenso Newtoniano, el MRMF es capaz de obtener buenas clasificaciones de la imagen; incluso cuando se inicia con un modelo distinto al real, permitiendo al mismo tiempo que el modelo de firmas converja al modelo real de firmas espectrales.

Cabe mencionar que en algún momento se añadió a la función de energía un término más: $\gamma \sum_{k,b} |\theta_{k,b} - \hat{\theta}_{k,b}|^2$ con el que se penalizaba la diferencia entre el modelo de firmas espectrales inicial y el que se obtiene en cada iteración; pero nos dimos cuenta de que no era necesario incluirlo, ya que todos los experimentos se realizaron con un valor de $\gamma=0$ y el modelo final en los resultados siempre fue muy parecido al real (aunque se iniciara el método con un modelo muy distinto). Ahora veamos el comportamiento del método partiendo de modelos de firmas espectrales todavía más distintos al real.

4.4 Pruebas de segmentar la imagen sintética partiendo de un modelo estimado

Ya que calibramos los parámetros del método de Descenso Newtoniano y que sabemos que el modelo de firmas converge al modelo real, podemos considerar el caso más común en la práctica, en el que no se conocen los parámetros de las clases.

Para abordar este problema, en la sección 3.3 se propusieron dos métodos para estimar el modelo inicial, en esta sección veremos las pruebas correspondientes, donde los parámetros de Descenso Newtoniano se mantienen constantes con los valores definidos en la sección 4.2.

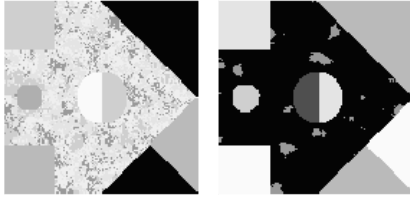
4.4.1 Modelo inicializado mediante muestras aleatorias

A continuación veremos los resultados obtenidos de segmentar las imágenes sintéticas partiendo del modelo de firmas espectrales inicializado con el método propuesto en la sección 3.3.1.

Pruebas:

- a) Segmentación de la imagen sintética con ruido gaussiano de $\sigma=3$ en cada banda.

Número de clases para la segmentación: **10**



Parámetros: $\lambda=1$, $\beta=0.035$ en ambos casos.

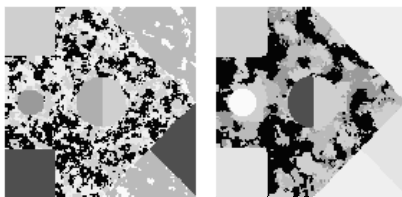
Figura 4.6 Resultados de la prueba 4.4.1_a

Para obtener el resultado del lado izquierdo se aceptaron las estadísticas de algunas muestras donde la desviación estándar en al menos una banda era mucho mayor a 3 (por ejemplo 10) y cuyos histogramas no necesariamente eran unimodales, es por eso que hay tanta granularidad en la región 3.

Para el resultado de la derecha sólo se permitieron muestras cuya sigma en cada banda fuera menor o igual a 3.5, lo que permitió disminuir la confusión entre clases. Ahora bien, otros casos que permiten disminuir la granularidad son aumentar la λ y/o disminuir la β . Para pruebas de este estilo, donde no se tiene un nivel de ruido muy alto, es posible obtener la clasificación perfecta variando esos parámetros.

- b) Segmentación de la imagen sintética con ruido gaussiano de $\sigma=5$ en cada banda.

Núm de clases para la segmentación: **10**



Parámetros:

$\lambda=1$, $\beta=0.035$ en el caso de la izquierda

$\lambda=3$, $\beta=0.035$ en el caso de la derecha.

Figura 4.7 Resultados de la prueba 4.4.1_b

En ambos casos se aceptaron para el modelo inicial, las medias de las muestras cuya sigma en cada banda fuera menor o igual a 7. En el primer caso se obtuvo granularidad en las regiones 2 y 3; pero al aumentar el parámetro λ sólo se obtuvo granularidad en la región 3, como lo muestra el segundo caso, donde se identifican perfectamente las otras 5 regiones de la imagen. Al aumentar más este parámetro disminuirá la granularidad en la región 3.

- c) Segmentación de la imagen sintética con ruido gaussiano de $\sigma=5$ en cada banda.

Núm de clases para la segmentación: **8**



Parámetros: $\lambda=3$, $\beta=0.035$

Para el modelo inicial, se aceptaron las medias de 8 muestras cuya sigma en cada banda fuera menor o igual a 7 y cuyos histogramas fueran unimodales. Se identificaron perfectamente 5 regiones de la imagen; sólo la región 3 tiene granularidad.

Fig. 4.8 Resultado de la prueba 4.4.1_c

Haciendo una comparación con la imagen derecha de la Fig. 4.7 podemos notar que en esencia produjeron el mismo resultado; la diferencia es que en la Fig. 4.8 hay menos clases confundiendo en la región 3, lo que puede solucionarse de dos maneras: aumentando más el parámetro λ o disminuyendo el número de clases para la segmentación.

- d) Segmentación de la imagen sintética con ruido gaussiano de $\sigma=5$ en cada banda.

Núm de clases para la segmentación: **6**



Parámetros: $\lambda=3$, $\beta=0.035$

En este caso sólo se aceptaron las muestras cuya sigma en cada banda fuera menor o igual a 6 y cuyos histogramas fueran unimodales.

Fig. 4.9 Resultado de la prueba 4.4.1_d

Podemos notar una gran mejoría respecto al resultado de la figura 4.8; aunque la región 3 de la imagen tiene pocos pixeles que fueron mal clasificados y la región 5 se ve un poco deformada.

Conclusiones :

Para obtener muy buenos resultados con esta propuesta de inicialización, hay que restringirse a aceptar las muestras donde la desviación estándar no sea muy grande (por lo menos no mayor a la σ del ruido que contiene la imagen) y donde los histogramas se vean unimodales; el único inconveniente de este método es que en ocasiones hay que esperar a que el sistema capture las muestras con dichas características.

4.4.2 Modelo inicializado mediante segmentaciones monobanda

A continuación veremos algunos resultados obtenidos de segmentar las imágenes sintéticas partiendo del modelo de firmas espectrales inicializado con el método propuesto en la sección 3.3.2

Pruebas:

- a) Segmentación de la imagen sintética con ruido gaussiano de $\sigma=3$ en cada banda.

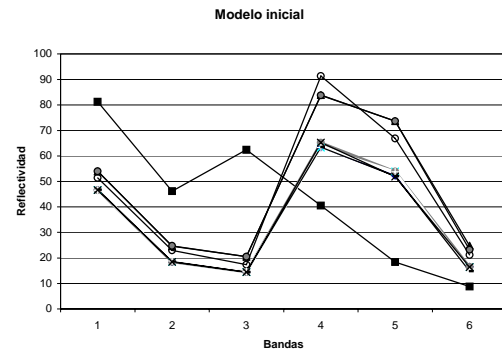
Número de clases para la segmentación monobanda: **10**

Clasificación obtenida por banda con $\lambda = 4.5$ y $\beta=0.015$:



Figura 4.10 Segmentaciones monobanda de la prueba 4.4.2_a

Después de formar la imagen que contiene a todas las particiones identificadas en las segmentaciones monobanda, se tomaron las medias de las **10** regiones con mayor soporte para iniciar la segmentación multibanda. La gráfica muestra dichas medias, sólo 4 clases son perfectamente distinguibles, las otras 6 no se alcanzan a ver porque están muy cercanas a éstas.



Con tales medias se inició la segmentación multibanda, obteniendo los resultados de la figura 4.11. Parámetros: $\lambda=3$ y $\beta=0.035$ (izq), $\lambda=5$ y $\beta=0.015$ (centro) y $\lambda=7$ y $\beta=0.015$ (der.) .



Figura 4.11 Segmentaciones multibanda de la prueba 4.4.2 a

Podemos notar que conforme aumentamos el parámetro λ la granularidad de la región 3 disminuyó. Teníamos firmas muy parecidas al iniciar la segmentación multibanda, pero al obligar que las etiquetas de los pixeles se parezcan a las de sus vecinos exigimos que algunas de esas firmas se queden sin soporte. Ahora bien, este hecho también provocó que la región 5 cambiara su forma sin conseguir una clasificación perfecta. Veamos qué sucede si aumentamos enormemente el parámetro λ en la segmentación monobanda y disminuimos el número de clases para cada tipo de segmentación.

- b) Segmentación de la imagen sintética con ruido gaussiano de $\sigma=3$ en cada banda.
 Número de clases para la segmentación por banda: 8

Clasificación obtenida por banda con $\lambda=25$ y $\beta=0.035$:



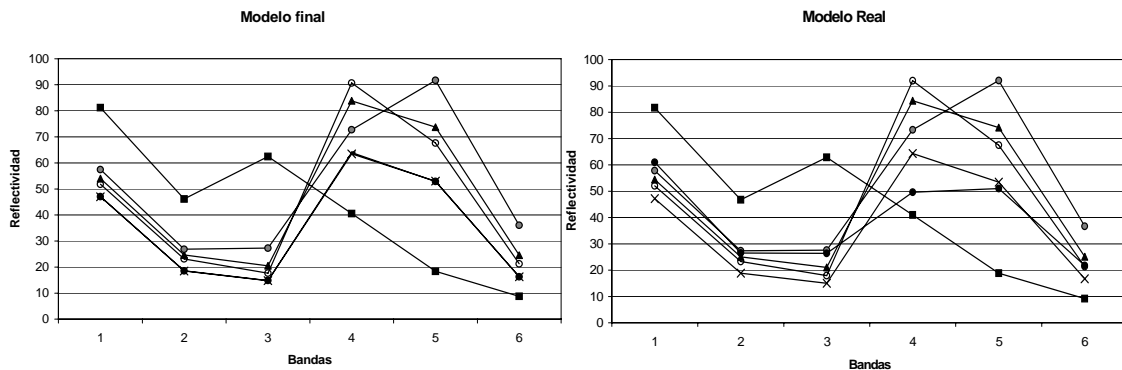
Figura 4.12 Segmentaciones monobanda de la prueba 4.4.2_b

Después de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las 6 regiones con mayor soporte para iniciar la segmentación multibanda. Con dichas medias y con los parámetros $\lambda=1.5$ y $\beta=0.035$ se obtuvo el resultado que muestra la figura 4.13.



Fig. 4.13 Segmentación multibanda de la prueba 4.4.2 b

Comparemos ahora el modelo final de este resultado con el modelo real de la imagen sintética:



Casi todas las clases se ajustaron de manera perfecta a las reales. La diferencia mínima entre medias reales y ajustadas es de 0.077, cuyo valor corresponde a la clase de la región 1 (marcada con círculos vacíos) en la banda 6; y las diferencias máximas son de 13.9 y 13.58, que corresponden a la clase de la región 5 (marcada con círculos rellenos de negro) en las bandas 1 y 4 respectivamente. También podemos notar que las clases de las regiones 5 y 3 finales, obtuvieron una firma "muy parecida" y aunque sus medias difieren en promedio tan solo un valor de 0.096, el sistema fue capaz de obtener la clasificación perfecta.

NOTA: En el caso de iniciar la segmentación multibanda considerando las medias de más de 6 regiones, también se obtiene una clasificación perfecta.

- c) Segmentación de la imagen sintética con ruido gaussiano de $\sigma=5$ en cada banda.
 Núm de clases para la segmentación por banda: **8**

Clasificaciones monobanda con $\lambda = 25$ y $\beta=0.035$:

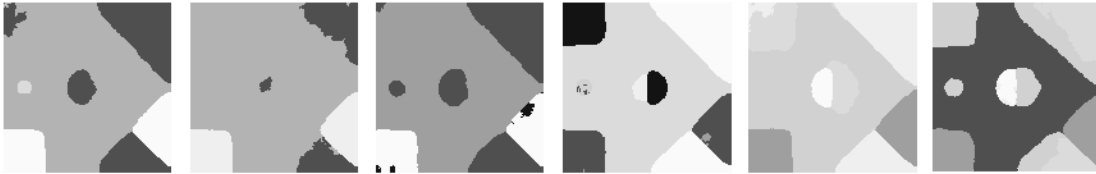


Fig. 4.14 Segmentaciones monobanda de la prueba 4.4.2 c

Después de considerar una imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las **6** regiones con mayor soporte para iniciar la segmentación multibanda. Con dichas medias y con los parámetros $\lambda=8$ y $\beta=0.035$ se obtuvo el resultado que muestra la figura 4.15.



Fig. 4.15 Primera segmentación multibanda de la prueba 4.4.2_c

No se obtuvo una clasificación perfecta debido al nivel de ruido de la imagen. De hecho, la figura 4.15 es muy parecida a la imagen obtenida de hacer la segmentación multibanda con $\lambda=5$ y $\beta=0.035$ (ver figura 4.16).



Fig. 4.16 Segunda segmentación multibanda de la prueba 4.4.2_c

Considerando precisamente el ruido que contiene la imagen (la varianza), podríamos pensar en disminuir el parámetro β para ver si así se mejoran los resultados, la figura 4.17 muestra dos resultados, el de la izquierda se obtuvo de hacer la segmentación multibanda con los parámetros $\lambda=1$ y $\beta=0.015$; intentando disminuir la granularidad, se incrementó el parámetro λ a 5 y se obtuvo el resultado de la

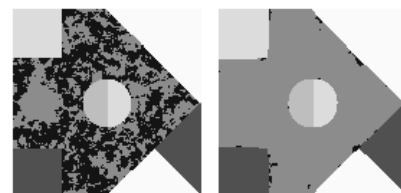


Fig. 4.17 Tercera y cuarta segmentación multibanda de la prueba 4.4.2_c

derecha, definitivamente desapareció la granularidad en la región 3 pero lamentablemente desapareció la región 5. Finalmente, para este caso, los mejores resultados son los que se ven en las figuras 4.15 y 4.16.

Conclusiones :

Al usar el parámetro λ con valores grandes para realizar las segmentaciones monobanda, se obtienen segmentaciones que no muestran el detalle de las cosas, identifican cosas grandes, las formas más representativas de las bandas. Así, después de generar la imagen que combina a dichas segmentaciones, es más probable que se elijan las clases más significativas de la imagen, lo que causa que ya no sea necesario realizar la segmentación multibanda con un valor tan grande para el parámetro λ (a menos que se tenga un nivel de ruido bastante alto).

Este método automático es muy robusto al ruido; en esta sección mostramos clasificaciones perfectas para imágenes con ruido de hasta $\sigma=3$. Para imágenes con niveles de ruido más altos se obtuvieron muy buenos resultados, como el que se ve en la figura 4.15 con $\sigma=5$, donde el porcentaje de pixeles mal clasificados fue poco más del 1%.

Esto nos dio confianza para realizar las pruebas con imágenes satelitales (se muestran en el siguiente capítulo), ya que es muy poco probable que un ND sea modificado en un rango mayor o igual a ± 20 .

Capítulo 5

Segmentación de imágenes del satélite LANDSAT TM

Un punto fundamental de este trabajo fue que un experto validara los resultados del método de MRMF para el caso multibanda, partiendo de modelos de firmas espectrales estimados. Hasta el momento se tiene la experiencia de realizar segmentaciones con el MRMF en imágenes de resonancia magnética del cerebro donde expertos en ese campo han determinado que los resultados obtenidos son muy buenos.

