



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

POSGRADO DE INGENIERÍA ELECTRICA

**ANÁLISIS DE ESCENAS URBANAS DEL
DISTRITO FEDERAL MEDIANTE
DATOS TEXTURALES**

T E S I S

QUE PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERIA

P R E S E N T A :

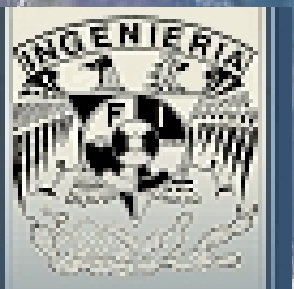
LEONARDO REYES SÁNCHEZ

DIRECTOR DE TESIS:

DR. MIGUEL MOCTEZUMA FLORES

Universidad Nacional Autónoma de México, DEPEI

MÉXICO, D.F. a Agosto del 2005.





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

ANÁLISIS DE ESCENAS URBANAS DEL DISTRITO FEDERAL MEDIANTE DATOS TEXTURALES

Universidad Nacional Autónoma de México, DEPFI. MÉXICO, D.F. a Agosto del 2005.

Agradecimientos

Agradezco a todas aquellas personas que hicieron posible este trabajo especialmente a:

Mi novia y compañera Liz que gracias a su apoyo y comprensión eh logrado subir algunos de los peldaños que se me han puesto en la vida y que sin su compañía difícilmente los lograría, muchas gracias hermosa.

A mi profesor Dr. Miguel Moctezuma Flores que gracias a su confianza y apoyo logre la culminación de este peldaño en mi vida, gracias.

A mis padres que gracias a su apoyo y comprensión eh logrado convertirme en el hombre que soy.

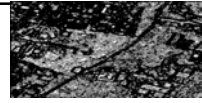
Y a mis compañeros de la maestría en telecomunicaciones de la F.I. Michael, Jacobo, Toño, etc.

Índice

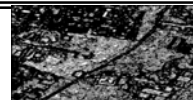
CAPÍTULO 1	Introducción 1
1.0	Introducción 2
CAPÍTULO 2	Fundamentos (Procesamiento digital y percepción remota) 6
2.0	Introducción 7
2.1	El espectro electromagnético 7
2.1.1	Características de la energía que emite el astro rey 7
2.1.2	La radiación electromagnética y la transferencia de energía 8
2.1.3	La energía electromagnética y la percepción remota 8
2.1.4	Características de las distintas regiones del espectro electromagnético 9
2.1.4.1	Las ondas de radiofrecuencia 9
2.1.4.2	Radiación de microondas 9
2.1.4.3	Radiación infrarroja 10
2.1.4.4	La luz visible 10
2.1.4.5	Radiación ultravioleta 11
2.1.4.6	Rayos x 11
2.1.4.7	Rayos gamma 11
2.2	Adquisición del espectro electromagnético 12
2.2.1	Sensores remotos 12
2.2.2	Nivel de adquisición de datos 13
2.2.3	Componentes de un sistema sensor 14
2.2.4	Tipos de sensores y su resolución 15
2.2.5	Adaptación a la iluminación y discriminación 16
2.3	Imagen digital 16
2.3.1	Imágenes monocromáticas 17
2.3.2	Imágenes en pseudo color 17
2.3.3	Imágenes en color real 18
2.3.4	Imágenes en falso color 18
2.3.5	Resolución y escala en las imágenes digitales 18
CAPÍTULO 3	Técnicas de segmentación y clasificación 20
3.0	Introducción 21

3.1	Crecimiento de regiones	22
3.2	Detección de cuencas	22
3.3	La clasificación	22
3.3.1	Clasificadores píxel a píxel	23
3.3.2	Clasificadores por regiones	23
3.3.2.1	Entrenamiento supervisado	24
3.3.2.2	Entrenamiento no supervisado	24
3.3.3	Texturas	25
3.4	Antecedentes de los modelos de campos aleatorios de Markov	25
3.4.1	Sistemas de vecindarios y asociación de tipos de cliques	26
3.4.2	El campo aleatorio de Markov (CAM) y la distribución de Gibbs	27
3.4.3	Modelos de campos aleatorios de Markov	28
3.4.4	Auto modelos	28
3.4.5	Modelo logístico multinivel	31
3.4.6	Ejemplo de segmentación por CAM y el algoritmo de optimización de recocido simulado (RS)	32
3.4.6.1	Algoritmo basado en CAM y optimizado mediante recocido simulado	33
3.5	Segmentación bayesiana	36
3.5.1	Presegmentación bayesiana	36
3.5.2	Segmentación bayesiana	37
3.5.3	Función discriminante y función óptima de Bayes	39
CAPÍTULO 4	El modelo textural	42
4.0	Introducción	43
4.1	Técnicas espectrales	43
4.2	Técnicas estructurales	44
4.3	Técnicas estadísticas	44
4.3.1	Estadísticas de primer orden	45
4.3.2	Estadística de segundo orden (G LCM)	45
4.4	Descriptores de textura	48
4.5	Modelo textural empleado	51
4.5.1	Algoritmo rápido para el cálculo de la matriz de co-ocurrencia	51

CAPÍTULO 5	Método de binarización y homogeneización de imágenes	56
5.0	Introducción	57
5.1	Método de binarización óptima	57
5.2	Método de binarización y homogeneización subjetiva propuesta	59
CAPÍTULO 6	Modelo de fusión	64
6.0	Introducción	65
6.1	Modelo de fusión utilizado en este trabajo	65
6.1.1	Caso binario	67
6.1.2	Caso multiclase	67
6.1.3	Evaluación de confiabilidad	67
6.2	Modelo de fusión	69
6.2.1	Ejemplo de fusión binaria	69
6.2.2	Modelo de fusión multiclasas propuesto	70
CAPÍTULO 7	Resultados	77
7.0	Resultados de segmentación bayesiana y el método contextual (Markov)	78
7.1	Descriptores de textura obtenidos mediante la matriz de co-ocurrencia	81
7.2	Resultados del proceso de binarización y homogeneización de las imágenes de textura	86
7.3	Imágenes resultantes del proceso de fusión binaria y multiclasas	98
CAPÍTULO 8	Conclusiones	118
8.0	Conclusiones	119
BIBLIOGRAFÍA		123



1 *Introducción*



1.0 Introducción

En este trabajo se pretende realizar el análisis de imágenes de fotografía aérea de alta resolución para escenas de la gran urbe del Distrito Federal. Con diversas texturas presentes en las mismas, el análisis de estas imágenes se realiza con objeto de aportar información sobre los problemas que representa el desarrollo urbano, ya que existen algunas zonas de crecimiento dinámico en donde no se cuenta con información actualizada sobre los diferentes elementos urbanos. En el caso de la presente tesis, nos interesa identificar de forma clara, las zonas urbanas y no urbanas.

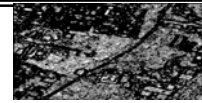
Una imagen pancromática proporciona información parcial, dada su geometría óptica monocular y el limitado intervalo espectral de adquisición. Las modas en el histograma pueden presentar una mezcla y en consecuencia, un proceso de segmentación puede arrojar elementos de escena solo parcialmente detectados. Por tales razones, a partir de una imagen pancromática, en la presente tesis se busca obtener descripciones estadísticas de los diversos elementos de escena, mediante operadores texturales. A su vez, las diversas matrices texturales nos proporcionan información parcial pero que conjuntamente pueden integrarse mediante un proceso de fusión de datos, a efectos de contar con resultados de segmentación más completos que los que proporcionarían el análisis de una sola imagen pancromática. El objetivo de esta tesis consiste en definir un esquema de análisis textural y de fusión de datos. Una aportación de esta tesis consiste en proponer un esquema de fusión de datos multiclases, a partir de un esquema binario de fusión.

El procesamiento digital de imágenes, es una disciplina que desarrolla las bases teóricas y algorítmicas mediante las cuales se puede extraer información del mundo real [AJB05], a partir de una imagen adquirida por medios satelitales o aéreos, lo cual es realizado por un área denominada percepción remota.

La percepción remota es un área del conocimiento que se encarga de la obtención, interpretación y estudio de imágenes (escenas) obtenidas por medio de sensores aéreos o satelitales, los cuales nos dan una interpretación de ciertos rasgos de la tierra. Esta área del conocimiento parte sus ramas en dos técnicas que son la percepción remota satelital y la percepción remota aérea [EC90] [GBTEL] en la cual se basa este trabajo. Con ayuda de la percepción remota y mediante técnicas de procesamiento digital de imágenes se puede extraer información de fotografías e imágenes aéreas. Las imágenes aéreas son analizadas y procesadas mediante diferentes algoritmos para poder obtener cierta información que a simple vista no es fácil de distinguir o diferenciar.

Entre los métodos más utilizados en el campo de la percepción remota y el procesamiento de imágenes se encuentran: la segmentación de imágenes, la clasificación y la fusión de imágenes. Estos tres métodos a su vez tienen diversas ramas de estudio y de aplicación así como diversos algoritmos de análisis.

Entre los métodos empleados en esta tesis se mencionan: la extracción de información textural mediante la matriz de co-ocurrencia, métodos de homogeneización de imágenes y por último la fusión de imágenes de fotografía aérea de alta resolución. Tales métodos se aplican de forma supervisada.