Para este caso, el experto que analizó los resultados fue la M.C. Aurora Breceda, investigadora del Centro de Investigaciones Biológicas del Noroeste (CIBNOR), quien ha realizado diversos estudios sobre vegetación en el estado de Baja California Sur y en su reciente trabajo de tesis doctoral identificó distintos tipos de vegetación en la reserva de la biósfera Sierra de La Laguna, en B.C.S.

En la sección 4.1 se promediaron las firmas espectrales de algunas categorías que la M.C. Aurora Breceda presenta en su documento de tesis doctoral [4], esto con el fin de contar con un modelo de firmas para generar la imagen sintética que se trató a lo largo de aquel capítulo. Ahora veremos algunos resultados de segmentar imágenes captadas en septiembre de 1996 por el satélite LANDSAT TM; el experto sugirió que los experimentos se realizaran bajo el supuesto de *no conocer nada*, es decir, no usar sus firmas espectrales como modelo de inicio, de tal manera que en cada prueba el sistema identificara simultáneamente un modelo de firmas y las regiones de la imagen.

Con la finalidad de realizar la segmentación completamente automática, se utilizaron en todos los casos modelos de firmas espectrales inicializados mediante la propuesta de segmentaciones monobanda. Veamos qué resultados se obtuvieron.

5.1 Segmentaciones de la zona de San José del Cabo

San José del Cabo es una ciudad pequeña que se encuentra en el municipio de Los Cabos, al sur del estado de Baja California Sur. En este apartado se muestran algunas segmentaciones que corresponden a imágenes del satélite LANDSAT TM, donde se captaron algunas zonas de la ciudad y particularmente el área del Estero de San José. En la figura 5.1 podemos ver las bandas de la imagen.

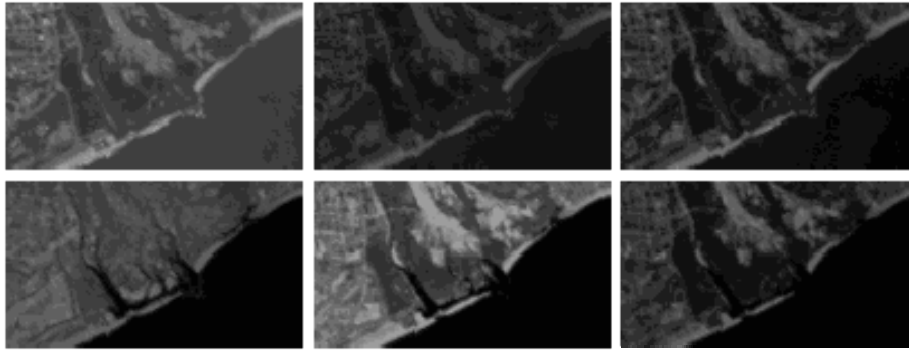


Fig 5.1 Zona de San José del Cabo, vista desde el satélite LANDSAT TM

Pruebas:

a) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: **8**

Después de generar la imagen de todas las regiones contenidas en las segmentaciones monobanda, se tomaron las medias de las **6** regiones con mayor soporte para iniciar la segmentación multibanda con los siguientes parámetros: $\beta=0.015$ para cada resultado de la figura 5.2 y $\lambda=0.5$ (arriba izquierda), $\lambda=1$ (arriba centro), $\lambda=1.5$ (arriba derecha), $\lambda=2$ (abajo izquierda), $\lambda=2.5$ (abajo centro) y $\lambda=5$ (abajo derecha).

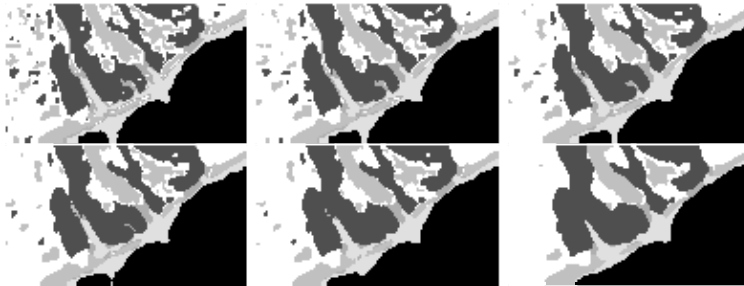
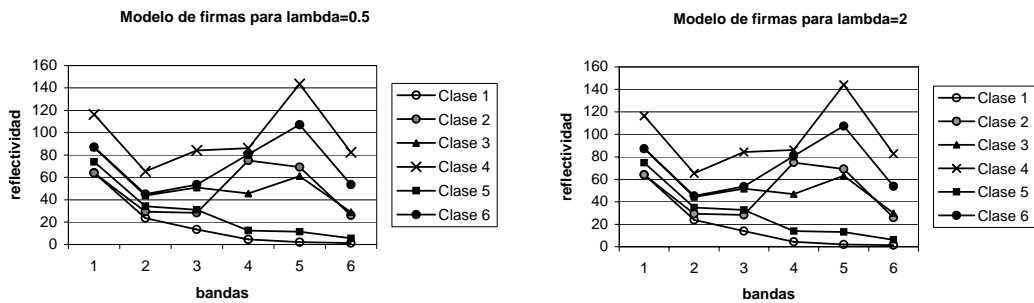


Figura 5.2 Resultados de la prueba 5.1_a

En cada clasificación aparecen 6 tonos de gris distintos, el más oscuro corresponde a la primera clase identificada (clase 1) y el tono más claro a la última (clase 6). De modo que siguiendo esta escala de grises es como vamos a diferenciar las clases de la imagen. Ahora veamos el modelo de firmas obtenido.

Para cada caso se llegó básicamente al mismo modelo, donde cada patrón espectral identificado es distinto a los otros.



Análisis del experto:

- La clase 1 corresponde al mar y está bien delimitada en todas las imágenes. Sólo hay una pequeña diferencia entre los experimentos, para λ menores que 5 se identificó una mezcla de agua marina con agua dulce como parte de la clase 5 (agua dulce) y para $\lambda=5$ la misma mezcla se clasificó como agua de mar.
- La clase 2 corresponde a vegetación. En ninguno de los resultados la clasificación es fina porque no se distinguen "tipos" de vegetación; sin embargo, para un análisis general los resultados son buenos. Pensando en realizar una reclasificación de esa zona, la imagen que permitiría hacerlo de una manera más eficiente es la que se segmentó con $\lambda = 0.5$.
- La clase 3 corresponde a una interfase muy difícil y es agua-suelo. Son áreas donde el espejo del agua es muy somero y hay mucho ruido en la respuesta espectral.
- La clase 4 corresponde a suelos y arena y se ve mejor en la imagen segmentada con $\lambda = 0.5$.
- La clase 5 es agua dulce y también se ve mejor clasificada en la imagen segmentada con $\lambda = 0.5$; donde se ve claramente una entrada de agua dulce en la vegetación. El problema es la mezcla de agua dulce con mar, porque esta imagen marca más una prolongación del agua dulce dentro del mar.
- La clase 6 es una mezcla de muchas cosas: cultivos, zonas urbanas y áreas de poca vegetación. Entre las imágenes se ve poca diferencia; sin embargo, en la imagen segmentada con $\lambda = 0.5$ al lado izquierdo se ve mayor fragmentación (son zonas urbanas).

Por lo que el experto comenta respecto a la clase 6, es evidente que la imagen contiene más clases, lo que nos llevaría a hacer una segmentación tratando de identificar más cosas; pero antes de hacerla pasemos a otra prueba identificando otra vez 6 cosas.

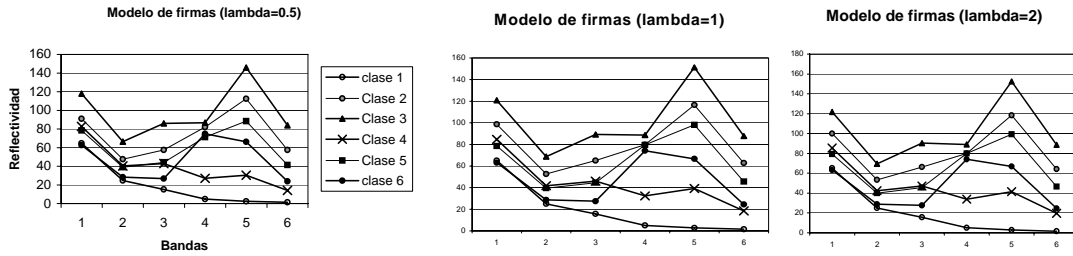
b) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: 6

Después de considerar la imagen de todas las regiones contenidas en estas segmentaciones monobanda, se tomaron las medias de las 6 regiones con mayor soporte para iniciar la segmentación multibanda con los siguientes parámetros: $\beta=0.015$ para los tres resultados que muestra la figura 5.3 y $\lambda=0.5$ (izquierda), $\lambda=1$ (centro) y $\lambda=2$ (derecha).



Figura 5.3 Resultados de la prueba 5.1_b

Ahora veamos el modelo de firmas obtenido en cada resultado.



Con $\lambda=0.5$ se obtuvieron 2 firmas parecidas (clases 2 y 5) que corresponden al patrón espectral identificado por la clase 6 de la prueba anterior. Al realizar la segmentación con $\lambda=1$ y $\lambda=2$ se modificaron un poco esas firmas pero no se desacoplaron, sino que se ajustaron de modo que el promedio de ellas difiere muy poco de la clase 6 de la prueba anterior (en la prueba con $\lambda=2$ lo más que difieren es 1.3 en la banda 1). Lo que explica la confusión entre estas dos clases en los resultados.

La siguiente figura resalta los pixeles etiquetados por las clases 5 (arriba) y 2 (abajo).



Figura 5.4 Pixeles etiquetados por las clases 5 y 2

Ahora veamos que no hay mucha diferencia entre la unión de pixeles de estas clases y los pixeles etiquetados por la clase 6 en la prueba anterior. La fig. 5.5 muestra arriba los pixeles de las clases 5 y 2 de esta prueba y en la parte de abajo resalta la clase 6 de la prueba anterior para los resultados con $\lambda=0.5$, $\lambda=1$ y $\lambda=2$ respectivamente.

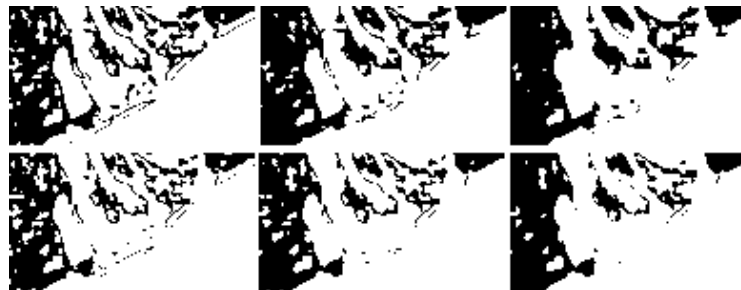


Figura 5.5 Comparación entre categorías

Regresando a la comparación entre firmas, encontramos que la clase 4 identificada en estos resultados corresponde al promedio de las clases 3 y 5 de la prueba anterior.

Análisis del experto:

Clase 1 = mar y agua dulce

Clase 2 = zonas urbanas, cultivos y áreas de poca vegetación

Clase 3 = suelo-arenas

Clase 4 = suelo-agua

Clase 5 = zonas urbanas, cultivos, vegetación, agua y suelo; esta clase es muy poco clara.

Clase 6 = vegetación

En estas imágenes se mezclan más el mar y el agua dulce; la clase 5 es menos clara que la 2, ya que identificó mezcla de agua-suelo, zonas urbanas, cultivos y vegetación.

Hasta este momento, la imagen que el experto prefirió fue la de la figura 5.2 con $\lambda=0.5$, aunque la parte urbana se distingue un poco mejor en la imagen de la figura 5.3 con $\lambda=2$. Entonces, con motivo de identificar más detalles, sugirió realizar la segmentación con una mayor cantidad de clases.

c) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: 12

Luego de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las 12 regiones con mayor soporte para iniciar la segmentación multibanda con los parámetros $\beta=0.015$ y $\lambda=0.15$.

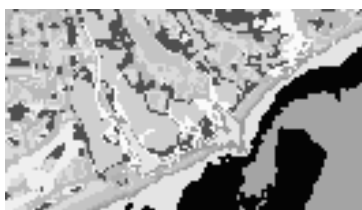
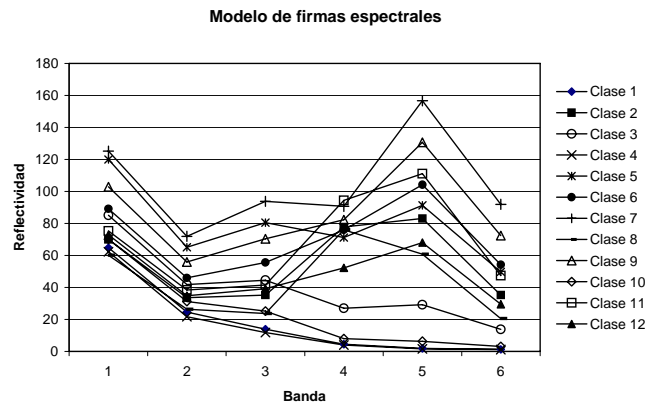


Figura 5.6 Resultado de prueba



Análisis del experto:

En general esta imagen es muy detallada, probablemente muestra demasiados detalles, por lo que hay que reclasificarla, agrupando sólo algunas firmas espectrales parecidas; pero en general es bastante buena. A continuación veremos lo que se identificó en esta imagen y algunos comentarios sobre las firmas.

CLASES:

- clase 1 = mar
- clase 2 = áreas de escasa vegetación
- clase 3 = agua-suelo
- clase 4 = mar
- clase 5 = agua-suelo
- clase 6 = suelos y vegetación
- clase 7 = suelos y arena
- clase 8 = vegetación (guamúchil y mezquitales)
- clase 9 = cultivos
- clase 10 = agua dulce y mezcla con mar
- clase 11 = trazos urbanos
- clase 12 = vegetación (palmares)

COMENTARIOS:

Aunque las firmas 3 y 5 son parte de la interfase agua-suelo, sus patrones espectrales son muy distintas, esto ha de deberse a la predominancia de alguno de los elementos (en un caso agua y en el otro tierra). Luego, las clases 1 y 4 sí son muy parecidas y se podrían agrupar; la clase 6 parece contener zonas mezcladas de suelo y poca vegetación, más o menos parecida a la clase 2. Y el resto de las clases se ven bastante bien, sobre todo las correspondientes a vegetación.

Haciendo la reclasificación, obtenemos la imagen de la siguiente figura.

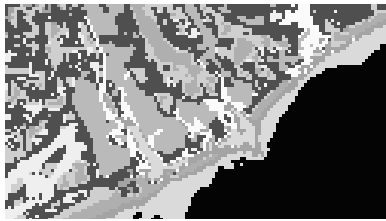


Figura 5.7 Imagen reclasificada 5.1_c

CLASES:

- clase 1 = mar
- clase 2 = suelos y áreas de escasa vegetación
- clase 3 = agua-suelo
- clase 4 = suelos y arena
- clase 5 = vegetación (guamúchil y mezquitales)
- clase 6 = cultivos
- clase 7 = agua dulce y mezcla con mar
- clase 8 = trazos urbanos
- clase 9 = vegetación (palmares)

De esta manera, la imagen ya no muestra categorías repetidas y como se había comentado, se distinguen bastante bien unas cosas de otras. Al experto le gustó mucho esta clasificación, sobre todo porque se identificaron algunos tipos de vegetación.

5.2 Segmentaciones en la reserva de la biósfera Sierra de La Laguna

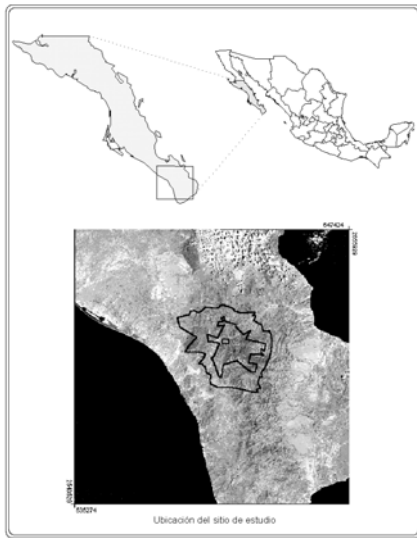


Figura 5.8 Ubicación de la reserva de la biósfera Sierra de La Laguna

La figura 5.8 muestra la ubicación de la reserva. La imagen del satélite que se muestra en la parte de abajo (en la misma figura) contiene 4496 renglones y 4486 columnas. Para trabajar el área de la reserva se hizo un recorte de la imagen de 2000 renglones y 2080 columnas. Como todavía es una imagen muy grande, se localizaron distintas zonas dentro de la reserva, de modo que cada zona contiene algunos puntos geográficos donde el experto realizó estudios de campo [4]. En esas zonas se hicieron pruebas y el experto evaluó los resultados. A continuación veremos algunos de ellos.

5.2.1 Segmentaciones en la zona de la vertiente del Golfo de California.

Para darnos una idea de la zona a clasificar, la figura 5.9 muestra las bandas 4 y 5 del satélite LANDSAT TM.

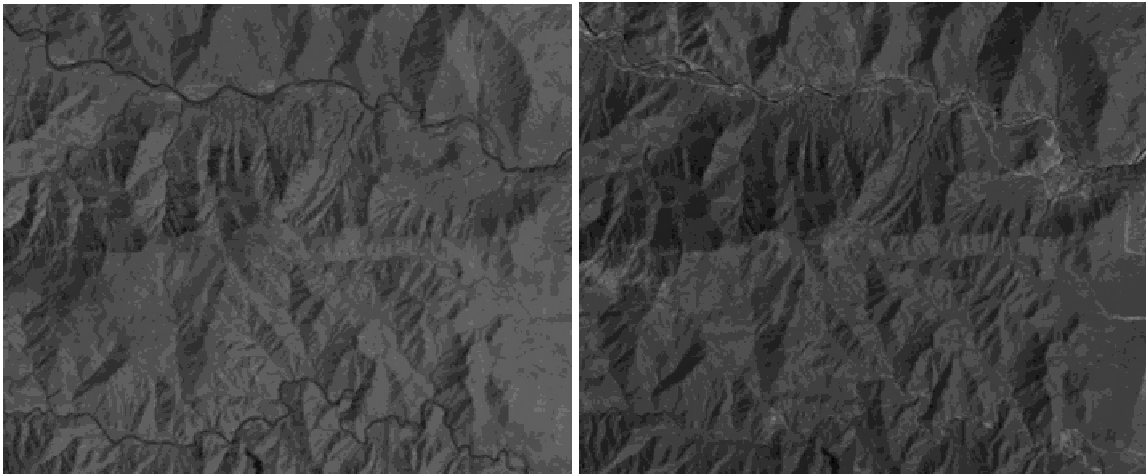


Figura 5.9 Zona de prueba en la vertiente del Golfo de California, vista desde las bandas 4 y 5 del sensor Thematic Mapper.

Pruebas:

- a) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: 8

Luego de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las **8** regiones con mayor soporte para iniciar la segmentación multibanda con los parámetros $\beta=0.015$ y $\lambda=0.1$.



El experto encontró que la mayoría de firmas espectrales obtenidas de esta segmentación coinciden con las que identificó en su trabajo de tesis doctoral e hizo una correspondencia entre las categorías de esta imagen y las suyas.

Figura 5.10 Resultado de la prueba 5.2.1_a

CLASES:

La clase 1 corresponde a *selva amatorialada*.

La clase 2 mezcla *selva* y *matorral*.

La clase 3 define perfectamente los *arroyos*.

Las clases 4 y 5 se parecen mucho, podrían agruparse dado que pertenecen a *selva cerrada*.

La clase 6 mezcla varias cosas, se parece a la clase 2 pero incluye al *encinar*.

La clase 7 corresponde a las *sombras*, define muy bien las respuestas más oscuras.

La clase 8 corresponde a la *selva abierta*.

La reclasificación de esta imagen consistiría sólo en fusionar las clases 4 y 5, por lo que resultaría una imagen de cuatro clases bien identificadas y dos confusas. Así pues, con el propósito de desagrupar las cosas mezcladas en las clases 2 y 6, se realizó la siguiente prueba con más clases en la segmentación multibanda.

- b) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: 8

Luego de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las **10** regiones con mayor soporte para iniciar la segmentación multibanda con los siguientes parámetros: $\beta=0.015$ y $\lambda=0.1$.

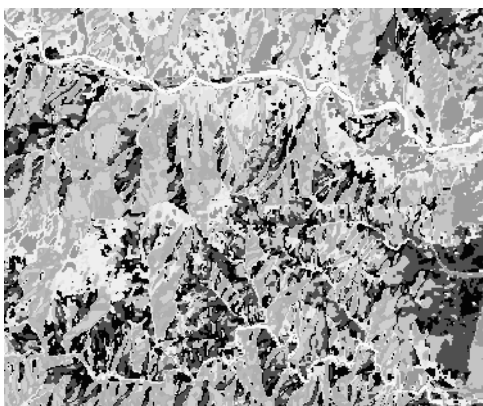


Figura 5.11 Resultado de la prueba 5.2.1_b

Cabe mencionar que el experto hizo una comparación detallada entre estos resultados y su mapa de vegetación⁵. Para obtener dicho mapa, realizó una clasificación supervisada de la imagen de 2000x2080 píxeles, utilizando el algoritmo de máxima verosimilitud y con esta clasificación, el índice de suelo TM5/TM4⁶ e información acerca de la altimetría del terreno y cubiertas de vegetación, generó entonces una reclasificación, permitiéndole diferenciar firmas parecidas. Y una vez hecha esa reclasificación, suavizó la imagen con un filtro de moda de 5x5 píxeles.

Análisis del experto:

- Las clases 1 y 2 tienen firmas espectrales muy parecidas y en ambos casos corresponden a la clase de *selva amatorialada*.
- La clase 3 está muy dispersa, es decir, hay muchos píxeles que en la figura 5.11 fueron identificados como pertenecientes a la clase 3 y en el mapa de vegetación del experto los mismos píxeles se consideraron de otras categorías. Sin embargo, corresponde principalmente a *áreas de escasa vegetación*. En la figura 5.12 se muestran los píxeles etiquetados como clase 3.
- Las clases 4 y 5 corresponden notablemente a *selva cerrada* y aunque algunos píxeles mal clasificados están muy dispersos entre las otras clases, no se considera un problema crítico, ya que no se está tomando en cuenta la altimetría del terreno; cosa que ayudaría a validar si un píxel pertenece verdaderamente a la clase que se le ha etiquetado dada la altitud a la que se encuentra y la altitud a la que generalmente se encuentra esa clase.
- Las clases 6, 8 y 9 tienen firmas espectrales muy parecidas y corresponden a *selva abierta*. Hay que agruparlas para hacer la reclasificación.
- La clase 7 corresponde a *sombras*.
- La clase 10 corresponde a arroyos.

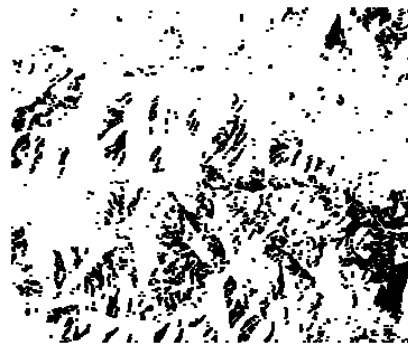


Figura 5.12 Píxeles etiquetados por la clase 3

⁵ El proceso que siguió para obtenerlo lo explica cuidadosamente en su tesis doctoral [4].

⁶ Tiene un significado similar al índice de vegetación (mencionado en la sección 2.4.1).

La reclasificación de esta imagen se muestra a continuación.

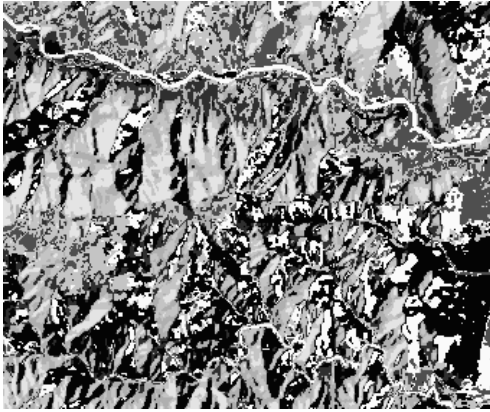


Figura 5.13 Imagen reclasificada 5.2.1_b

CLASES:

- clase 1 = selva amatorialada
- clase 2 = áreas de escasa vegetación
- clase 3 = selva cerrada
- clase 4 = selva abierta
- clase 5 = sombras
- clase 6 = arroyos

Con el propósito de detectar más detalles en la zona, se realizó la siguiente prueba .

c) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: 12

Después de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las 12 regiones con mayor soporte para iniciar la segmentación multibanda con los parámetros $\beta=0.015$ y $\lambda=0.1$. La figura 5.14 muestra el resultado obtenido.

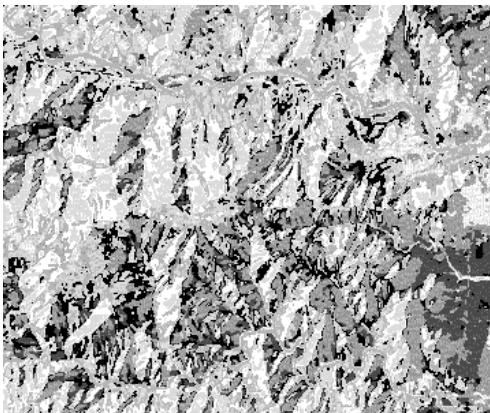


Figura 5.14 Resultado de la prueba 5.2.1 c

CLASES:

- Las clases 1, 2 y 3 son *selva amatorialada*.
- Las clases 4 y 8 son *selva abierta*.
- Las clases 5 y 11 corresponden a *áreas de escasa vegetación* y *suelos*; la clase 5 identifica principalmente arroyos muy delgados.
- La clase 6 son *arroyos*.
- Las clases 7, 10 y 12 corresponden a *selva cerrada*.
- La clase 9 corresponde a *sombras*.



Figura 5.15 Imagen reclasificada 5.2.1 c

Agrupando categorías se obtiene la reclasificación que muestra la fig. 5.15 con las siguientes clases:

- clase 1 = selva amatorialada
- clase 2 = áreas de escasa vegetación
- clase 3 = selva cerrada
- clase 4 = selva abierta
- clase 5 = sombras
- clase 6 = arroyos

Esta última imagen le gustó mucho al experto, a continuación veremos por separado cada clase identificada.

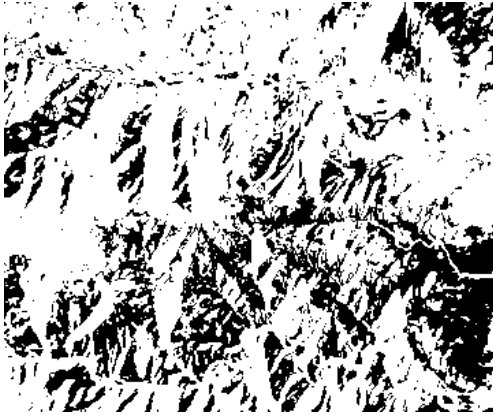


Figura 5.16 Selva amatorralada

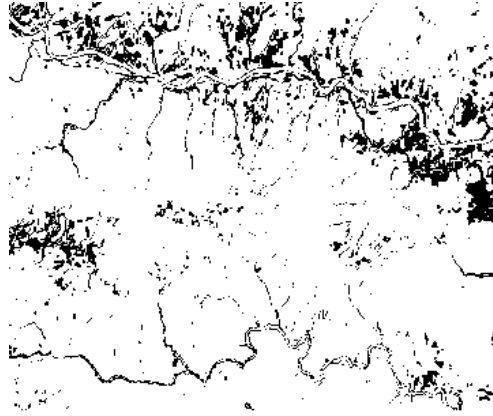


Fig. 5.17 Áreas de escasa vegetación y suelos

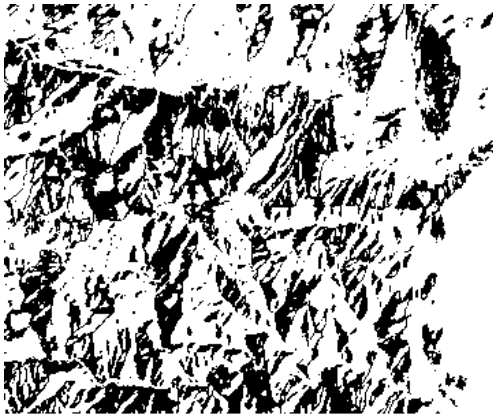


Figura 5.18 Selva cerrada

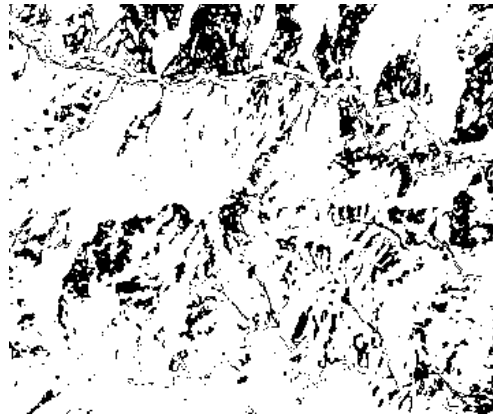


Figura 5.19 Selva abierta

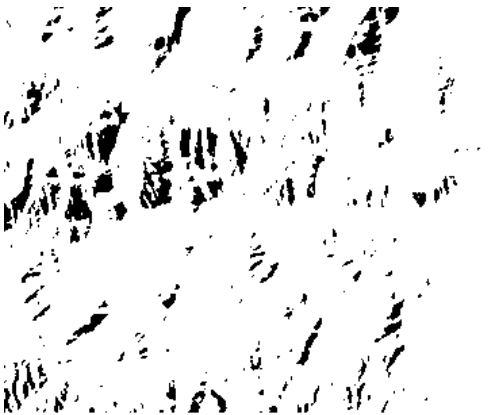


Figura 5.20 Sombras



Figura 5.21 Arroyos

5.2.2 Segmentaciones en la zona de la vertiente del Océano Pacífico.

Para darnos una idea de la zona a clasificar, la figura 5.22 muestra las bandas 4 y 5 del satélite LANDSAT TM.