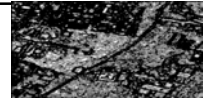


Con el propósito de comparar la funcionalidad del método propuesto, también se implementa un esquema empleando un algoritmo basado en campos aleatorios de Markov y teoría de Bayes.

En el presente documento se presenta una parte teórica y otra algorítmica experimental de los esquemas programados. A continuación se presenta una breve descripción del temario:

- El capítulo 1: Presenta una breve introducción a la presente tesis con el propósito de ofrecer un panorama general de la misma.
- En el capítulo 2: Se presentan algunos fundamentos del procesamiento digital de imágenes y de la percepción remota, entre los cuales se abarcan temas como: El espectro electromagnético y sus características, la imagen digital, etc.
- En el capítulo 3: Son mostradas algunas de las técnicas y principios de segmentación y clasificación de imágenes, como lo son los campos aleatorios de Markov (CAM) y el teorema de Bayes.
- En el capítulo 4: Se presenta una breve introducción al análisis de texturas, así como la descripción del modelo estadístico empleado en este trabajo. Se utilizan matrices de co-ocurrencia (**Grey Level Co-occurrence Matriz -GLCM-**) a fin de obtener descriptores de textura para su posterior análisis en la etapa de fusión de información.
- En el capítulo 5: Se explican los fundamentos y métodos de homogeneización supervisada y no supervisada de imágenes en n-clases. Estos métodos facilitan la recuperación de la información contenida en los descriptores de textura obtenidos previamente en el capítulo anterior.
- En el capítulo 6: Se presenta una descripción del proceso de fusión y de sus técnicas. También se expone el modelo de fusión empleado. Una parte importante de esta tesis es la propuesta de un esquema generalizado a procesos de fusión multiclases.
- En el capítulo 8: Son mostrados los resultados obtenidos de los procesos de segmentación textural, homogeneización de los descriptores de textura en n-clases y fusión. Para la fusión se emplearon los resultados obtenidos de los procesos mencionados anteriormente.
- En el capítulo 9: Son presentadas las conclusiones generales de este trabajo, así como de los resultados obtenidos.

Para cumplir con el objetivo de esta tesis, se trabajó con 2 imágenes pancromáticas de diferentes zonas del sur de la Ciudad de México, a las que se proceso inicialmente por medio de un algoritmo rápido para el cálculo de la matriz de co-ocurrencia, con el propósito de extraer información de diversos descriptores texturales.

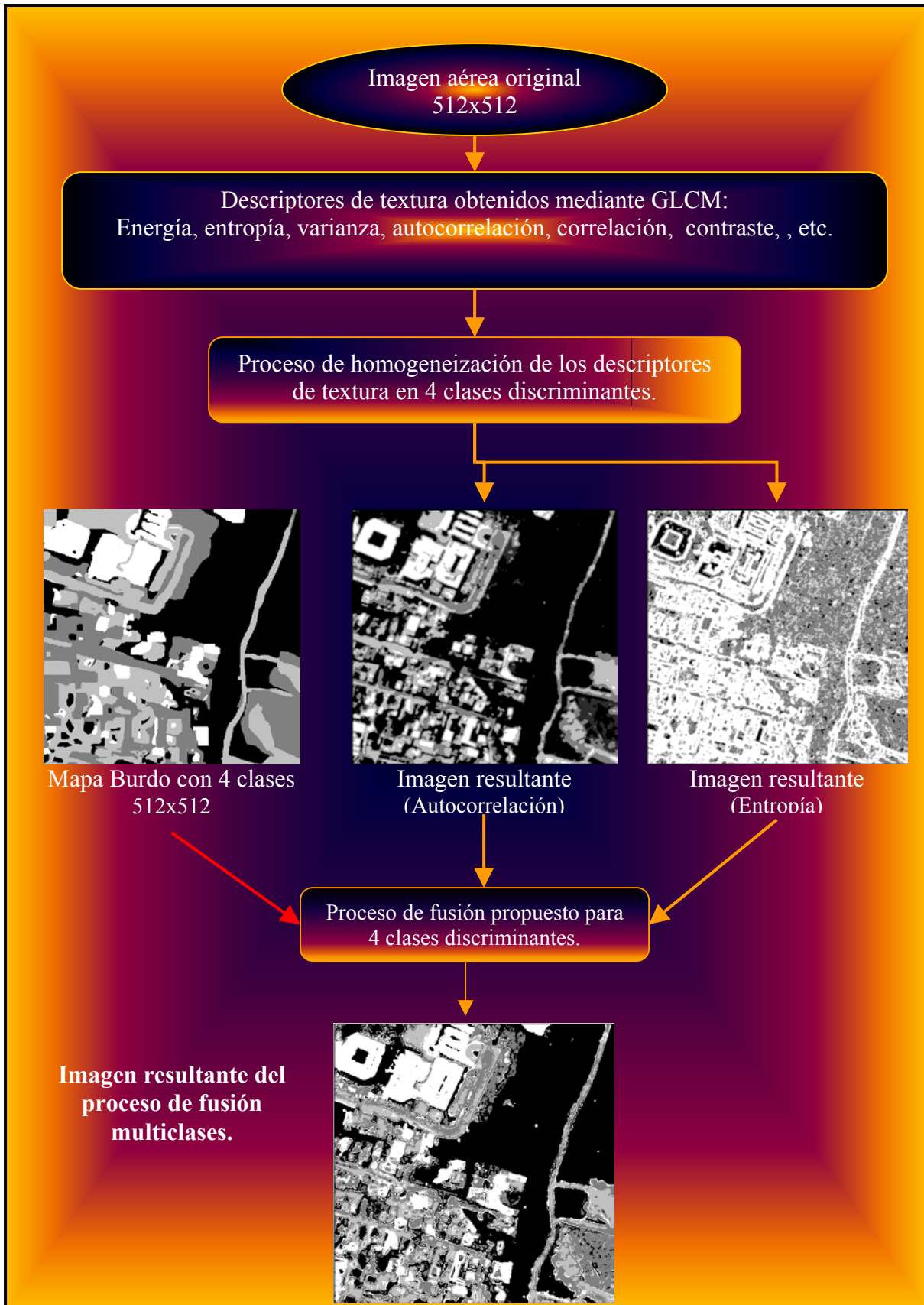
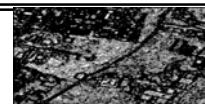


Después, a las imágenes obtenidas se les aplicó un algoritmo de homogeneización supervisado para 4 clases. Estas representan diferentes texturas o clases presentes en la imagen, es decir; éstas pueden estar representadas por áreas verdes, zonas urbanas, senderos de terracería, etc.

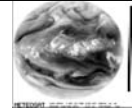
Por último se aplicó el modelo de fusión multiclases propuesto con el mismo criterio de umbralaje y experimentación empleada en el proceso de homogeneización.

Con objeto de comparar la calidad de los resultados obtenidos mediante el método de fusión propuesto, se realizaron algunos ejemplos de fusión binaria y se verifico la calidad y eficacia del proceso de fusión multiclases.

A continuación se presenta un diagrama del proceso seguido a través de este trabajo:

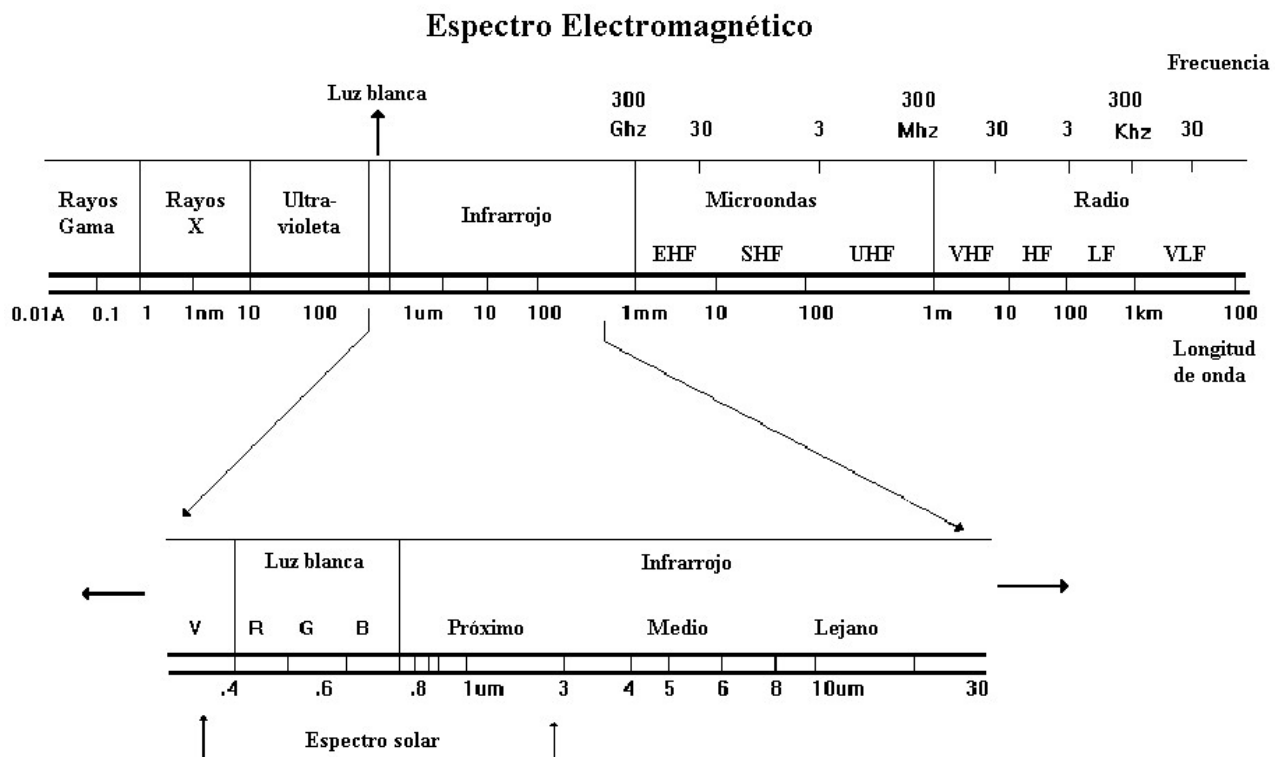


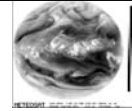




2 Fundamentos

(Procesamiento digital y percepción remota)