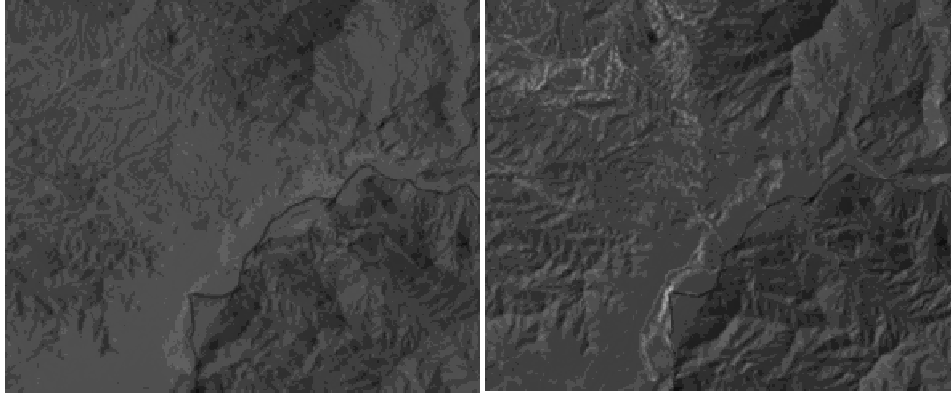


Figura 5.22 Zona de prueba en la vertiente del Océano Pacífico, vista desde las bandas 4 y 5 del sensor Thematic Mapper.

Pruebas:

- a) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: 10

Luego de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las 7 regiones con mayor soporte para iniciar la segmentación multibanda con los parámetros $\beta=0.015$ y $\lambda=0.1$.

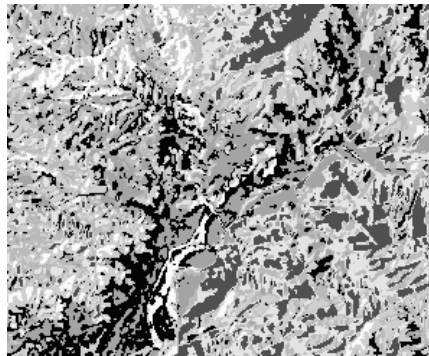


Figura 5.23 Resultado de la prueba 5.2.2_a

Análisis del experto:

- La categoría 1 mezcla dos tipos de vegetación muy distintas: *selva amatorialada* y *selva abierta*.
- La categoría 2 mezcla las clases *selva cerrada* y *sombras*.
- La clase 3 corresponde en su mayoría a *selva abierta*.

- La clase 4 muestra mucha confusión entre *selva abierta* y *selva cerrada*.
- La clase 5 presenta una equitativa mezcla de píxeles pertenecientes a *áreas de escasa vegetación* y *selva abierta*, lo que agradó al experto.
- La clase 6 corresponde en un 86% a *selva cerrada*.
- La categoría 7 identifica muy bien las *áreas de escasa vegetación*.

b) Parámetros de clasificación monobanda $\lambda = 25$, $\beta=0.035$, número de clases: 10

Luego de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las **10** regiones con mayor soporte para iniciar la segmentación multibanda con los parámetros $\beta=0.015$ y $\lambda=0.1$.

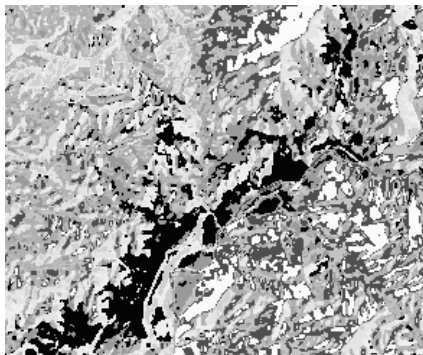


Figura 5.24 Resultado de la prueba 5.2.2_b

Análisis del experto:

- La categoría 1 presenta la misma confusión que en la imagen anterior entre: *selva amatorialada* y *selva abierta*.
- La categoría 2 corresponde en un 85.2% a la categoría que el experto definió como *selva cerrada*, un 3% de los píxeles de esta categoría se confundió con la clase *sombras* debido a la respuesta espectral de los mismos y el único aspecto preocupante es que un 9% de los píxeles se identificaron como de un tipo de vegetación muy abierta.
- La clase 3 corresponde en un 82% a *selva abierta*, un 17% de los píxeles de esta clase se identificaron como pertenecientes a *selva cerrada*.
- La clase 4 corresponde bastante bien a *selva abierta* (en cuanto al número de píxeles identificados en esta clase y a su patrón espectral); sin embargo, hay varios píxeles que en el mapa de vegetación del experto están catalogados como pertenecientes a una firma asociada a la clase *áreas de escasa vegetación*, la recomendación es hacer una depuración más fina de esta categoría.
- La clase 5 corresponde en un 87.4% a *selva abierta*.
- La clase 6 corresponde a la categoría *selva cerrada*.

- Las categorías 7 y 8 corresponden a *áreas de escasa vegetación* y se podrían unir para una reclasificación.
- La cat 9 está muy mezclada entre *selva abierta* y *áreas de escasa vegetación*, se recomienda hacer una diferenciación entre píxeles más fina, de modo que algunos se reclasifiquen entre las categorías definidas por 4 y 5 y otros en la categorías 7 y 8.
- La clase 10 corresponde a *sombras*.

La reclasificación de la imagen se muestra en la fig. 5.25

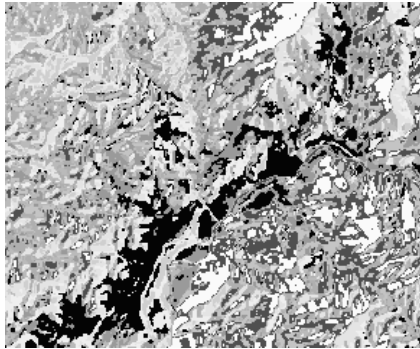


Figura 5.25 Imagen reclasificada 5.2.2_b

CLASES:

- clase 1 = mezcla de selva amatorralada y selva abierta
- clase 2 = selva cerrada
- clase 3 = selva abierta
- clase 4 = áreas de escasa vegetación
- clase 5 = mezcla selva abierta y áreas de escasa vegetación
- clase 6 = sombras

c) Parámetros de clasificación monobanda $\lambda=25$, $\beta=0.035$, número de clases: 12

Luego de generar la imagen con todas las regiones contenidas en estas segmentaciones, se tomaron las medias de las 12 regiones con mayor soporte para iniciar la segmentación multibanda con los parámetros $\beta=0.015$ y $\lambda=0.1$.

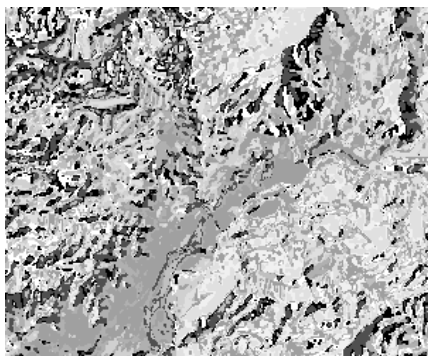


Figura 5.26 Resultado de prueba 5.2.2_c

CLASES:

- Las clases 1, 5, 8, 9 y 12 son *selva abierta*.
- Las clases 2 y 7 representan *áreas de escasa vegetación*.
- Las clases 3, 6 y 11 son *selva cerrada*.
- La clase 4 es *selva amatorralada*.
- La clase 6 son *arroyos*.
- Las clase 10 corresponde a *sombras*.

La figura 5.27 ilustra la imagen reclasificada como se indica a continuación y las figuras 5.28-5.32 muestran cada clase por separado.

- clase 1 = selva amatorralada.
- clase 2 = selva cerrada
- clase 3 = selva abierta
- clase 4 = áreas de escasa vegetación

clase 5 = sombras

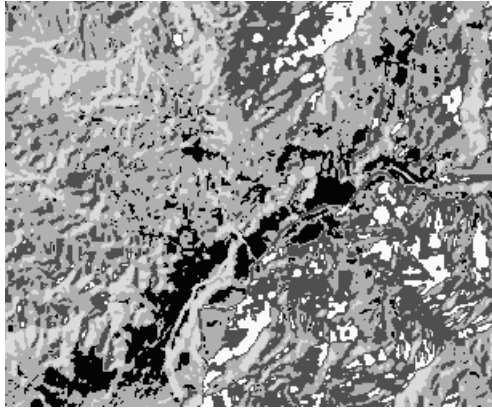


Figura 5.27 Imagen reclasificada

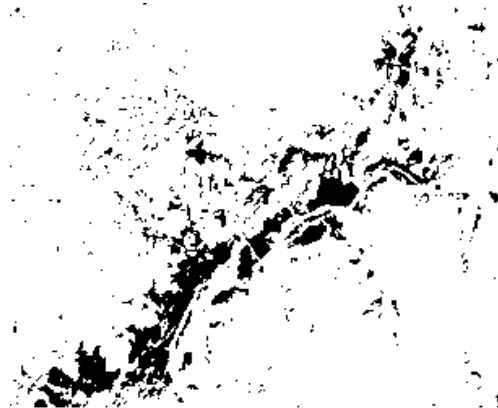


Figura 5.28 Selva amatorralada

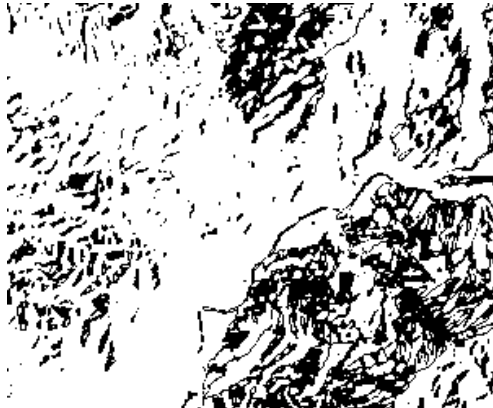


Figura 5.29 Selva cerrada



Figura 5.30 Selva abierta



Figura 5.31 Áreas de escasa vegetación

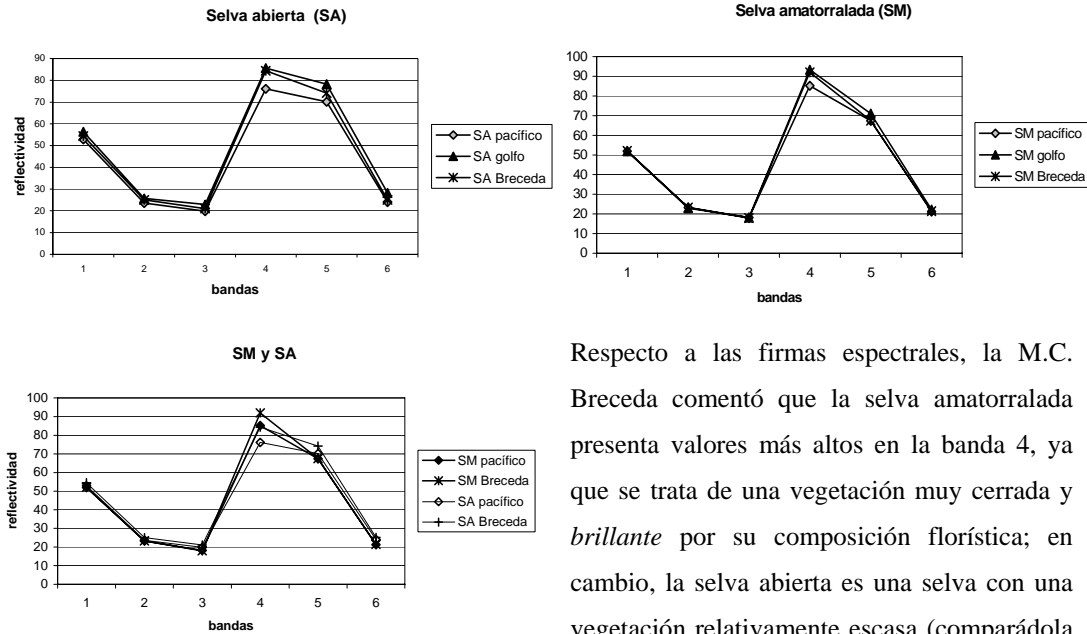


Figura 5.32 Sombras

5.3 Conclusiones y comentarios:

- De los análisis del experto y debido a que siempre prefirió aquellos resultados donde se pudieran identificar muchas clases (como 10 ó 12 clases), se concluye que es conveniente realizar la segmentación multibanda obligando al sistema a detectar más categorías de las que en realidad se desean identificar, efectuando posteriormente una reclasificación donde se agrupan patrones espectrales parecidos.
- Cuando la M.C. Breceda comparó los resultados contra su mapa de vegetación, encontró una consistente “confusión” entre las clases *selva amatorralada* y *selva abierta* en las pruebas efectuadas sobre la vertiente del Océano Pacífico. Comenta que esto puede deberse a dos cosas: primero, que una parte de la zona de prueba (en la fig. 5.22) no está dentro de la reserva e incluso contiene mucho matorral y segundo, que ella no trabajó mucho la selva amatorralada para obtener su mapa, lo que considera que puede ser una fuente de error.

Al observar las firmas de estas clases podemos darnos cuenta de que el MRMF identificó diferencias para cada clase entre vertientes y que las firmas que el experto utilizó se parecen mucho más a las que el MRMF encontró en la vertiente del Golfo de California. Lo que explica las confusiones a las que se refiere el experto.



Respecto a las firmas espectrales, la M.C. Breceda comentó que la selva amatorralada presenta valores más altos en la banda 4, ya que se trata de una vegetación muy cerrada y brillante por su composición florística; en cambio, la selva abierta es una selva con una vegetación relativamente escasa (comparádola con selva cerrada y/o selva amatorralada), por lo que su firma presenta valores menos marcados para la banda 4 y aunque no son tan notorias estas características en los patrones espectrales identificados por el MRMF en la vertiente del pacífico, concluyó que la figura 5.27 es una clasificación muy buena de esa zona.

- Finalmente, el experto comentó que las zonas de prueba dentro de la reserva de la biósfera Sierra de la Laguna son más complejas que el área de San José; debido a que buena parte de la reserva corresponde a vegetación, resulta más difícil hacer una segmentación con el propósito de distinguir tipos de vegetación que realizar la segmentación con el propósito de identificar cosas más distintas como mar, agua dulce, zonas urbanas, vegetación, etc. Por lo que dijo sorprenderse aún más de los resultados.

Conclusiones generales y trabajo futuro

Conclusiones generales:

- Los resultados que se obtienen con el método de MRMF definitivamente son muy buenos. Incluso cuando los niveles de ruido no son pequeños es posible llegar a la clasificación óptima de la imagen seleccionando los parámetros adecuados. Y aunque resulta difícil decir cuáles son los parámetros apropiados para cada caso (empezando por calibrar los parámetros del método de Descenso Newtoniano), por la experiencia de realizar más pruebas de las que se presentan en este documento, fue posible indicar cómo hacerlo.
En el apéndice C⁷ se menciona cómo *definir* y *ajustar* cada uno de los parámetros involucrados en este proceso de segmentación.
- El método de MRMF es robusto al modelo de inicialización de firmas espectrales. En el capítulo 4 pudimos confirmarlo con imágenes sintéticas y en el capítulo 5 el experto encontró que muchas de las firmas espectrales obtenidas en estos experimentos, coincidieron con las que había identificado en su trabajo de tesis doctoral. Ahora bien, las pruebas que se muestran en el capítulo 5 se realizaron inicializando el modelo de firmas mediante segmentaciones monobanda, básicamente por la comodidad de usar un método completamente automático (además de que las segmentaciones monobanda se pueden reutilizar).
- Finalmente, cabe mencionar que en el contexto de percepción remota, la *clasificación* se considera un arte. Hay que tomar en cuenta diversos factores para poder decir qué es lo que se encuentra en una imagen satelital (o aérea). Y para decidir qué factores se deben considerar en el proceso, se requiere de que quien realiza este trabajo tenga suficiente conocimiento acerca del terreno en estudio, por lo que es importante recurrir a investigaciones de campo.