2.0 Introducción

El objetivo de este capítulo es la introducción a varios conceptos relacionados con las imágenes digitales así como fundamentos y notación que se empleara a lo largo de la tesis [RJ86] [RCG96]. En primera instancia se resumirán conceptos fundamentales, tales como el espectro electromagnético, sus características, usos y aplicaciones de este en el procesamiento de imágenes digitales. También se hablará de los elementos usados en la adquisición del espectro electromagnético y de las imágenes digitales adquiridas por dichos elementos.

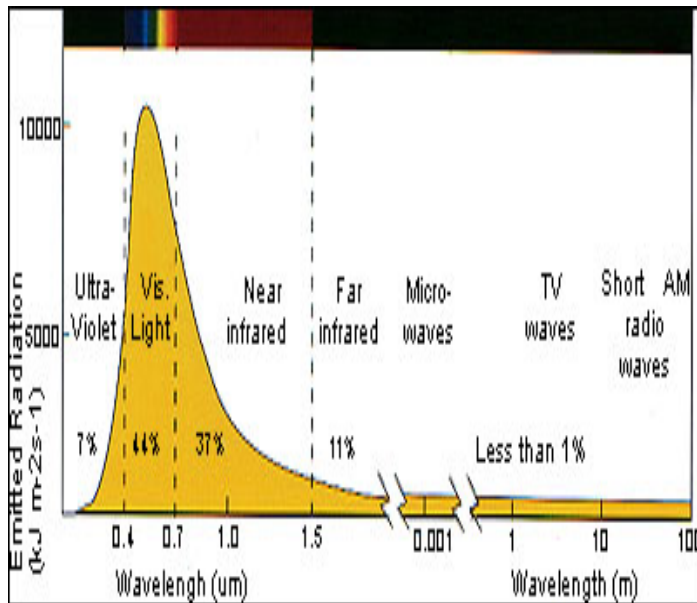
2.1 El espectro electromagnético

2.1.1

Características de la energía que emite el astro rey

La radiación solar incide sobre todo el planeta constantemente. A la distancia que se encuentra la Tierra del Sol, el planeta recibe un flujo constante de energía de alrededor de 1,38 kilovatios por metro cuadrado en la capa superior de la atmósfera. Este valor posee muy pocas variaciones y se conoce como constante solar [FCAGR].

Podríamos medir en la parte superior de la atmósfera, qué longitudes de onda posee la energía que proviene del Sol, y así desglosar sus cantidades, o bien podemos calcular el "espectro de emisión" de la energía solar, asumiendo que el Sol es un absorbedor y emisor de energía perfecto, es decir, lo que se conoce físicamente como un "cuerpo negro".



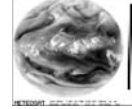
La temperatura de la fotosfera, la región más superficial del Sol, es de alrededor de 5700 grados centígrados.

Esto significa que la potencia radiada a través de cada metro cuadrado de la superficie solar es de 63 megavatios.

En otros términos, el flujo total de energía que radia el Sol asciende a $3,9 \times 1.026$ Joules por segundo.

La figura 2.1 muestra el espectro de emisión del Sol asumiendo que su temperatura es la antes dicha. Nótese que hay un pico de energía cercano a los 0,5 micrómetros o micrones, en el intervalo de longitudes de onda del visible, es decir; entre los 0,4 y los 0,7 micrones.

Figura 2.1 Espectro de emisión del Sol: Espectro de emisión de la radiación solar



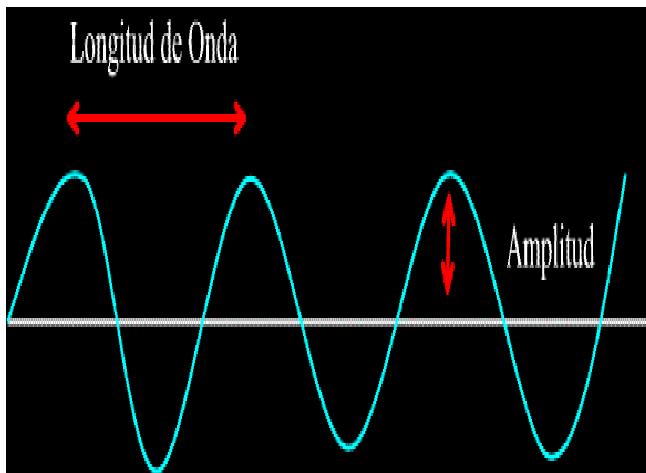
2.1.2 La radiación electromagnética y la transferencia de energía

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

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

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

La radiación de energía, a diferencia de la conducción o la convección, no necesita de ningún "medio" para la transferencia de energía. Por ejemplo, la energía radiactiva viaja por el vacío del espacio, mientras que la convección transporta calor solo a través de cuerpos o a través de fluidos como lo son la atmósfera o los océanos.



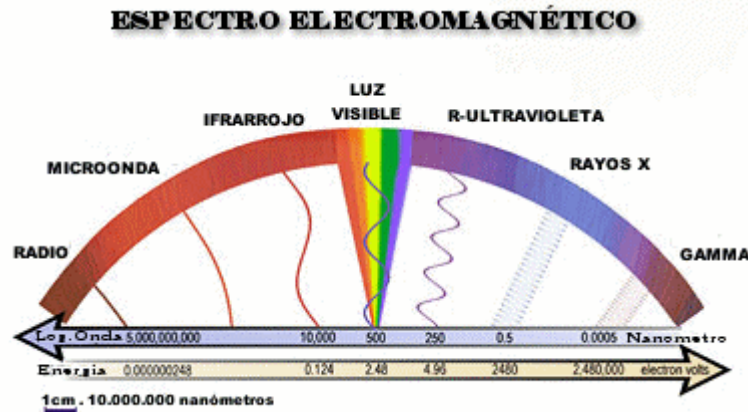
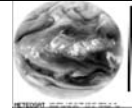
La intensidad de la energía que un objeto radia depende básicamente de su temperatura en una relación directa con su longitud de onda.

Usamos el término "longitud de onda" para definir esta energía radiada (o radiación) porque la forma en la que la energía viaja por el espacio o a través de cuerpos y fluidos se asemeja a una onda sinusoidal. La energía asume una gran cantidad de longitudes de onda, las cuales dependen de sus características particulares y peculiares.

Figura 2.2 Longitud de onda

2.1.3 La energía electromagnética y la percepción remota

Cuando nos referimos a energía electromagnética o radiación en percepción remota, la característica más común a tener en cuenta es la "longitud de onda", debido a que los "emisores de energía" en la naturaleza suelen poseer temperaturas muy extremas, desde el frío aire antártico del mes de Julio hasta la candente superficie del astro solar, el intervalo de longitudes de onda también es muy amplio. Si las ondas electromagnéticas se organizan en un continuo de acuerdo a sus longitudes se obtiene el "espectro electromagnético"[TCK05], donde las ondas más largas (longitudes desde metros a kilómetros) se encuentran en un extremo (radio) y las más cortas en el otro, al intervalo completo de longitudes de onda se le conoce como espectro electromagnético.

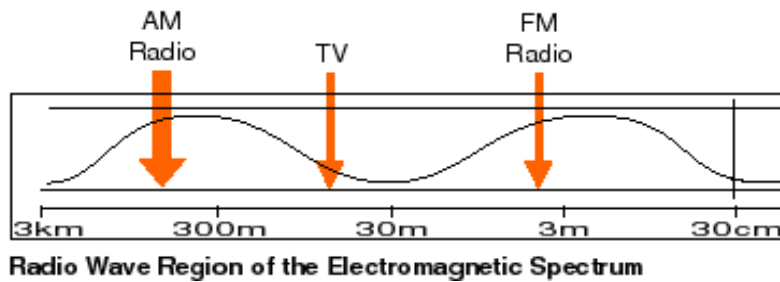


Figuras 2.3 El espectro electromagnético [PTD05]

2.1.4 Características de las distintas regiones del espectro electromagnético

2.1.4.1 Las ondas de radiofrecuencia

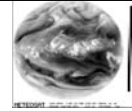
Estas ondas poseen frecuencias que van de 0 a 10^9 Hz. y son usadas en los sistemas de radio y televisión, generándose mediante circuitos oscilantes. Las ondas de radiofrecuencia y las microondas son especialmente útiles por que en esta pequeña región del espectro las señales producidas pueden penetrar las nubes, la niebla y las paredes.