Trabajo futuro:

- Es notorio que la debilidad del método son los parámetros, se deben calibrar los de descenso newtoniano antes de hacer pruebas y para cada prueba hay que ajustar λ y β .
Y aunque en el apéndice C ya se definieron algunos factores a considerar para ajustarlos, definitivamente ahorraría tiempo si la calibración de los parámetros del método de descenso fuera automática y se pudiera mantener fijo alguno de los otros dos parámetros para que el usuario sólo tuviera que ajustar uno.

⁷ El apéndice C es el manual de usuario del software desarrollado para hacer segmentaciones con MRMF.

- En los experimentos realizados con imágenes de satélite fue necesaria la ayuda del experto para hacer una reclasificación. Si se tomara en cuenta que algunas firmas son parecidas y partiendo del supuesto de que firmas parecidas corresponden a una misma clase, podría mejorarse el sistema incluyendo una técnica sofisticada para agrupar dichas firmas y así automatizar la reclasificación. Ahora bien, en el caso de implementar alguna propuesta, ésta también tendrá que contemplar que en ocasiones varias clases tienen la misma respuesta espectral aunque se traten de cosas distintas.
- Con la finalidad de depurar y validar resultados, los expertos en teledetección recomiendan hacer una *reclasificación*, considerando factores externos al proceso anterior. Por ejemplo, la M.C. Breceda consideró la altimetría del terreno para distinguir entre dos clases que parecían ser una sola, debido a que tuvieron la misma respuesta espectral en el proceso automático. También consideró el índice de suelo TM5/TM4, esto para verificar las clases que presentan menor cobertura vegetal en el área de estudio.

En ocasiones podría considerarse este tipo de información en el algoritmo de segmentación, permitiendo obtener una clasificación completamente automática y confiable. Por ejemplo, podríamos tomar el mapa de altimetría del terreno como si fuera una banda más de la imagen, de modo que el sistema haga distinción de clases cuando éstas tengan la misma respuesta espectral pero se encuentren a distintas alturas. Aunque sin duda, el problema de esta propuesta se presentará cuando una clase pueda localizarse a cualquier altura y el sistema la particione en tantas clases como rangos de altimetría se tengan en el terreno; por lo que si la minoría de las clases se distinguen de otras sólo por esta característica, sería conveniente no incluir el mapa como otra banda y hacer la distinción de clases en la reclasificación posterior.

Ahora bien, en el caso de que el experto conociera perfectamente a qué altura es más probable que se encuentre cada tipo de vegetación identificado en la imagen tratada, pudiera modelarse este conocimiento en forma de distribución de probabilidad e incluirlo como información a priori al proceso de segmentación.

Apéndice A

Modelo Oculto de Campo de Medidas Aleatorio Markoviano⁸ para Segmentación de Imágenes [34]

A.1 Modelos clásicos de Campos Aleatorios Markovianos

Los modelos probabilísticos que muchas veces se usan para formular problemas de segmentación se ajustan a la descripción general de la figura A.1.

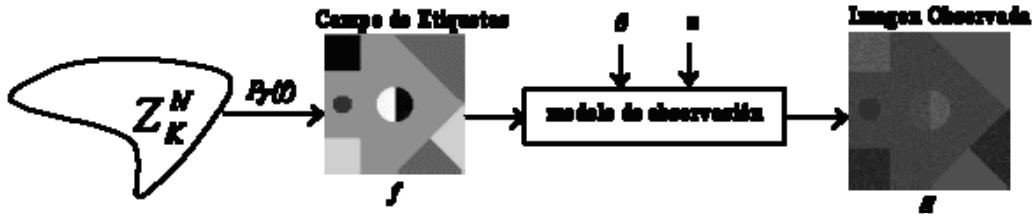


Figura A.1 Modelo clásico de Campos Aleatorios Markovianos para segmentación de imágenes

Para entenderla, introduzcamos la siguiente notación: sea L la retícula de píxeles (o voxeles, si es un caso tridimensional) donde se observa la imagen g . El modelo asume que existen K regiones, $\{R_1, \dots, R_K\}$, tales que $L = \bigcup_{k=1}^K R_k$; $R_i \cap R_j = \emptyset$, $i \neq j$, de modo que el valor observado en el píxel $r \in L$ está dado por:

$$g(r) = \sum_{k=1}^K \Phi(r, \theta_k) b_k(r) + n(r) \quad (1)$$

donde $n(r)$ es un campo de ruido blanco con distribución conocida P_n (esto es, $\{n(r), r \in L\}$ son variables aleatorias Gaussianas con media 0 y distribución estándar σ , independientes e idénticamente distribuidas); $\Phi(\cdot, \cdot)$ es un modelo paramétrico; θ_k es el vector de parámetros que corresponde a la región R_k , y $b_k(r)$ es la función indicadora correspondiente: $b_k(r) = 1 \Leftrightarrow r \in R_k$; nótese que $b_k(r)$ satisface las restricciones:

$$\sum_{k=1}^K b_k(r) = 1, b_k(r) \in \{0, 1\} \forall r \in L \quad (2)$$

Existe un campo de etiquetas f asociado a b , con $f(r) \in Z_K = \{1, \dots, K\}$, el cual indica a qué región pertenece el píxel r , esto es, $b_k(r) = \delta(f(r) - k)$. En este modelo, se asume que el campo f es una muestra de un CAM, es

⁸ Campo de Medidas Aleatorio Markoviano, conocido como MRMF del inglés (Markov Random Measure Field).

decir, una muestra del espacio de estados Z_K^N (donde N es la cardinalidad de L), obtenida con una

distribución de Gibbs:

$$P_f(f) = \frac{1}{Z_f} e^{-\sum_C V_C(f)} \quad (3)$$

donde Z_f es una constante de normalización, la sumatoria en el exponente corre sobre todos los cliques de L para un sistema de vecindad dado, y $\{V_C(f)\}$ son "funciones de energía potencial", cada una de las cuales depende exclusivamente del valor de f en los sitios que pertenecen al clique C (ver [16,17] para detalles). Estas funciones de energía potencial junto con el sistema de vecindad elegido, controlan la apariencia del campo de muestra f , y por ende las propiedades de la segmentación estimada. Una función de energía potencial muy usada es la generalización del modelo de Ising, la cual considera cliques de grado 2 (pares de sitios que están separados menos de 2 unidades) y funciones de energía potencial de la forma:

$$\begin{aligned} V_C(f_i, f_j) &= -\lambda, \text{ si } f_i = f_j \\ &= \lambda, \text{ en otro caso} \end{aligned}$$

donde λ es un parámetro que controla la granularidad del campo. Como el campo f no es directamente observable, suele llamársele modelo de CAM oculto.

A.1.1 Algoritmos de Estimación

El problema de segmentación consiste en encontrar estimadores óptimos tanto para el campo f como para el vector de parámetros $\theta = (\theta_1, \dots, \theta_K)$, dadas las observaciones g . Para esto, usando la teoría de estimación Bayesiana encontramos los siguientes pasos [37]:

1. Encontrar la verosimilitud de las observaciones $P(g|f, \theta)$.
2. Con la regla de Bayes, encontrar la distribución a posteriori $P(f, \theta | g)$ usando la distribución a priori $P_f(f)$ (y $P_\theta(\theta)$ si también se conoce).
3. Definir una función de costo apropiada $C(\hat{f}, \hat{\theta}, f, \theta)$ que asocie el costo de estimar $\hat{f}, \hat{\theta}$, dado que los verdaderos valores son f, θ .
4. Encontrar los estimadores óptimos f^*, θ^* minimizando la función

$$Q(\hat{f}, \hat{\theta}) = E [C(\hat{f}, \hat{\theta}, f, \theta) | g] \quad (4)$$

Veamos estos pasos con detalle.

La verosimilitud de las observaciones se obtiene del modelo de observación (1) y de la distribución de ruido P_n (asumiendo que la conocemos):

$$P(g | f, \theta) = \prod_{r \in L} \nu_{f(r)}(r) \quad (5)$$

donde cada vector de K entradas $\nu(r)$ se define como:

$$\nu_k(r) = P(g(r) | f(r) = k, \theta) = P_n(g(r) - \Phi(r, \theta_k)) \quad (6)$$

Por ejemplo, para el caso de ruido gaussiano tenemos:

$$v_k(r) = \sqrt{\frac{\beta}{\pi}} \exp(-\beta |g(r) - \Phi(r, \theta_k)|^2) \quad (7)$$

donde β es un parámetro que depende de la varianza del ruido.

Usando (7), (3) y la regla de Bayes, encontramos la distribución a posteriori como:

$$P(f, \theta | g) = \frac{1}{Z_p} e^{-U(f, \theta)}$$

donde Z_p es una constante de normalización, y

$$U(f, p) = -\sum_{r \in L} \log v_{f(r)}(r) + \sum_C V_C(f) - \log P_\theta(\theta) \quad (8)$$

donde se usará una distribución a priori P_θ no informativa (una constante) en caso de no existir restricciones a priori sobre θ .

La minimización de Q (ecuación (4)) generalmente se realiza en 2 pasos iterativos, a los algoritmos que lo hacen así se les llaman algoritmos de Segmentación-Estimación del Modelo (SM), éstos, en el paso S buscan la mejor segmentación de la imagen, dados los estimadores actuales del modelo de parámetros θ , y en el paso M buscan la mejor estimación de θ , dada la segmentación actualmente estimada. Es bien sabido que al encontrar una segmentación "suave" o "difusa" en el paso S, el algoritmo SM se hace más robusto respecto a la estimación inicial de θ , que siempre hay que definir. Por esto, es más conveniente trabajar con un campo de vectores indicadores b con las restricciones: $\sum_{k=1}^K b_k(r) = 1$ y $b_k(r) \geq 0, \forall r \in L$, en lugar del campo f con las "estrictas" restricciones (2).

Esta es la forma general de los algoritmos SM:

1. Encontrar un estimador inicial $\bar{\theta}$ para θ ;
2. Repetir hasta tener convergencia:

$$\text{Paso S: actualiza } \bar{b} = \arg \min_b Q(\hat{b}, \bar{\theta})$$

$$\text{Paso M: actualiza } \bar{\theta} = \arg \min_\theta Q(\bar{b}, \hat{\theta})$$

Las funciones de costo que normalmente se usan (aunque en muchos casos no se definan explícitamente) son separables, expresadas en términos de b son de la forma:

$$C(\hat{b}, \hat{\theta}, b, \theta) = C_b(\hat{b}, b) + C_\theta(\hat{\theta}, \theta)$$

Dependiendo de C_b y C_θ , se obtienen distintos algoritmos SM. Por ejemplo, siendo

$$C_\theta(\hat{\theta}, \theta) = 0, \text{ si } \hat{\theta} = \theta \\ = 1, \text{ en otro caso}$$

$$\text{y } C_b(\hat{b}, b) = \sum_{r \in L} |\hat{b}(r) - b(r)|^2$$

tenemos el algoritmo EM [12], (aunque no es la forma usual de derivar el método) en este caso tenemos que:

$$Q_b(\hat{b}) = E[C_b(\hat{b}, b) | g] = \sum_{r \in L} \sum_{k=1}^K E[(\hat{b}_k(r) - b_k(r))^2 | g]$$

y con $\frac{\partial Q_b}{\partial \hat{b}_k(r)} = 0$ obtenemos que \bar{b} en el paso S está dada por $\bar{b}_k(r) = E[b_k(r) | g] = P(f(r) = k | g)$, es

decir, \bar{b} corresponde a la distribución de probabilidad marginal a posteriori para f . En el paso M, $\bar{\theta}$ es simplemente el estimador de Máximo a Posteriori (MAP) dado $b = \bar{b}$, que se obtiene de minimizar la ecuación (8) expresada en términos de b , es decir, minimizando con respecto a θ la función

$$U(\bar{b}, \theta) = - \sum_{r \in L} \sum_{k=1}^K \bar{b}_k(r) \log v_k(r) - \log P_\theta(\theta)$$

Otras opciones para C_b (con la misma elección para C_θ) dan otros algoritmos SM; por ejemplo, usando

$$C_b(\hat{b}, b) = \sum_{r \in L} |1 - \delta(\hat{b}(r) - b(r))|$$

donde $\delta(x)$ es 1 si y sólo si x es el vector cero (esto quiere decir que C_b indica el número de píxeles mal clasificados), en el paso S uno obtiene el estimador máxima marginal a posteriori (MPM Maximizer of the Posterior Marginals) [37]. Usando

$$\begin{aligned} C_b(\hat{b}, b) &= 1, \text{ si } \hat{b} = b \\ &= 0, \text{ en otro caso} \end{aligned}$$

en el paso S obtenemos para b el estimador MAP [45][52], etc.

El problema con esta clase de algoritmos es que no es posible minimizar con precisión $Q(\hat{b}, \bar{\theta})$ con respecto a \hat{b} en el paso S, es decir, ni las marginales a posteriori ni los estimadores óptimos MPM o MAP se pueden calcular con exactitud debido a que los cálculos involucran una sumatoria con K^N términos (para las marginales a posteriori) o la solución de un problema de optimización combinatoria con $N=|L|$ variables (para el estimador MAP). Por lo que debemos recurrir a aproximaciones; las más precisas están basadas en métodos estocásticos, como los algoritmos de Monte Carlo con Cadenas de Markov (MCMC) [16,40], lo malo es que computacionalmente son muy caros; las aproximaciones rápidas (como el algoritmo ICM⁹, [2]) son muy sensibles al ruido. Las aproximaciones basadas en la teoría de campo medio (Mean Field) [50,51,15,49] son más rápidas que MCMC, pero todavía caras y además sensibles al ruido (ver capítulo 3). En [36] se presenta un algoritmo reciente para el estimador MPM, basado en una aproximación Gaussiana para las marginales a posteriori, es rápido y resistente al ruido; sin embargo, este proceso MPM-MAP es muy sensible al estimador inicial para θ .

⁹ Iterated Conditional Models

A.2 Modelos de Campos de Medidas Markovianos Ocultos

La dificultad mencionada con anterioridad puede resolverse si usamos un modelo probabilístico distinto para generar el campo de etiquetas f . En lugar del procedimiento de un paso descrito en la Fig. A.1, la propuesta es usar un modelo probabilístico de 2 pasos, con un campo adicional oculto llamado p . Este modelo se muestra en la Fig. A.2. En el primer paso se genera el campo aleatorio markoviano de vectores, donde cada vector $p(r)$ toma valores dentro del simplejo de K vértices S_K :

$$S_K = \{u \in \mathfrak{R}^K : \sum_{k=1}^K u_k = 1, u_k \geq 0, k = 1, \dots, K\} \quad (9)$$

Por lo que $p(r)$ se interpreta como una medida de probabilidad discreta sobre Z_K (el espacio de estados de etiquetas). En el segundo paso se genera el campo de etiquetas f , de modo que cada $f(r)$ es una muestra independiente de la distribución $p(r)$, así que

$$P(f | p) = \prod_{r \in L} p_{f(r)}(r) \quad (10)$$

Nótese que la distribución a priori para f es:

$$P_f(f) = \int_{S_K^N} P(f | p) dP(p)$$

lo que no es una distribución Gibbsiana (ver apéndice B); por lo tanto, aunque p sea Markoviano, f de ninguna manera es un CAM, así que esta clase de modelos es diferente de los clásicos.

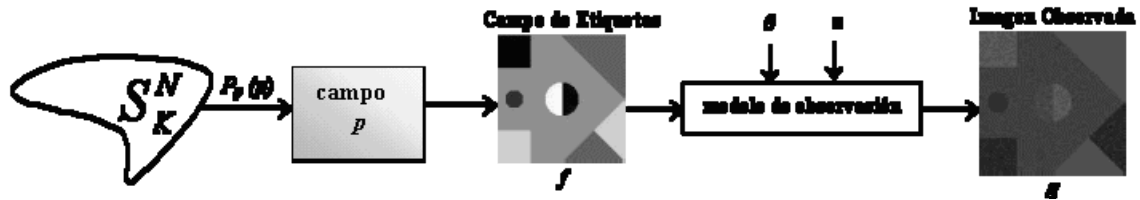


Figura A.2 Campo Aleatorio Markoviano Oculto para segmentación de imágenes

Sin embargo, como en los modelos clásicos, las funciones potenciales (en este caso para el campo p) se usan para cumplir las restricciones del campo de etiquetas. Por ejemplo, se tiene coherencia espacial en las regiones $\{R_1, \dots, R_K\}$ cuando se requiere que cada vector $p(r)$ sea similar a sus vecinos. Un potencial cuadrático que expresa esta condición es:

$$V_{\langle r, s \rangle}(p(r), p(s)) = \lambda |p(r) - p(s)|^2 = \lambda \sum_{k=1}^K (p_k(r) - p_k(s))^2 \quad (11)$$

donde λ es un parámetro positivo, y $\langle r, s \rangle$ son pixeles vecinos en L . Se pueden definir otros potenciales para cumplir restricciones más complejas, pero en este caso nos concentraremos en la (11).

La distribución a posteriori $P(p, \theta | g)$ se obtiene de la regla de Bayes como:

$$P(p, \theta | g) = \frac{1}{Z} P(g | p, \theta) P_p(p) P_\theta(\theta) \quad (12)$$

donde Z es una constante de normalización. La distribución condicional $P(g/p, \theta)$ se obtiene de la siguiente forma:

$$P(g | p, \theta) = \prod_{r \in L} P(g(r) | p, \theta)$$

La distribución condicional $P(g(r)/p, \theta)$ puede obtenerse de primero calcular la distribución condicional conjunta $P(g(r), f(r) | p, \theta) = P(g(r) | f(r), p, \theta) P(f(r) | p, \theta)$, y luego marginalizar sobre $f(r)$:

$$P(g(r) | p, \theta) = \sum_{k=1}^K P(g(r) | f(r) = k, p, \theta) P(f(r) = k | p, \theta)$$

Considerando que $P(g(r) | f(r), p, \theta) = P(g(r) | f(r), \theta)$ y que $P(f(r) = k | p, \theta) = p_k(r)$, obtenemos:

$$P(g(r) | p, \theta) = \sum_{k=1}^K v_k(r) p_k(r) = v(r) \cdot p(r) \quad (13)$$

donde $v_k(r)$ está dada por la ecuación (7).

Luego, de las ecuaciones (13) y (12) finalmente obtenemos que:

$$P(p, \theta | g) = \frac{1}{Z} e^{-U(p, \theta)} \quad (14)$$

con

$$U(p, \theta) = -\sum_{r \in L} \log(v(r) \cdot p(r)) + \sum_C V_c(p) - \log P_\theta(\theta) \quad (15)$$

donde consideramos cliques de tamaño 2 y potenciales dados por la ecuación (11). Para obtener el estimador óptimo f^* (para el campo de etiquetas), usamos el siguiente procedimiento de 2 pasos:

1. Encontrar los estimadores MAP p^* y θ^* para p, θ :

$$p^*, \theta^* = \arg \max_{p \in S_K^N, \theta} P(p, \theta | g) \quad (16)$$

2. Encontrar f^* como el máximo de $P(f | p = p^*, \theta^*, g)$

El primer paso es equivalente a minimizar $U(p, \theta)$, dada en la ecuación (15), sujeta a las siguientes restricciones:

$$p(r) \in S_K, \forall r \in L \quad (17)$$

con S_K definida en (9), mientras el segundo paso simplemente consiste en hallar la moda de cada medida discreta $p^*(r)$ en forma desacoplada:

$$f^*(r) = \arg \max_k p_k^*(r) \quad (18)$$

La carga computacional, en este caso, recae en el primer paso; sin embargo, como la función (15) es diferenciable, se puede minimizar muy eficientemente, como se muestra en el siguiente apartado.

A.3 Algoritmo de minimización de la función de energía

Puede suceder que no obtengamos un resultado muy bueno al momento de minimizar la función (15) si es que usamos técnicas de optimización de propósito general; pero si consideramos la simplicidad de las restricciones (17) y la estructura del campo p debido a la Markovianidad, es posible obtener buenísimos resultados usando un método de descenso Newtoniano con la proyección del gradiente (GPND¹⁰) [33,29]. El método se basa en la idea de que para cada iteración hay que moverse en la dirección de d , de tal manera que $\nabla U \cdot d < 0$ (esta es la dirección de descenso) y que la nueva posición esté dentro de una región factible. Esto se logra cuando se elige d como la proyección del gradiente negativo sobre el subespacio tangente, definido por el sistema de restricciones activas (ver [29], pp 331-339). Luego, la convergencia se puede acelerar si consideramos a cada elemento $p_k(r)$ (o $\theta_j(r)$) como la posición de una partícula de masa unitaria, sujeta a una fuerza igual a $-\partial U/\partial p_k(r)$ (o bien $-\partial U/\partial \theta_j$). Las ecuaciones de movimiento para estas partículas se obtienen de la segunda ley de Newton:

$$\begin{aligned}\ddot{\theta} &= -\nabla_{\theta} U - 2\alpha\dot{\theta} \\ \ddot{p} &= -\nabla_p U - 2\alpha\dot{p}\end{aligned}$$

donde α es el coeficiente de fricción. La discretización de estas ecuaciones nos lleva a un algoritmo de descenso de gradiente con inercia; para satisfacer las restricciones (17), cada nueva posición de la partícula $p_k(r)$ debe ser proyectada de nuevo al simplejo S_k , para obtener la iteración completa como:

$$\begin{aligned}\theta^{(t+h)} &= \frac{2}{\alpha h + 1} \theta^{(t)} + \frac{\alpha h - 1}{\alpha h + 1} \theta^{(t-h)} - \frac{h^2}{\alpha h + 1} \nabla_{\theta} U(p^{(t)}, \theta^{(t)}) \\ \tilde{p} &= \frac{2}{\alpha h + 1} p^{(t)} + \frac{\alpha h - 1}{\alpha h + 1} p^{(t-h)} - \frac{h^2}{\alpha h + 1} \nabla_p U(p^{(t)}, \theta^{(t)}) \\ p^{(t+h)}(r) &= \prod_{S_k}(\tilde{p}(r)), \text{ para toda } r \in L.\end{aligned}\tag{19}$$

donde h es el tamaño del paso que damos en cada iteración y el operador $\prod_{S_k}(u)$ encuentra el punto más aproximado al vector $u \in \mathfrak{R}^K$ dentro de S_k .

Para encontrar esta proyección, consideramos las siguientes observaciones: un simplejo $S_{K,A}$, con K vértices, definido como :

$$S_{M,A} = \{u \in \mathfrak{R}^K : \sum_{k=1}^K u_k = 1; u_k \geq 0, k \in A; u_k = 0, k \notin A\}$$

donde $A \subseteq Z_K$ es un conjunto de índices y $|A| = K$, está contenido en el hiperplano:

$$H_A = \{u \in \mathfrak{R}^K : \sum_{j \in A} u_j = 1 \text{ y } u_j = 0 \text{ si } j \notin A\}$$

¹⁰ Gradient Projection Newtonian Descent

La proyección ortogonal $x = \prod_A(u)$ de un punto $u \in \mathcal{H}^K$ sobre el hiperplano H_A satisface $\sum_{j \in A} (u_j - x_j) = 0$; $x_j = 0$, $j \notin A$, y entonces la obtenemos de la fórmula:

$$x_k = u_k - \frac{\sum_{i \in A} u_i - 1}{K}, \text{ si } k \in A \quad (20)$$

Ahora, si x es el punto en $S_{K,A}$ más cercano a \tilde{p} , y x se encuentra en el interior de $S_{K,A}$, entonces x deberá ser igual a $\prod_A(\tilde{p})$. Por otro lado, si x no se encuentra en el interior de $S_{K,A}$, entonces debió caer en el simplejo $S_{K,A'}$, con $|A'| < |A|$ sobre la frontera de $S_{K,A}$. Este simplejo activo corresponde precisamente a $A' = \{k \in A : x_k \geq 0\}$, donde $x = \prod_{A'}(\tilde{p})$. Esta observación sugiere que el punto en S_K más cercano a un punto dado \tilde{p} debiera ser encontrado proyectando recursivamente \tilde{p} en H_A , y luego actualizando A en caso de ser necesario, así es que esto se refiere a un simplejo activo en la frontera de S_K . Esto se hace de una manera muy simple, eliminando de A aquellos índices que corresponden a componentes negativos de x . Veamos el siguiente algoritmo para encontrar $x = \prod_{S_K}(\tilde{p}(r))$:

1. Sea $x = \tilde{p}(r)$ y $A = Z_K$
2. Mientras $x \notin S_K$ hacer:
 - a. $x = \prod_A(\tilde{p})$ usando (20)
 - b. $A = \{k : x_k \geq 0\}$

Nótese que este algoritmo convergerá en un máximo de K iteraciones.

La minimización de la función (15) puede acelerarse más si se usa una propuesta multiescala: podemos tomar características de la imagen observada g de escalas cada vez más gruesas de ella $\{g = g_0, \dots, g_M\}$ mediante suavizado y submuestreo recursivo de la misma (pirámide Gaussiana estándar [5]). En cada escala m , se obtiene un campo de verosimilitudes $v^{(m)}$ correspondiente, se reemplaza g por g_m en (7) y se minimiza la función de energía $U_m(p^{(m)}, \theta^{(m)})$. En la escala M , el procedimiento concluirá eficientemente debido al reducido número de variables. Una vez que se encuentran los mínimos $p^{*(m)}$ y $\theta^{*(m)}$, se pasan como valores de inicio a la escala $m-1$, hasta que se realiza el proceso en la escala 0 (con la imagen original). Debe tenerse mucho cuidado al pasar la solución de $p^{*(m)}$ a una escala más fina, ya que el proceso de interpolación debe garantizar que el campo p interpolado también esté dentro de S_K^N . Si consideramos métodos de interpolación de la forma:

$$p^{(k-1)}(r) = \sum_{s \in N^m(r)} w_{rs} p^{*m}(s)$$

donde $N^m(r)$ denota el conjunto de sitios en la retícula del nivel m , en la que se basa la interpolación del sitio r en la rejilla $m-1$, entonces se cumplirá la restricción $p^{(m-1)} \in S_K^N$ si $p^{(m)} \in S_K^N$ y $\sum_{s \in N^m(r)} w_{rs} = 1$, para toda r . Este es el caso de la interpolación bilineal (o trilineal, en el caso de tener datos en 3-D); nótese

que las variables θ no se interpolan, sólo se transmiten de nivel a nivel. Finalmente, el parámetro λ que controla la fuerza de la interacción espacial entre pixeles vecinos y por tanto, la granularidad de la solución, debe ajustarse; como la distancia entre pixeles se duplica cuando vamos de una escala fina a otra más gruesa, el parámetro λ correspondiente deberá dividirse entre 2, así que $\lambda_M = 2^{-M} \lambda_0$. Este proceso de optimización multiescala no es usado en los resultados reportados en esta tesis pero es interesante probarlo cuando se hace segmentación de imágenes monobanda.

Apéndice B

En este apéndice se demuestra que f no es un CAM, aunque p sí lo sea (ver figura A.2).

Para que f sea un CAM, deberíamos tener que $P(f(r)|f(s), s \neq r) = P(f(r)|f(s), s \in N_r)$, donde N_r es la vecindad del pixel r . Ahora,

$$P(f(r) | f(s), s \neq r) = \frac{\int_{S_M^N} P(f(r), f(s), s \neq r | p) dP(p)}{\int_{S_M^N} P(f(s), s \neq r | p) dP(p)} = \frac{\int_{S_M^N} \prod_{s \in L} p_{f(s)}(s) dP(p)}{\int_{S_M^N} \prod_{s \in L - \{r\}} p_{f(s)}(s) dP(p)}$$

Es posible factorizar $P(p)$ en 2 términos para que esta expresión dependa únicamente de $f(r)$, $f(s)$, $s \in N_r$: $P(p) = P_1(p(s), s \in N_r \cap \{r\}) \cdot P_2(p(s), s \notin N_r \cap \{r\})$ de modo que se puedan cancelar algunos términos, pero esto implicaría que $p(s)$, $s \in N_r \cap \{r\}$ es independiente $p(s)$, $s \notin N_r \cap \{r\}$, lo que contradice al modelo Gibbsiano propuesto para p .

Apéndice C

Manual de usuario del software MRMF

En este apartado podrá ver de una manera breve, cómo utilizar el software del MRMF para realizar segmentaciones de imágenes multibanda y/o visualizar resultados almacenados o las propias bandas de una imagen.