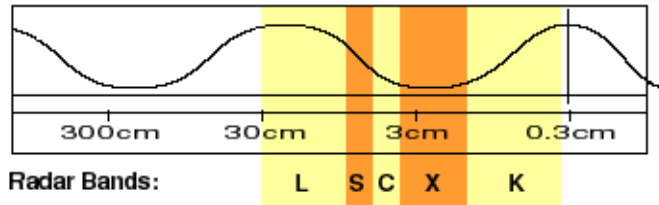
Figuras 2.4 Las ondas de radiofrecuencia

2.1.4.2 Radiación de microondas

La región de las microondas se encuentra entre los 10^9 hasta aproximadamente 3×10^{11} Hertz (con longitud de onda entre 30 cm. a 1 mm), estas son utilizadas en el radar y otros sistemas de comunicación, así como en el análisis de detalles muy finos de la estructura atómica y molecular.



Microwave region of the Electromagnetic Spectrum

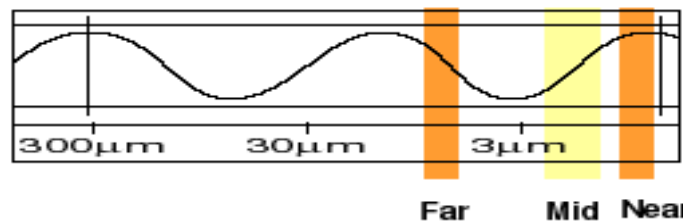


Figuras 2.5 Radiación de microondas

2.1.4.3 Radiación infrarroja

La radiación infrarroja se localiza en el espectro entre 3×10^{11} hz. hasta aproximadamente los 4×10^{14} Hz. Se subdivide en tres regiones, infrarrojo lejano, medio y cercano (Lejana 50000nm – 1mm, Intermedia 2500 - 50000 nm y cercana 780 - 2500 nm).

Infrared Region of the Electromagnetic Spectrum



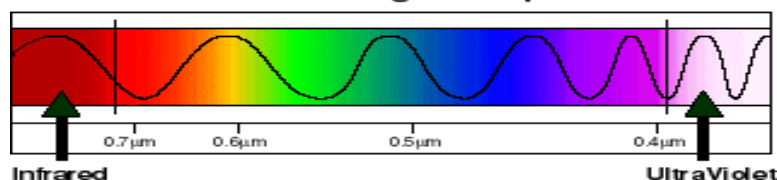
Figuras 2.6 Radiación infrarroja

Toda molécula que tenga un temperatura superior al cero absoluto (-273° K) emite rayos infrarrojos y estos serán mayores entre mas temperatura tenga el objeto. En la región infrarroja de la emisión terrestre la mayoría de la energía emitida escapa hacia el espacio debido a que la atmósfera es "transparente" a ella. Esto es debido a la composición de la mezcla de gases que integran la atmósfera y de la temperatura de nuestro planeta.

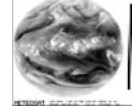
2.1.4.4 La luz visible

Esta es una región muy estrecha pero una de las más importantes, ya que nuestra retina es sensible a las radiaciones de estas frecuencias.

Visible Light Region of the Electromagnetic Spectrum



Figuras 2.7 La luz visible



Esta región se subdivide en seis intervalos que definen los colores básicos (rojo, naranja, amarillo, verde, azul y violeta). La radiación visible va desde 384×10^{12} hasta 769×10^{12} Hz. Las frecuencias más bajas de la luz visible (longitud de onda larga) se perciben como rojas y las de más alta frecuencia (longitud corta) aparecen violetas.

2.1.4.5

Radiación ultravioleta

La región ultravioleta ocupa una extensa banda del espectro que va de 10 nm a 380 nm. Esta banda es utilizada para la detección de minerales por luminiscencia y contaminación marina. Uno de los grandes obstáculos para la utilización de esta región del espectro, es la fuerte atenuación atmosférica [DPI]. Los átomos y moléculas sometidos a descargas eléctricas producen este tipo de radiación.

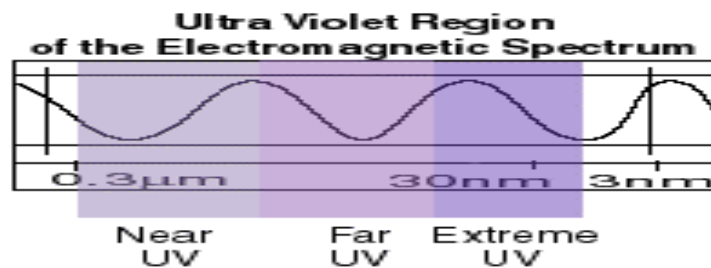


Figura 2.8 Radiación ultravioleta

2.1.4.6

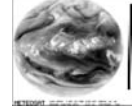
Rayos X

En 1895 Wilhelm Röntgen inventó una máquina que producía radiación electromagnética con una longitud de onda menor a 10 nm a la cual, debido a que no conocía su naturaleza la bautizó como X. Si se aceleran electrones y después se les hacen chocar con una placa metálica, la radiación de frenado produce rayos X. Los rayos X se han utilizado en medicina desde el mismo momento en que los descubrió Röntgen debido a que los huesos absorben mucho más radiación que los tejidos blandos. Debido a la gran energía de los fotones los rayos X son muy peligrosos para los organismos vivos.

2.1.4.7

Rayos gamma

Se localizan en la parte del espectro que tiene las longitudes de onda más pequeñas, entre 10 y 0.01 nm. Se producen en los procesos nucleares, por ejemplo: cuando se desintegran las sustancias radioactivas. La enorme energía de los fotones gamma los hace especialmente útiles para destruir células cancerosas. Pero son también peligrosos para los tejidos sanos, por lo que la manipulación de rayos gamma requiere de un buen blindaje de protección.



2.2 Adquisición del espectro electromagnético

Para la adquisición digital de imágenes se necesitan dos elementos básicos:

El primero es un dispositivo físico sensible a una determinada banda del espectro de energía electromagnética, que produce una señal eléctrica proporcional al nivel de energía detectado, y el segundo es un dispositivo denominado digitalizador, encargado de convertir la señal de salida de un sistema sensible a una forma digital [DPI].

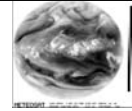
Una amplia categoría de sensores está relacionada con la luz visible e infrarroja; entre los dispositivos empleados con mayor frecuencia se encuentran los microdensitómetros, diseccionadores de imágenes, cámaras vidicon, matrices de detectores fotosensibles de estado sólido, etc. Los sensores son equipos capaces de captar energía proveniente de un objeto, convirtiéndola en una señal posible de ser registrada y presentada en forma adecuada para la extracción de informaciones [RCG96].

Para entender mejor podemos dividir la fase de adquisición de datos en los siguientes puntos:

- **Plataforma:** Aérea o espacial, la selección de la plataforma depende de los objetivos del programa. La combinación de los dos tipos (aéreo y espacial) genera excelentes resultados.
- **Sensor:** Existen diversos tipos y su selección está íntimamente relacionada con el tipo de fenómeno a ser investigado.
- **Trayectoria:** Depende de la localización del área de interés, así como de la necesidad de recubrimiento espacial y temporal.
- **Verdad terrestre:** Los datos de campo (factor vital). Es a través del muestreo que se consigue el entrenamiento, la verificación y la evaluación de todo el sistema de percepción remota.
- **Preprocesamiento:** Engloba las modificaciones de las informaciones obtenidas por los sensores antes de retornar la plataforma a la Tierra o antes de ser transmitidas a las estaciones de rastreo. En esta etapa se incluyen las correcciones radiométricas, geométricas, conversión de datos analógicos en digitales, etc.

2.2.1 Sensores remotos

Los sensores son sistemas fotográficos u óptico-electrónicos capaces de detectar y registrar en forma de imágenes o no, el flujo de energía radiante reflejado o emitido por objetos distantes. Un flujo de radiación electromagnética (energía) al propagarse por el espacio puede interactuar con superficies u objetos, siendo reflejado, absorbido o emitido por dichas superficies u objetos.



Todos los instrumentos que captan y transforman esa energía pueden ser clasificados como sensores. Una parte de esta radiación solar reflejada retorna hacia el espacio y es captada por sensores y satélites para formar sus imágenes satelitales.

Durante la fase de adquisición de datos por los sensores, se pueden distinguir los siguientes elementos básicos: energía radiante, fuente de radiación, objeto (albo), trayectoria y sensor (sistema de imagen óptica y detector).