C.1 Pantalla principal

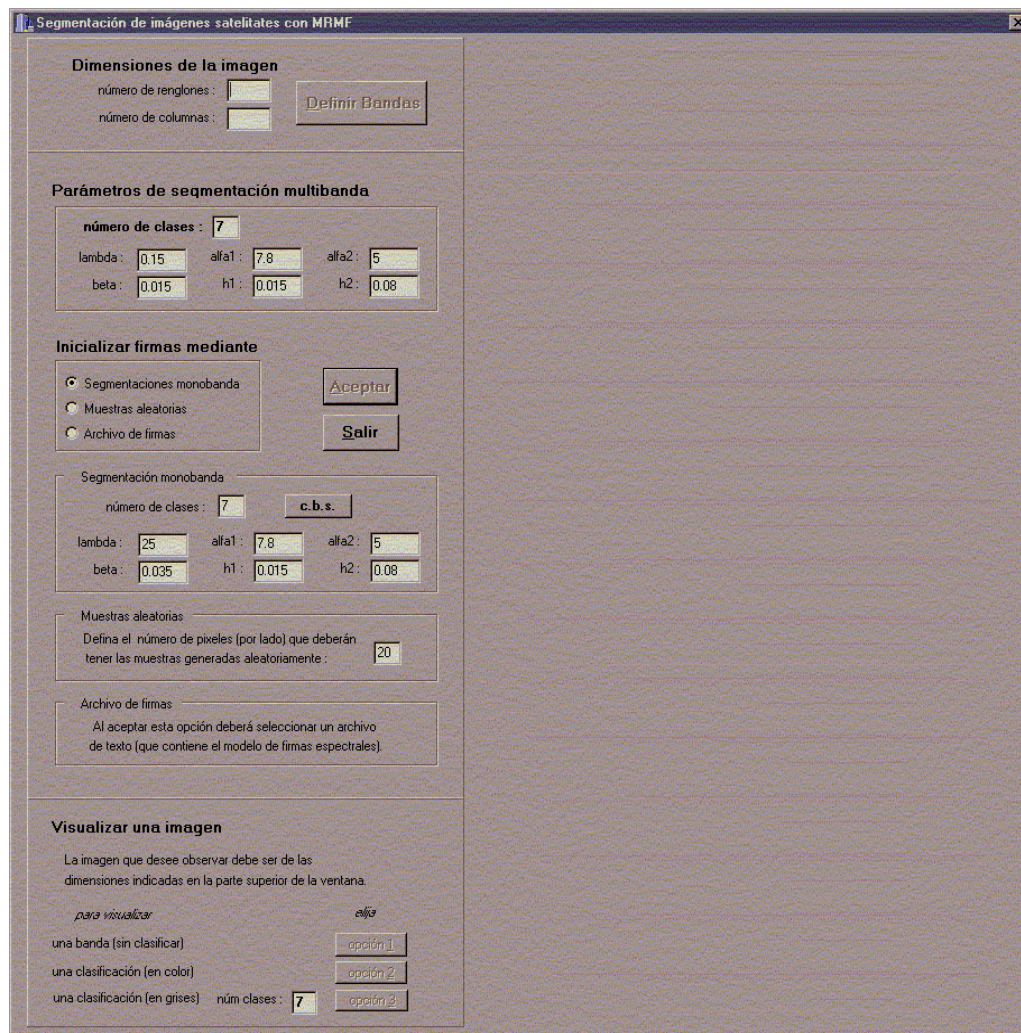


Figura C.1 Pantalla principal

La figura C.1 muestra la pantalla que aparece al momento de ejecutar el programa. En esta pantalla se encuentran todas las tareas que realiza el software. Podrá notar que parecen muchas secciones deshabilitadas, esto es porque todas las opciones requieren de otros datos para poder llevarse a cabo. Veamos qué pasos hay que seguir para realizar cada tipo de tarea.

C.2 Dimensiones de la imagen

El primer paso consiste en determinar las dimensiones de la imagen que se va trabajar¹¹. Al momento de introducir estos datos, se habilitan cuatro botones. El primero de ellos es para cargar las bandas de la imagen que se desea segmentar (ver figura C.2).



Figura C.2 Dimensiones de la imagen

Los otros 3 botones corresponden a la opción de visualizar una imagen. Al hacer click en alguno de estos botones, aparece la clásica ventanita del sistema operativo Windows que pregunta por el archivo que desea abrir. Donde usted sólo tiene que seleccionarlo y observar la imagen en la parte derecha de la pantalla.

En la figura C.3 vemos que esta sección es bastante clara en la descripción de cada botón, pues le indica que si desea visualizar alguna banda de una imagen deberá seleccionar *opción 1*, si desea observar alguna clasificación a colores deberá hacer click en el botón *opción 2* o si prefiere ver la clasificación en escala de grises deberá elegir *opción 3*. Para lograr una mejor distinción entre niveles de gris, se requiere que antes de presionar el botón correspondiente indique el número de clases que se indentifican en la imagen.

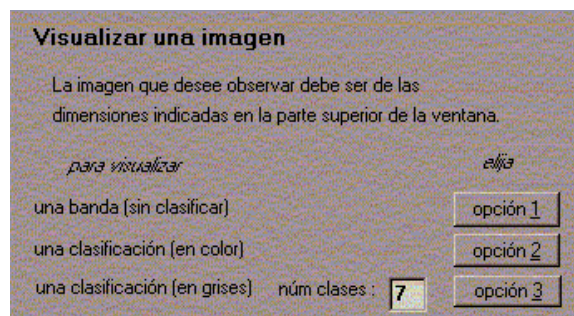


Figura C.3 Visualizar imagen

¹¹ Sugerimos que tanto el número de renglones como el número de columnas sean menores a 400 pixeles

C.3 Lectura de bandas

Éste es el segundo paso para realizar la segmentación multibanda y consiste en cargar las bandas de la imagen. Al hacer click en el botón *Definir Banda* (Fig. C.2) aparece un mensaje que indica el número de banda que está por elegir (Fig. C.4), seguido de la ventana que pregunta el nombre del archivo y así, continúa esta iteración de ventanas hasta que se cargan las 6 bandas de la imagen.

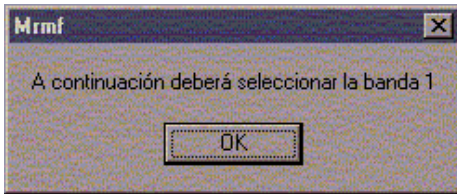


Fig. C.4 Mensaje para definir bandas

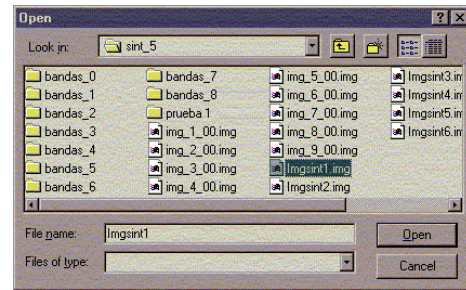


Fig. C.5 Abrir archivo

Con esto, ya es posible definir los parámetros de la segmentación, por lo que se habilitan las opciones centrales de la pantalla y el botón *Aceptar*, que indica que también se puede realizar la segmentación.

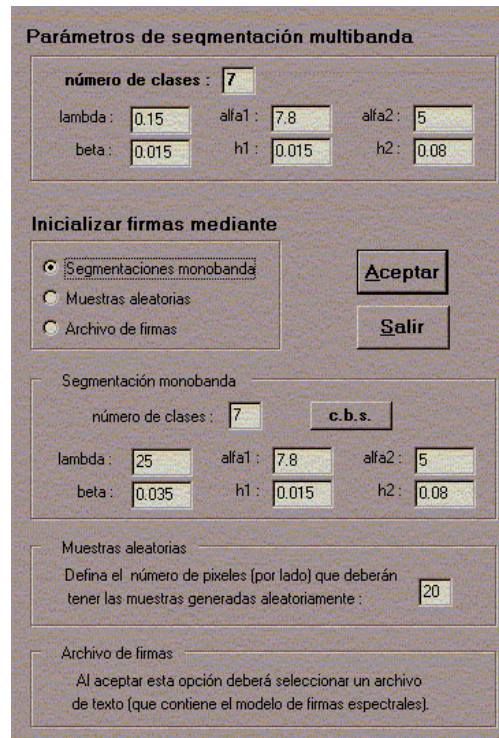


Figura C.6 Parámetros de segmentación

C.4 Parámetros de segmentación multibanda

Aunque el sistema defina por default ciertos valores para los parámetros de segmentación, se pueden modificar todos. De hecho, es primordial que usted defina el número de clases que desea identificar en la segmentación, cuya cantidad se considera prudente siendo menor o igual a 15;

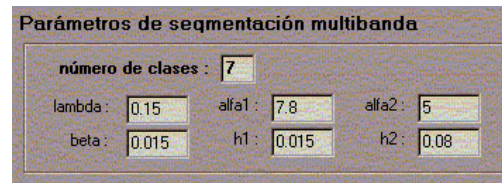


Figura C.7 Parámetros de segmentación multibanda

El parámetro *lambda* corresponde a la coherencia espacial, indica qué tanto deben parecerse los pixeles vecinos, cuanto más grande sea este valor menor granularidad tendrá la imagen clasificada, aunque también se perderán algunos detalles. Para imágenes con poco ruido se obtienen buenos resultados manteniendo este parámetro con un valor menor o igual 1, pero para imágenes con altos niveles de ruido (de $\sigma \geq 5$) conviene subirlo hasta 8 aproximadamente (ver las pruebas 4.2b, 4.4.1b y 4.4.2a).

Respecto al parámetro *beta* debe considerarse la varianza de la imagen. Si se ven muy ruidosas las bandas, deberá disminuir un poco el valor de este parámetro para obtener mejores resultados. En casos donde se tienen altos niveles de ruido a veces sólo es necesario aumentar el valor de *lambda* para eliminar la granularidad manera adecuada y otras veces también se requiere de disminuir este parámetro (ver las pruebas 4.2 b, 4.4.2a y 4.4.2c).

Los parámetros *h1* y *h2* indican el tamaño de paso del método de Descenso Newtoniano para el *modelo de firmas* y el *campo de medidas* respectivamente y los parámetros *alfa1* y *alfa2* indican el coeficiente de fricción del método de Descenso Newtoniano para el *modelo de firmas* y el *campo de medidas* respectivamente.

Se sugiere que estos parámetros se modifiquen sólo en aquellos casos donde no se obtenga clasificación perfecta al tratar una imagen *sin ruido*, y bien, en esos casos habrá que hacerlo de la siguiente manera:

- Si en el archivo de firmas espectrales se encuentran valores muy distintos (mucho mayores) a los que debieron obtenerse, tendrá que aumentar (poco) el tamaño de paso *h1* y observar si se mejoran un poco (fijando *lambda* en un valor menor o igual a 1 y $\beta \approx 0.035$); en caso de que no sea así, habrá que aumentar más el valor de *h1* y/o, decrementar el coeficiente de fricción *alfa1*, y así, una vez que se obtienen resultados parecidos a los que se esperan, habrá que hacer modificaciones muy pequeñas a

estos dos parámetros, chequeando al mismo tiempo cuáles valores de $h2$ y $alfa2$ son apropiados para la clasificación.

- Los parámetros $h2$ y $alfa2$ deberán modificarse de la misma manera que $h1$ y $alfa1$, principalmente cuando se visualiza el resultado con mucha granularidad sin que pueda disminuirse al modificar los parámetros lambda y beta.

C.5 Inicialización del modelo de firmas espectrales

Esta sección de la pantalla permite elegir la forma de inicialización de firmas espectrales para posteriormente realizar la segmentación multibanda. Contiene las propuestas presentadas en la sección 3.3 de esta tesis y una tercera opción que corresponde a utilizar un modelo de firmas almacenado en archivo.

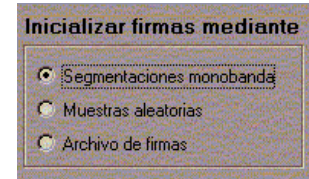


Figura C.8 Opciones de inicializar el modelo de firmas

C.5.1 Modelo de firmas inicializado mediante segmentaciones monobanda

Esta es la opción que el sistema sugiere de inicio, cuyos parámetros predeterminados se muestran en la figura C.9.

Es indispensable que defina el número de clases a identificar en estas segmentaciones y se sugiere que el parámetro lambda no sea mayor a 30 (esto es porque las segmentaciones monobanda identifican las estructuras más representativas de la imagen y si este parámetro es muy grande, prodría suceder que se pierdan), mantener beta en 0.035 y que los demás parámetros sean iguales a los que se definan en la segmentación multibanda.

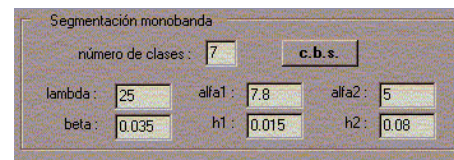


Figura C.9 Segmentación monobanda

Definidos los parámetros se puede iniciar la segmentación, entonces, el siguiente paso consiste en presionar

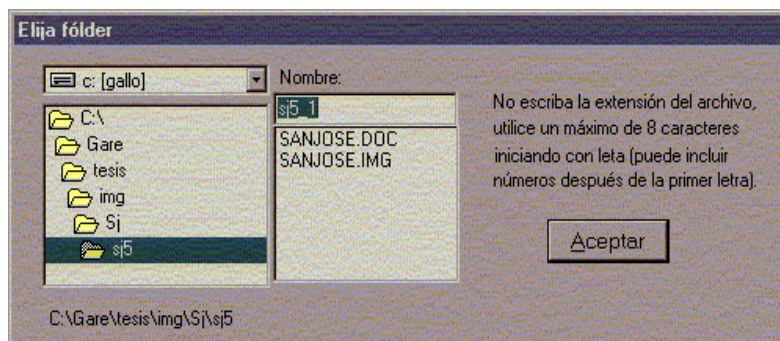


Figura C.10 Selecci6n de ruta de almacenamiento

el bot6n *Aceptar*; en ese momento aparecer6 la ventana que muestra la figura C.10, en la que debe elegir el f6lder y nombre de archivo (sin indicar extensi6n) donde desea almacenar el resultado de la segmentaci6n.

Por ejemplo, con los datos de la figura C.10 se van a generar los archivos: sj5_1_b1.img, sj5_1_b2.img, sj5_1_b3.img, sj5_1_b4.img, sj5_1_b5.img, sj5_1_b6.img, sj5_1.img, sj5_1.txt y sj5_1.doc en la ruta "c:\gare\tesis\img\sj\sj5\".

Los seis primeros archivos corresponden a las bandas segmentadas, el archivo sj5_1.img es la clasificación multibanda, sj5_1.txt contiene los parámetros utilizados en la segmentación y luego el modelo de firmas espectral encontrado (cada renglón corresponde a una clase y cada columna a una banda) y sj5_1.doc es el archivo de cabecera de la imagen clasificada (sin indicar la posición geográfica de la imagen). Las figuras C.11 y C.12 muestran el formato de los últimos dos archivos:

```

Parámetros de clasificación monobanda : lambda=1.000000, beta=0.035000
Parámetros de clasificación multibanda: lambda=25.000000, beta=0.035000

51.999996      29.000002      16.999996      98.999992      66.999992      20.003299
56.999985      27.000010      26.999990      72.999969      91.999962      36.000023
54.000004      25.000002      21.000004      83.999992      74.000000      24.999998
47.000000      18.000004      15.000005      63.999992      53.000000      16.000002
81.000000      45.999992      61.999996      48.999992      17.999998      8.999999
59.999855      25.999981      25.999866      49.000172      51.000000      20.999502
34.000004      15.000002      21.000004      83.977001      64.000005      14.999918

```

Figura C.11 Formato de archivo del modelo de firmas

```

File title : Clasificación
data type : byte
file type : binary
columns : 128
rows : 128
ref. system : plane
ref. units : m
unit dist. : 1
min. x :
max. x :
min. y :
max. y :
pos'n error : unknown
resolution : unknown
min. value : 0
max. value : 7
value units : unspecified
value error : unknown
flag value : none
flag def'n : none
legend cats : 8
category 0 :
category 1 : cat1
category 2 : cat2
category 3 : cat3
category 4 : cat4
category 5 : cat5
category 6 : cat6
category 7 : cat7

```

Figura C.12 Formato de archivo de cabecera

En la figura C.11 se observa el modelo de firmas en 7 renglones y 6 columnas, lo que indica que se realizó la segmentación multibanda considerando 7 clases. El número 14.999918 posicionado en el séptimo renglón y sexta columna indica la media de la clase 7 en la banda 6.