TYPICAL ALBEDOS	
Surface or object	Albedo(%)
Fresh snow	75 to 95
Clouds (thick)	60 to 90
Clouds (thin)	30 to 50
Ice	30 to 40
Sand	15 to 45
Grassy field	10 to 30
Dry ploughed field	5 to 20
Water (daily average)	10
Forest	3 to 10
Earth and atmosphere	30
Moon	7

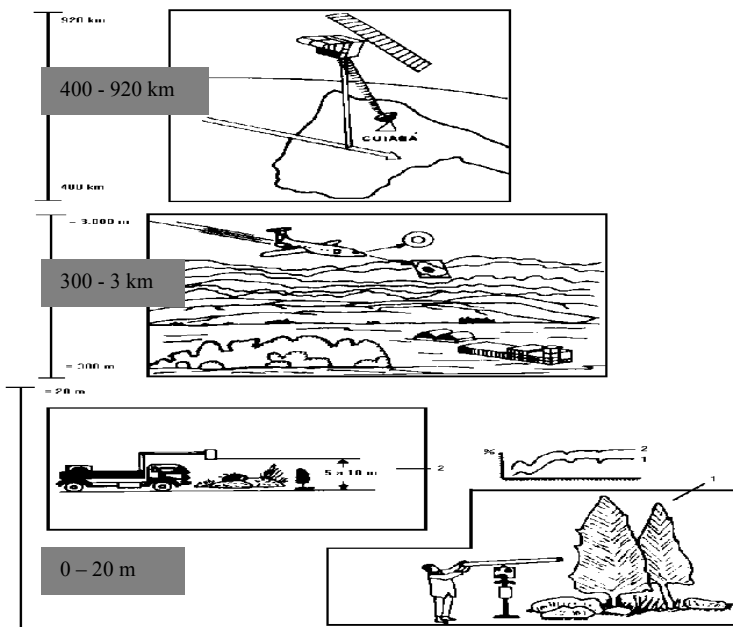
Los objetos de la superficie terrestre son "vistos" por los sensores remotos de la misma manera que los seres humanos vemos con los ojos, lo cual depende de la luz reflejada por los objetos.

La cuantificación de la energía que los distintos objetos son capaces de reflejar se le conoce como "albedo" y es expresado en valores porcentuales (figura 2.9).

Figura 2.9 Albedos típicos

2.2.2

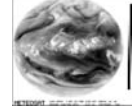
Nivel de adquisición de datos



En la siguiente figura se muestran los niveles de adquisición de datos en percepción remota, ya que la altitud del sensor con relación a la superficie abarcada, también es un factor de gran interferencia, no sólo en la intensidad y calidad de la señal sino también en las formas de registro y análisis de los datos.

Al pasar de un nivel a otro, se modifican las dimensiones del área a ser adquirida, consecuentemente se obtienen imágenes con diferentes resoluciones. A continuación se muestran estos niveles de adquisición de datos.

Figura 2.10 Niveles de adquisición de datos

**2.2.4****Tipos de sensores y su resolución**

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

En función de la fuente de energía:

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

En función del tipo de producto:

- 1) **No generadores de imagen:** No generan una imagen de la superficie observada. Por ejemplo, los radiómetros (salida en dígitos o gráficos) y los espectros radiómetros (firma espectral). Son esenciales para la adquisición de informaciones minuciosas sobre el comportamiento espectral de los objetos de la superficie terrestre.
- 2) **Generadores de imagen:** Se obtiene como resultado una imagen de la superficie observada. Suministran informaciones sobre la variación espacial de la respuesta espectral de la superficie observada.

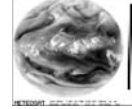
Algunos tipos de sensores utilizados en satélites son:

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

MSS (Multi Spectral Scanner): Es un sistema sensor que permite la adquisición de líneas del terreno en una faja de 185 Km. perpendicularmente a la órbita del satélite. El barrido del terreno es realizado con ayuda de un espejo que oscila perpendicularmente al desplazamiento del satélite.

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

La resolución es una medida de la habilidad que posee un sistema de sensor en distinguir entre las respuestas espectralmente semejantes o próximas espacialmente. La resolución puede ser clasificada en espacial, espectral y radiométrica.



Resolución espacial: Mide la menor separación angular o lineal entre dos objetos; es decir, una resolución de 20 metros implica que objetos distanciados entre sí menos de 20 metros, no serán discriminados por el sistema.

Resolución espectral: Es una medida del ancho de las bandas espectrales del sistema sensor; es decir, un sensor que opera en el intervalo de 0,4 a 0,45 μm tiene una resolución espectral menor que un sensor que opere en el intervalo de 0,4 a 0,5 μm .

Resolución radiométrica: Está asociada a la sensibilidad del sistema sensor para distinguir dos niveles de intensidad de la señal de retorno, es decir, una resolución de 10 bits (1024 niveles digitales) es mejor que una de 8 bits.

2.2.5

Adaptación a la iluminación y discriminación

Debido a que las imágenes digitales son representadas como un conjunto de puntos brillantes, es de considerable importancia la capacidad del ojo para discriminar diferentes niveles de iluminación, por lo cual esto es de considerarse importante para poder presentar resultados de un procesamiento de imagen. El intervalo de niveles de intensidad de luz a los que es capaz de adaptarse el ojo es enorme, del orden de 10^{10} desde el umbral escotópico hasta el límite de deslumbramiento. Existen evidencias experimentales que indican que la iluminación subjetiva es una función logarítmica de la intensidad de la luz incidente sobre el ojo.

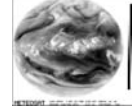
El intervalo total de niveles de iluminación que puede discriminar simultáneamente es bastante pequeño comparado con el intervalo total de adaptación, para cada conjunto de condiciones, el nivel de sensibilidad del sistema visual se denomina nivel de adaptación a la iluminación.

2.3

Imagen digital

En percepción remota se utiliza un gran número de datos para representar una imagen, los cuales pueden ser manipulados en formato digital, con la finalidad de extraer cierta información de estas imágenes. Cada punto captado por los sensores, corresponde a un área mínima denominada "píxel" (picture element), el cual está bien identificado geográficamente, y para el cual se registran valores digitales relacionados con la intensidad de energía reflejada en diferentes bandas bien definidas del espectro electromagnético [LSS05].

La formación de una imagen digital o digitalización de una imagen, es la conversión de la función continua $f(x,y)$ en una serie de valores discretos (generalmente valores enteros), por muestreo tanto del plano espacial como de la intensidad [RCG96].



De esta forma una imagen digital es representada como una tabla de valores o matriz bidimensional (conjunto de valores numéricos ordenados en filas y columnas $I [M,N]$), donde cada elemento corresponde al valor de intensidad de cada punto de muestreo de la imagen, a lo que denominamos como elemento mínimo de la imagen o píxel.

Un píxel es designado con el nombre de la matriz y los subíndices que indican su pertenencia a una fila y columna determinada, es decir; $I_{(i,j)} = 127$ indica que el píxel que se encuentra en la fila i y columna j en la imagen I tiene un valor de intensidad o nivel de gris igual a 127.

Las dimensiones virtuales de una imagen (número de píxeles) dependen del número de filas y columnas que tenga la matriz que la define. De esta manera una imagen de 512 x 512 píxeles, corresponde a una matriz de 512 columnas y 512 filas, con un tamaño total de 262144 píxeles.

Para realizar el procesamiento digital de imágenes de datos en percepción remota, es necesario que la imagen esté en formato digital. Existen básicamente dos maneras de obtener una imagen digital: (1) Adquirir la imagen de teledetección en formato analógico (por ejemplo, fotografía aérea) y posteriormente digitalizarla; (2) Adquirir la imagen de percepción remota ya en formato digital, tales como los datos en cintas CCT (Computer compatible tape), grabados por satélites como Landsat y SPOT.

2.3.1

Imágenes monocromáticas



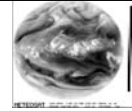
Desde el punto de vista físico, una imagen monocromática es una imagen con tonalidades de grises o fotografía en blanco y negro, la cual es una función continua $f(x,y)$, que expresa la variación en el espacio de la cantidad de luz emitida o reflejada por la imagen. Esta cantidad de luz o energía procedente de la imagen es lo que comúnmente denominamos brillo, intensidad o nivel de gris de los distintos objetos o detalles contenidos en la imagen.

Figura 2.12 Imagen monocromática

2.3.2

Imágenes en pseudo color

Frecuentemente a las imágenes monocromas se les representa en color, atribuyendo arbitrariamente un color fijo a cada valor de intensidad, mediante una paleta de colores.



En estos casos se habla de pseudo color, ya que la información sobre el color de cada píxel sólo puede restablecerse a partir de la paleta de colores pero nunca a partir de los propios valores de intensidad de la imagen. Dichas paletas de color pueden utilizarse en el realce de ciertos aspectos en imágenes monocromáticas con procesos de segmentación, es decir; asignando distintos colores a las distintas fases de una aleación o a los distintos intervalos de temperatura en una imagen térmica [LSS05].

Figura 2.13 Imagen pseudo color codificada en 32 colores

2.3.3

Imágenes en color real

En general se habla de imágenes en color real, cuando la codificación de la imagen contiene información sobre tres o más componentes de color. Ya sean las componentes RGB (rojo, verde y Azul), HSI (croma o tono, saturación e intensidad), CMY (cyan, magenta, amarillo), CMYK (cyan, magenta, amarillo y negro) u otras.



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

Figura 2.14 Imagen en color real

2.3.4

Imágenes en falso color

Se habla de imágenes en falso color cuando en uno o más de uno de los componentes normales del color se visualizan longitudes de onda del espectro electromagnético, que están fuera del espectro visible.

2.3.5

Resolución y escala en las imágenes digitales

Uno de los aspectos más importantes, para definir la calidad de las imágenes digitales, es la resolución física o superficial de los píxeles, ya que siendo el píxel el elemento más pequeño en el que puede dividirse una imagen digital, la superficie real que representa cada uno de ellos define los objetos o detalles más pequeños que pueden observarse en una imagen [LSS05].

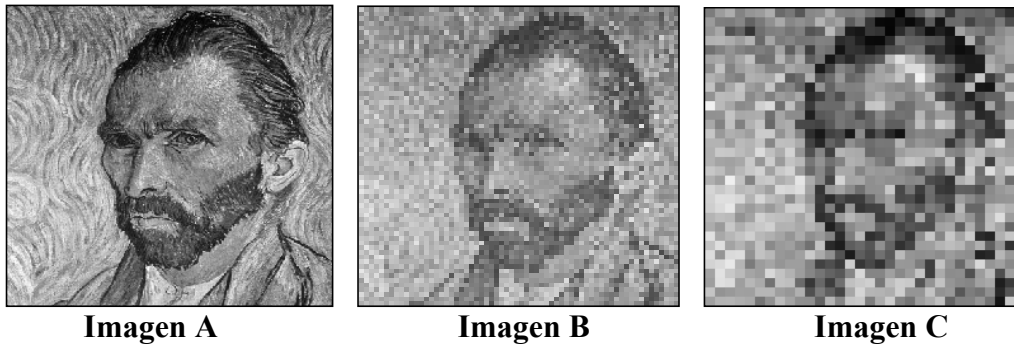
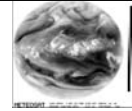


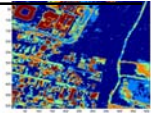
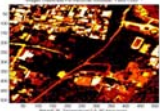
Figura 2.15 Imágenes con diferentes resoluciones

- Imagen A de 256x256 píxeles cada píxel ocupa un punto del monitor.
- Imagen B 64x64 píxeles y cada píxel se ha representado al mismo tamaño que en la imagen de 256x256 (un punto del monitor)
- Imagen C de 32x32 píxeles (cada píxel de la imagen ocupa 64 puntos de monitor).

Nota: Es importante no confundir la resolución de una imagen con su escala de representación. La escala de representación de una imagen es la relación que existe entre el tamaño real de los objetos y el tamaño que tienen en la representación gráfica.

Una imagen digital es en sí adimensional y puede ser representada de cualquier tamaño.

La escala gráfica de una imagen depende del tamaño físico con que se representa cada píxel. De forma contraria, la resolución de la imagen indica el tamaño de la ventana de muestreo o barrido, que se utilizó en la digitalización de la imagen.



3 Técnicas de segmentación y clasificación

*Imagen Clasificada Por Recocido Simulado "Falso Color" *

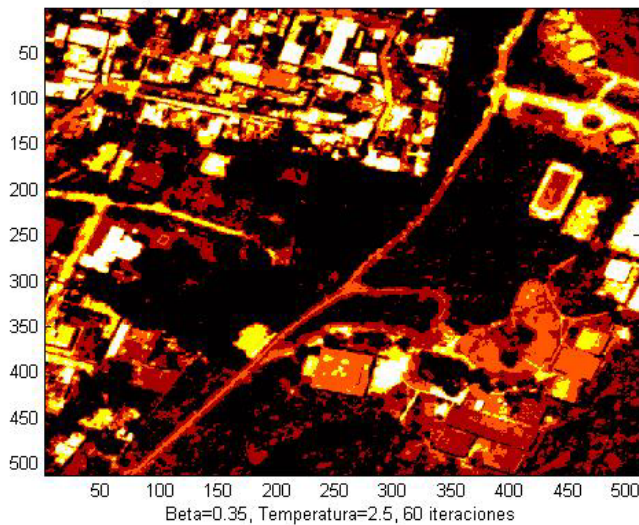
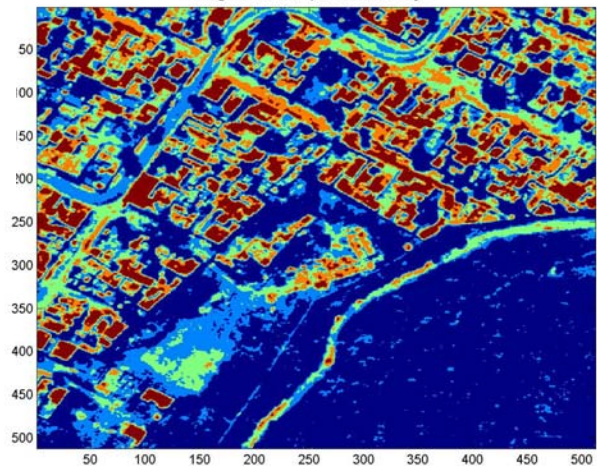
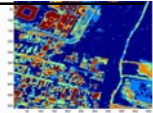
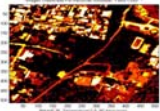


Imagen Clasificada por Criterio de Bayes





3.0 Introducción

La segmentación es el proceso de dividir una imagen en regiones o segmentos manipulables, la cual se emplea en el análisis de imágenes para separar propiedades o regiones de una textura dada en la imagen. La segmentación se puede basar en alguna de las siguientes propiedades:

Similitud: Mediante esta propiedad se segmenta la imagen en regiones que tienen niveles de gris dentro de un intervalo predeterminado, utilizando técnicas de fijación de umbrales y el crecimiento de regiones.

Discontinuidad: Esta propiedad se utiliza para segmentar la imagen en regiones de discontinuidad, donde existe un cambio abrupto en los valores de los niveles de gris. Lo anterior se utiliza para detectar las fronteras o bordes en la imagen y es consecuentemente conocida como detección de bordes.

Para clasificar las diferentes regiones obtenidas en el proceso de segmentación, se divide el conjunto de píxeles que componen la imagen en clases temáticas previamente definidas o por definir, según sea el método clasificatorio que se decida utilizar.

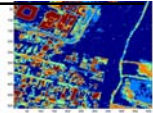
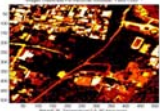
Durante la clasificación se requiere la adopción de métodos que incorporen reglas de decisión, los cuales se pueden agrupar en dos grandes categorías: los supervisados y los no supervisados. La diferencia entre estos métodos es que, en el caso de los primeros, se requiere de un conocimiento previo muy preciso de las clases a clasificar, mientras que en los no supervisados, la segmentación de las clases espectrales se obtiene con base a un procedimiento estadístico, lo cual genera clases estadísticas, las que posteriormente, deberán ser asignadas a clases o categorías de información de acuerdo con antecedentes de terreno.

La clasificación estadística es un procedimiento convencional de análisis digital de imágenes, el cual constituye un proceso de análisis de píxeles de forma aislada.

La clasificación presenta la limitación de que el análisis puntual se basa únicamente en atributos espectrales, y para superar estas limitaciones, se propone el uso de la segmentación de imágenes, como una fase previa a la fase de clasificación, en donde se extraen los objetos relevantes para la aplicación deseada.

En este proceso, se divide la imagen en regiones que deben corresponder a las áreas de interés de la aplicación. Dichas regiones son un conjunto de píxeles contiguos, que se dispersan bidireccionalmente y que presentan uniformidad entre sí.

La división en porciones consiste básicamente en un proceso de crecimiento de regiones, de detección de bordes o de detección de cuencas.



3.1 Crecimiento de regiones

El crecimiento de regiones es una técnica de agrupamiento de datos, en la cual sólo las regiones adyacentes espacialmente, pueden ser unidas. En principio, este proceso de segmentación rotula cada píxel como perteneciente a una región diferente y después calcula un criterio de similitud para cada par de regiones adyacentes espacialmente.

El criterio de similitud se basa en una prueba de hipótesis estadísticas, que prueba la media entre las regiones. Luego se divide a la imagen en un conjunto de subimágenes y se realiza la unión de regiones, según un umbral de agregación definido.

Para que dos regiones A y B vecinas sean unidas, se debe adoptar el siguiente criterio:

- Decidir si A y B son similares (prueba de las medias).
- Decidir si la similitud satisface el umbral establecido.
- A y B son mutuamente próximas (entre los vecinos de A, B es el más próximo, y entre los vecinos de B, A es el más próximo).

En el caso de que las regiones A y B satisfagan este criterio, dichas regiones serán agregadas, y en caso contrario el sistema reiniciará el proceso de prueba de agregación.

3.2 Detección de cuencas

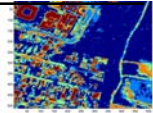
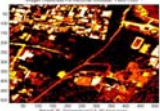
La segmentación por detección de cuencas es realizada sobre una imagen resultante de la extracción de bordes. La extracción de bordes es realizada por un algoritmo de detección de bordes, tal como el filtro de Sobel. Este algoritmo considera los gradientes de nivel de gris de la imagen original para generar una imagen gradiente o imagen de intensidad de borde.