La figura C.12 muestra el archivo de cabecera de la segmentación, este archivo se genera con el propósito de que otros programas puedan leer la clasificación del MRMF. Lo único que no contiene el archivo y que hay que añadir, son las coordenadas geográficas de la imagen (lo puede hacer en cualquier editor de texto).

Cuando se realizan diversas pruebas sobre la misma imagen, inicializando el modelo de firmas mediante segmentaciones monobanda, resulta conveniente utilizar las segmentaciones monobanda generadas con anterioridad (por ejemplo: sj5_1_b1.img, ..., sj5_1_b6.img) y partir de ellas para llevar a cabo la

segmentación multibanda. Para elegir esta opción debe presionar el botón *c.b.s.*¹² de la figura C.9 y después de cargar las bandas segmentadas podrá iniciarse la segmentación presionando el botón *Aceptar* (que nuevamente va a preguntarle la ruta y nombre del archivo donde desee guardar la clasificación).

C.5.2 Modelo de firmas inicializado mediante muestras aleatorias

Al momento de elegir esta opción en la sección de la pantalla que ilustra la figura C.8, se habilita la parte que ilustra la figura C.13, donde hay que determinar las dimensiones de

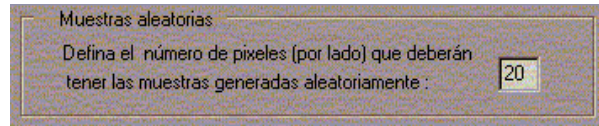


Figura C.13 Número de muestras aleatorias

la ventana cuadrada que capturará cada muestra. El sistema asigna un valor predeterminado en 20 píxeles por lado, pero las muestras de 400 píxeles resultan más confiables para imágenes de 128x128 que para imágenes de 300x300, por ejemplo. Así es que debe considerar las dimensiones de la imagen para definir el tamaño de las muestras y una vez definido este parámetro deberá seleccionar *Aceptar*, donde definirá ruta y nombre de archivo donde se almacenará el resultado para que posteriormente elija las estadísticas de las muestras que considere convenientes para la inicialización de firmas espectrales (ver la fig. C.14).

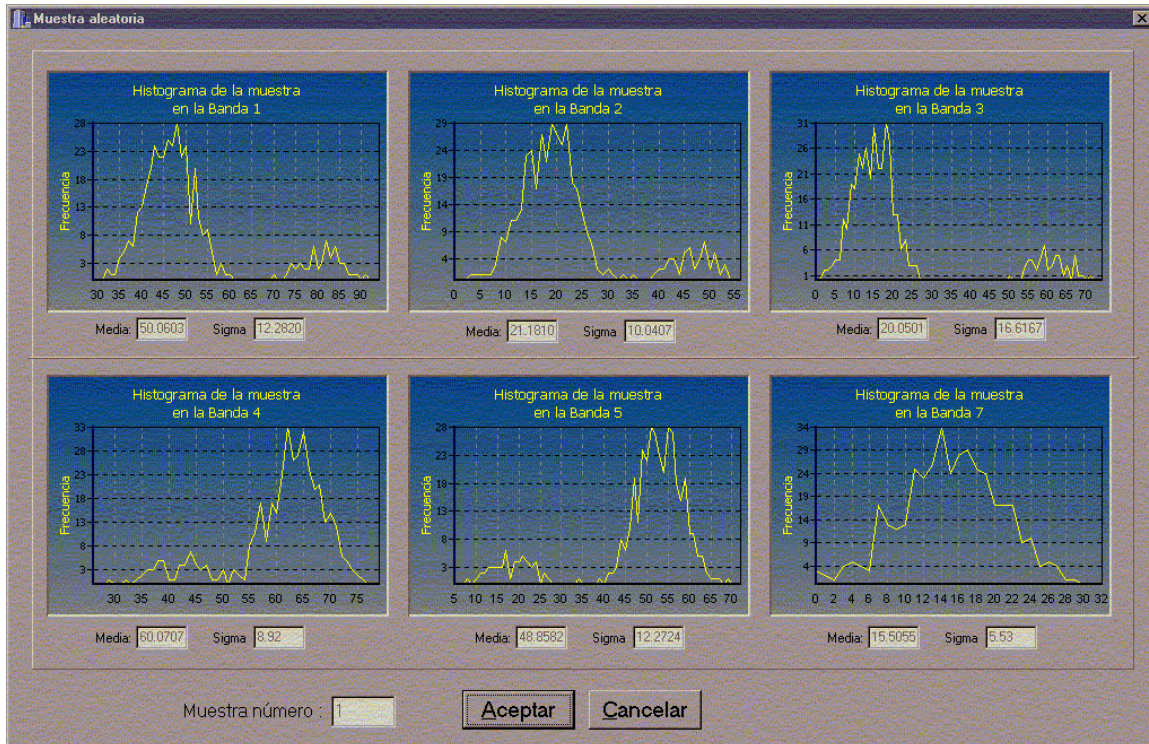


Figura C.14 Presentación de estadísticas de muestras aleatorias

¹² significa *cargar bandas segmentadas*

C.5.3 Modelo de firmas inicializado mediante archivo

Después de elegir esta opción en la sección de la pantalla que ilustra la figura C.8 podrá oprimir el botón *Aceptar*. En esta ocasión, antes de que el sistema le pregunte por la ruta y nombre de archivo donde se almacenarán los resultados, deberá elegir el archivo (con extensión txt) que contenga el modelo de firmas con el que iniciará la segmentación multibanda. Cabe mencionar que dicho archivo debe tener el formato mostrado en la figura C.11, donde no aparecerá el primer renglón (Parámetros de clasificación monobanda...) si es que ese modelo se obtuvo de segmentar la imagen inicializando el modelo de firmas mediante muestras aleatorias u otro archivo de firmas.

Antes de salir del sistema podrá realizar la cantidad de segmentaciones que desee (de distintas imágenes) y/o visualizar clasificaciones o bandas de imágenes, siempre y cuando el número de renglones y columnas de la imagen indiquen las dimensiones correctas.

C.6 Salir del sistema

Para salir del sistema sólo debe presionar el botón *Salir* o elegir cerrar la ventana desde la esquina superior derecha de la misma.

Bibliografía

- [1] L. Arriaga, and J. L. León de la Luz; *The Mexican tropical deciduous forest of Baja California Sur: a floristic and structural approach*, 1989 *Vegetation* 84: 45-52.
- [2] J. Besag; *On the statistical analysis of dirty pictures*. *Journal of the Royal Statistical Society B*, 48 (3):259-302, 1974.
- [3] A. Breceda; *La Selva Baja Caducifolia y la Vegetación de Fondo de Cañada en la Sierra de la Laguna, B.C.S.*, 1994 México. Tesis de Maestría. Fac. de Ciencias. UNAM. 95 pp.
- [4] A. Breceda; *El mosaico de vegetación en una selva baja caducifolia*. Tesis de doctorado en proceso. Fac. de Ciencias. UNAM.
- [5] P.J. Burt; *The pyramid as a structure for efficient computation*, In *Multiresolution image processing and analysis*, edited by A. Rosenfeld, Springer Series in information Sciences, Vol. 12 Springer, New York, 1984.
- [6] Ed Campbell; *Matching Remote Sensing Technologies and Their Applications*, London 1981, Remote Sensing Society, pp. 143-162.
- [7] J.B. Campbell; *Introduction to Remote Sensing*, New York 1987, The Guilford Press.
- [8] Emilio Chuvieco; *Fundamentos de teledetección espacial*, 1995.
- [9] R.G. Congalton; *Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data*, "*Photogrammetric Engineering and Remote Sensing*", vol. 54, 1988, pp. 587-592.
- [10] P.J. Curran; *Principles of Remote Sensing*, London 1985, Longman.
- [11] J.L. Cushnie; *The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies*, *International Journal of Remote Sensing*, vol. 8, 1987, pp. 15-29.
- [12] A.P. Dempster, N.M. Larid and D.B. Bubin; *Maximum likelihood form incomplete data via EM algorithm*. *Journal of the Royal Statistical Society, Series B* 39:1-38, 1977.
- [13] R.D.Duda and P.E. Hart; *Pattern Classification and Scene Analysis*, New York 1973, John Wilwy & Sons.
- [14] D.M. Gates, H.J. Keegan, J.C. Schleter and V.R. Weidner; *Spectral properties of plants*, *Applied optics*, vol. 4, 1965, pp. 11-20.
- [15] D. Geiger and F. Girosi; *Parallel and deterministic algoritms from MRFs: surface reconstruction*. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(5),401-412 1991.
- [16] D. Geman and S. Geman; *Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images*. *IEEE Transactions of Pattern Analysis and Machine Intelligence*. 6(6):721-741, 1984.
- [17] S. Geman and D. McClure; *Statistical methods for tomographic image reconstruction*. *Bull. Int. Stat. Inst.*, LII-4:5-21, 1987.
- [18] D. J. Hand; *Discrimination and Classification*, New York, 1981, John Wiley and Sons.

- [19] R. Harris; *Satellite Remote Sensing. An introduction*, London 1987, Routledge Kegan Paul.
- [20] R.M. Hoffer and P.H. Swain; *Computer processing of satellite data for assessing agriculture, forest and rangeland resources*. Proc. 14th Intern. Soc. Of Photog., Hamburg 1980, pp. 437-446
- [21] A.R. Huete; *Spectral signatures and vegetation indices*, II Reunión Nac. del grupo de trabajo en Teledetección (1987), pp 13-26.
- [22] D. Janzen; *Tropical dry forest. The most endangered major tropical ecosystem*, 1988. pp. 130-137. In. Wilson, E. O. (Ed.). Biodiversity. National Academic Press. Washington, D.C.
- [23] J.R. Jensen; *Introductory Digital Image Processing. A Remote Sensing Perspective*, Englewood Cliffs 1986, Prentice-Hall.
- [24] R.J. Johnston; *Multivariate Statistical Analysis in Geography*, London 1987, Longman.
- [25] C.O. Justice and J.R.G. Townshed; *Integrating ground data with remote sensing*, en J.R.G. Townshed Ed, *Terrain Analysis and Remote Sensing*, London 1981, George Allen & Unwin, pp. 38-58.
- [26] E.B. Knipling; *Physical and Physiological basis for the reflectance of visible and near-infrared radiation from vegetation*, Remote Sensing of Enviroment, vol. 1, 1970, pp. 155-159.
- [27] J.L. Labrandero; *Aplicación de Métodos Digitales al Reconocimiento de Suelos Mediante Sensores Remotos*, Madrid 1978. Tesis Doctoral. Univ. Autónoma de Madrid, Facultad de Ciencias.
- [28] Jorge Lira; *La percpción remota: nuestros ojos desde el espacio*. Fondo de la cultura económica, 1987.
- [29] D.G. Luenberger, *Linear and Nonlinear Programming*, Addison Wesley, reading, Mass., 1989.
- [30] K. Lulla; *Some observations on geobotanical remote sensing and mineral prospecting*, Canadian Journal of Remote Sensing, vol. 11, 1985, pp. 17-38
- [31] G.L. Markham and J.R.G. Townshend; *Land cover classification accuracy as a function of sensor spatial resolution*, Proc. 15th Intern. Symp. Rem. Sens. Of Environment, Ann Arbor (1981), pp. 1075-1090.
- [32] J.L. Marroquín; *apuntes del curso Visión Computacional* impartido por el Dr. José Luis Marroquín
- [33] J.L. Marroquín, *Deterministic Ineactive Particle Models for Images Processing and Comp. Graph.*, Graph. Mod. and Im. Proc., 55,5:408-417, (1993).
- [34] J.L. Marroquín, Edgar Arce, Salvador Botello; *Hidden Markov Measure Field Models for Image Segmentation*, Reporte Técnico No. I-02-05/12-04-2002 CC/CIMAT.
- [35] J.L. Marroquín, Edgar Arce, Salvador Botello; *Markov Random Measure Fields for Image Analysis*, IEEE 2002, International Conference of Image Processing MP-P3.4.
- [36] J.L. Marroquín, F. Velasco, M. Rivera and M. Nakamura, *Gauss-Markov Measure Field Models for Low-Level Vision*, IEEE Trans. On PAMI 23,4, p. 337-348 (2001).
- [37] J.L. Marroquín, S. Mitter and T. Poggio *Probabilistic solution of illposed problems in computational vision Journal of the American Statistical Association* 82(397):76-89, 1987.
- [38] P.M. Mather; *Computer Processing of Remotely Sensed Images*, Cichester 1987, John Wiley.

- [39] J. Meliá; *Signaturas Espectrales e Índices de Vegetación*, Burjassot (1986), Facultad de Ciencias Físicas Univ. Valencia.
- [40] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller; *Equations of state calculations by fast computational machine*. Journal of chemical Physics 21:1087-1902.
- [41] N.M. Milton, W. Collins, S.H. Chang and R.G. Schmidt; *Remote detection of metal anomalies on Pilot Mountain, Randolph Country, North Carolina*, Economic Geology, vol. 78, 1983, pp. 605-617.
- [42] M.A. Mulders; *Remote Sensing in Soil Science*, Amsterdam (1987), Elsevier.
- [43] P.G. Murphy, and A. E. Lugo; *Ecology of tropical dry forest*. Ann. Rev. Ecol. Syst. 17: 67-88 (1986)
- [44] P.A. Murtha; *Remote Sensing and vegetation damage: a theory for detection and assessment*, Photogrammetric Engineering and Remote Sensing, vol. 44, 1978, pp. 1147-1158.
- [45] J.C. Rajapakse and F. Krugge, *Segmentation of MR images with intensity inhomogeneities*, Image and Vis. Comp., 16:165-180, 1998.
- [46] M. Rebollo y F. Orti; *Pattern Recognition Techniques: Remote Sensing Applications at Centro de Investigación UAM-IBM*, Madrid (1980), Centro de Investigación UAM-IBM.
- [47] J. Rzedowski; *La Vegetación de México*. Limusa, México. 432p. (1978).
- [48] F.F. Sabins, *Remote Sensing*, New York, Freeman and Company (1996).
- [49] A.L. Yullie; *Generalized deformable models, statistical physics and matching problems*. Neural Computation 6:341-356, 1994.
- [50] J. Zhang; *The mean field theory in EM procedures for blind Markov random field image restoration*. IEEE transactions on Image processing 2(1):27-40, 1993.
- [51] J. Zhang; *The mean field theory in EM procedures for Markov random fields*. IEEE transactions on Image processing 40(10):2570-2583, 1992.
- [52] Y. Zhang, M. Brady and S. Smith; *Segmentation of brain MR images through a hidden markov random field model and the Expectation Maximization algorithm*, IEEE Trans. On Med. Im., 20,1:45-57, Jan. 2001.
- [53] <http://www.chem.vt.edu/chem-ed/light/em-rad.html>; Electromagnetic Radiation
- [54] <http://plata.uda.cl/minas/apuntes/Geologia/EXPLORAC/TEXT/02026satelite.html>
- [55] http://edcwww.cr.usgs.gov/glis/hyper/guide/landsat_tm