El algoritmo calcula un umbral para el seguimiento de bordes y cuando se encuentra un píxel con valor superior al umbral establecido, se inicia el proceso de seguimiento del borde. Se observa la vecindad para identificar el próximo píxel de mayor valor de nivel digital y se sigue en esa dirección hasta que se encuentre otro borde o la frontera de la imagen. De este proceso se origina una imagen binaria, en la cual el valor 1 representa los bordes y el nivel 0 representara a las regiones de no bordes [DPI].

3.3

La Clasificación

La clasificación es el proceso de extracción de información en imágenes para reconocer patrones y objetos homogéneos. Los métodos de clasificación se utilizan para mapear áreas de la superficie terrestre que presentan un significado igual en imágenes digitales.



La información espectral de una escena puede ser representada por una imagen espectral, donde cada píxel tiene las coordenadas espaciales (x,y) , además de la coordenada espectral L que representa la radiancia de un albedo en el intervalo de longitud de onda de una banda espectral. Cada píxel de una banda posee una correspondencia espacial con otro píxel, en todas las otras bandas; es decir, para una imagen de N bandas, existen N niveles de gris asociados a cada píxel siendo uno para cada banda espectral.

Dependiendo del proceso de clasificación a utilizar, los clasificadores pueden ser divididos en clasificadores píxel a píxel o clasificadores por regiones.

3.3.1 Clasificadores píxel a píxel

Este tipo de clasificadores utilizan sólo la información espectral aislada de cada píxel para encontrar regiones homogéneas. Estos clasificadores pueden ser divididos en métodos estadísticos (los cuales utilizan reglas de la teoría de probabilidad) y métodos determinísticos.

Las imágenes son composiciones complejas de píxeles, no de objetos, que producen patrones característicos variables aparentemente según la escala de observación. Por lo tanto, son fundamentales tanto el observador como la escala de observación para reconocer dichos patrones. Algunas de las técnicas automáticas que se han desarrollado para discriminar objetos y cubiertas, son:

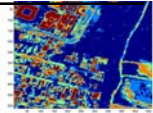
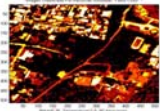
- Análisis de texturas (Haralick et al., 1973)
- Redes bayesianas
- Técnicas piramidales y multiescala
- Wavelets
- Métodos paramétricos, etc.

Diferentes criterios insisten en la aproximación multicriterio en la clasificación y segmentación de una imagen de alta resolución espacial, utilizando métodos basados en el análisis “píxel a píxel” junto con métodos que extraigan información de contexto (en este caso un patrón de textura). Esta combinación es especialmente relevante cuando la resolución es alta y las imágenes, al ofrecer un mayor nivel de detalle, ofrecen un aspecto muy texturizado.

3.3.2 Clasificadores por regiones

Este tipo de clasificadores además de utilizar la información espectral de cada píxel, utilizan la información espacial que envuelve la relación entre los píxeles y sus vecinos. Estos clasificadores buscan simular el comportamiento de un foto-intérprete, al reconocer áreas homogéneas en imágenes, basadas en las propiedades espectrales y espaciales de dichas imágenes.

El resultado de una clasificación digital se representa por clases espectrales.

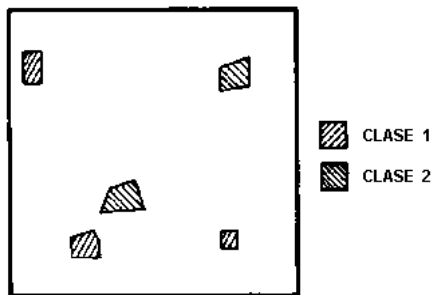


El proceso de clasificación digital transforma un gran número de niveles de gris en cada banda espectral en un pequeño número de clases en una única banda.

Las técnicas de clasificación que pueden ser aplicadas a un sólo canal espectral (banda de la imagen), son conocidas como “clasificaciones unidimensionales” y las técnicas en las que el criterio de decisión depende de la distribución de niveles de gris en varios canales espectrales son definidas como técnicas de “clasificación multiespectral”.

3.3.2.1 Entrenamiento supervisado

Se llama entrenamiento supervisado cuando existen regiones de la imagen en las que se dispone de información que permite la identificación de una clase de interés. Para un entrenamiento supervisado se debe identificar en la imagen el área representativa de cada clase.



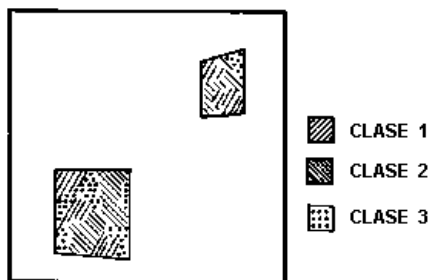
Es importante que el área de entrenamiento sea una muestra homogénea de la respectiva clase, pero al mismo tiempo se deben incluir todas las variabilidades de los niveles de gris.

Es recomendable que se adquiriera más de un área de entrenamiento utilizando el mayor número de información disponible, como trabajos de campo, mapas, entre otros. En la siguiente figura [DPI] se muestra como se deben seleccionar las áreas en el entrenamiento supervisado.

Figura 3.1 Selección de clases en el entrenamiento supervisado

3.3.2.2 Entrenamiento no supervisado

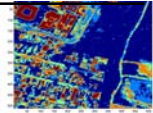
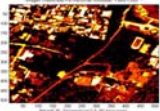
El entrenamiento es denominado “no supervisado” cuando se utilizan algoritmos para reconocer las clases presentes en la imagen. Al definir áreas para el entrenamiento no supervisado, no se



debe preocupar con la homogeneidad de las clases. Las áreas seleccionadas deben ser heterogéneas para asegurar que todas las posibles clases y sus variaciones sean incluidas.

Los píxeles dentro de un área de entrenamiento son sometidos a un algoritmo de agrupamiento ("clustering") que determina el grupo al cual pertenece el dato, en un espacio de dimensión igual al número de bandas presentes.

Figura 3.2 Selección de clases en el entrenamiento no supervisado

**3.3.3****Texturas**

Cuando se trabaja con imágenes de alta resolución se hace imprescindible recurrir otras técnicas que integren SIG (Geografic Information Systems) e información de contexto. En este trabajo veremos cómo se trata de incorporar la textura; es decir, la información espectral de contexto de los píxeles, no individualmente como hacen los métodos de clasificación.

Puede definirse la textura como la variación del contraste entre píxeles vecinos, o, dicho de otra forma, la variabilidad local de la reflectancia dentro de una unidad que a mayor escala se puede considerar como homogénea.

La interpretación y tratamiento de la textura tiene dos etapas: el reconocimiento de las regiones homogéneas que queremos que sirvan de “patrón” y la clasificación de toda la imagen según esas clases. Para aplicar el algoritmo de reconocimiento, Haralick [RMHS73] propone el cálculo de parámetros texturales a partir de matrices de co-ocurrencia de dependencia espacial de los niveles de gris, a las cuales se les aplican parámetros estadísticos (media, correlación, entropía, etc.).

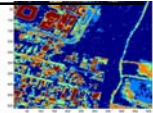
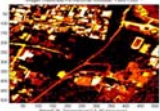
3.4**Antecedentes de los modelos de campos aleatorios de Markov**

Las restricciones de contexto son, en última instancia, necesarias en la interpretación de la información visual. Una escena se entiende en el contexto espacial y visual de los objetos que hay en ella, mientras que los objetos se reconocen en el contexto de las “características de objeto” en una representación de bajo nivel y las características de objeto se extraen en el contexto de los píxeles de la imagen en el nivel más bajo de abstracción.

La teoría de los campos aleatorios de Markov (CAM) proporciona una manera conveniente y constante de modelar entidades dependientes del contexto tales como píxeles de la imagen y otras características espaciales correlacionadas. El uso práctico de los modelos de campos aleatorios de Markov se atribuye en gran parte a la equivalencia entre las distribuciones de probabilidad de los campos aleatorios de Markov y de Gibbs establecidas por Hammersley y Clifford en 1971 [JHMC71] y posteriormente desarrolladas por Besag en 1974 [JB74]. Esto nos permite modelar matemáticamente problemas de visión manejables para el análisis de imagen en el marco bayesiano.

Los modelos de visión estocásticos completos basados en campos aleatorios de Markov están formulados dentro del campo de trabajo bayesiano. La solución óptima de un problema se define como el estimador de probabilidad máxima a posteriori o PMA (MAP), el mejor que se puede obtener a partir de observaciones aleatorias.

La mayoría de los problemas de visión pueden ser considerados como uno de etiquetado usando restricciones, debido al conocimiento a priori y a las observaciones. En este caso, la solución óptima es definida por el etiquetado PMA y se calcula minimizando la energía a posteriori. La probabilidad a posteriori se define usando la regla de Bayes, a partir de un modelo a priori y un modelo de probabilidad.



3.4.1 Sistemas de vecindarios y asociación de tipos de cliques

En general se habla de que las imágenes son universalmente definidas en un enrejado rectangular finito L , de $M \times N$ $L = \{(i,j) : 1 \leq i \leq M, 1 \leq j \leq N\}$

Este modelo estructura a la imagen como una consistencia de K diferentes tipos de regiones, $K=2$ en este caso y la intensidad de cada tipo sigue una distribución normal con un valor de media y varianza específica.

Define $Q = \{q_1, q_2, \dots, q_K\}$, entonces el proceso de región es denotado por $X = \{X_{ij} : X_{ij} \in Q\}$, gobernando el grupo de los píxeles en una imagen en distintas regiones, donde $X_{ij} = q_K$ con $q_K \in Q$ mediante esto el píxel en posición (i,j) pertenece o corresponde al tipo de región q_K .

Con la definición sobre de L , el vecindario de un píxel en la posición (i,j) , denotado por η_{ij} , es usualmente definido de tal manera que:

- $(i,j) \notin \eta_{ij}$ y si
- $(m, n) \in \eta_{ij}$ entonces
- $(i,j) \in \eta_{mn}, \forall (m, n) \in L$

La colección de todos los subconjuntos semejantes de L corresponden a una definición específica del vecindario, entonces es llamado como sistema vecindario η por L o:

$$\eta = \{\eta_{ij} : (i,j) \in L, \eta_{ij} \subseteq L\} \dots\dots\dots \text{ec.(3.1)}$$

Un subconjunto $C \subseteq L$ es llamado clique si C consiste de un simple píxel o un grupo de píxeles que son pares muy distintos de píxeles en los vecindarios de cada otro con respecto a un sistema particular de vecindario, entonces C es usada para denotar el conjunto de todo los cliques semejantes para el par (L, η) .

Es de especial interés para el procesamiento de imágenes, el llamado sistema de vecindarios homogéneos η_c con η_{ij} , el cual esta definido de la siguiente manera:

$$\eta_{ij} = \{(n, m) \in L : 0 < (m - i)^2 + (n - j)^2 = d^2 \leq c\} \dots\dots\dots \text{ec.(3.2)}$$

Si $c = 1$, el vecindario de un píxel simplemente consiste de 4 vecindarios cercanos y η_1 es llamado de primer orden o sistema de vecindario cercano. Para $c = 2$, el vecindario contiene 8 vecindarios cercanos de el píxel central y η_2 es determinado como sistema de vecindarios de 2º orden. En la siguiente figura se muestran algunos de los tipos de vecindarios y cliques de n orden.

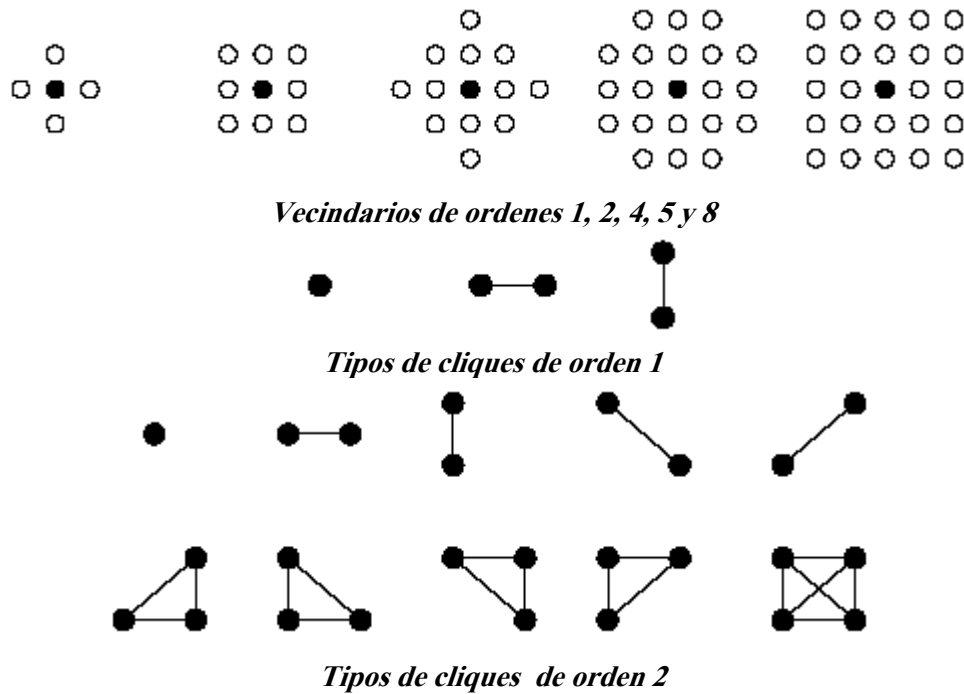
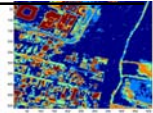


Figura 3.3 Tipos de vecindarios y cliques de n- orden para el modelo de textura

3.4.2 El campo aleatorio de Markov y la distribución de Gibbs

La restricción de continuidad espacial puede ser incorporada en el modelo de región para pretender que el proceso de región X es un campo aleatorio de Markov con respecto al sistema de vecindad.

Por definición el CAM x tiene la propiedad de la distribución condicional de x en un píxel particular, dado los valores de todos los otros píxeles en toda la imagen es igual a la distribución condicional dado solamente a los valores de los vecindarios.

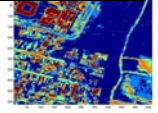
Con $x_{ij}, x_{mn} \in Q$,

$$P(X_{ij} = x_{ij} | X_{mn} = x_{mn}, (m,n) \neq (i,j)) = P(X_{ij} = x_{ij} | X_{mn} = x_{mn}, (m,n) \in \eta_{ij}) \dots\dots\dots ec.(3.3)$$

Esta propiedad es importante, y cuando tratamos con un píxel particular podemos dejar todos los píxeles fuera del vecindario.

Otra propiedad de gran importancia del CAM, es que bajo una condición positiva, el CAM X debe tener una distribución de Gibbs relativa a $\{L, \eta\}$.

Una distribución de Gibbs es una medida de probabilidad del atributo de un píxel en particular, derivado de la propiedad de los vecindarios con la siguiente forma:



$$P(x_{ij} = x) = 1/Z \exp \left\{ -1/T \sum_{c \in C} V_c(x) \right\} \dots\dots\dots ec.(3.4)$$

Donde $x \in Q$, C es el conjunto de cliques relacionado en el píxel (i,j) , $V_c(x)$ es la llamada función potencial asociada con el clique C , T es una constante de control llamada temperatura que por simplicidad puede ponerse a 1 sino se requiere una consideración especial y Z es la constante normalizada definida por:

$$Z = \sum_x \exp \left\{ -1/T \sum_{c \in C} V_c(x) \right\} \dots\dots\dots ec.(3.5)$$

Los CAM normalmente se asumen homogéneos, es decir, que $V_c(x)$ se asume ser no pertinente a la posición del clique en la rejilla L , pero solo depende del tipo y del valor del píxel en el clique afectado.

El CAM usado actualmente para procesamiento de imágenes tiene que ser una estimación discreta; esto es, el CA X consiste solamente de K valores de variables aleatorias discretas $\{X_{ij}\}$ tomando valores en Q .

Esta estimación discreta del CAM es a veces llamado un campo de multinivel logístico (MLL). La clave del campo MLL es que un parámetro positivo es asociado con cada tipo de clique.

3.4.3 Modelos de campos aleatorios de Markov

En este apartado vamos a introducir la definición de algunos modelos útiles de Campos Aleatorios de Markov para modelar características de la imagen tales como regiones y texturas.

Estamos interesados en sus distribuciones de probabilidad condicional y conjunta, y en sus correspondientes funciones de energía [MPRRPA00]. Dependiendo de cómo está definida la función de energía podemos agruparlos en:

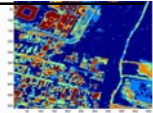
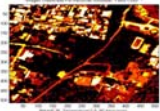
1. Auto modelos.
2. Modelo logístico multinivel.

3.4.4 Auto modelos

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

Entonces la función clique potencial estará definida por:

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{j \in S} V_2(f_i, f_j) \dots\dots\dots ec.(3.6)$$



A la que se le denomina energía de segundo orden porque se calcula para cada clique formada por dos estados. Podemos especificar un campo aleatorio de Gibbs o un campo aleatorio de Markov mediante la selección apropiada de V_1 y V_2 .

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

$$U(f) = \sum_{\{i\}} C_1 f_i G_i(f_i) + \sum_{\{i,j\}} C_2 \beta_{i,j} f_i f_j \dots \text{ec.(3.7)}$$

El modelo anterior recibe el nombre de auto modelo. Los auto modelos se pueden clasificar más a fondo de acuerdo a las suposiciones que se hagan sobre las f_i .

Un auto modelo se dice que es un modelo auto logístico, si las f_i toman valores en el conjunto de etiquetas discreto $L = \{0, 1\}$ o $(L = \{+1, -1\})$.

Cuando el sistema de vecinos N es una rejilla (4-vecinos en una rejilla 2D, o 2-vecinos en una rejilla 1D), entonces el modelo auto logístico se reduce al modelo de Ising [SGDG84]. La correspondiente función de energía es de la forma:

$$U(f) = \sum_{\{i\}} C_1 \alpha_i f_i + \sum_{\{i,j\}} C_2 \beta_{i,j} f_i f_j \dots \text{ec.(3.8)}$$

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

La probabilidad condicional para el modelo auto logístico con $L = \{0, 1\}$ es:

$$P(f_i | f_{N_i}) = \frac{e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}}{\sum_{f_i \in \{0,1\}} e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}} = \frac{e^{\alpha_i f_i + \sum_{j \in N_i} \beta_{i,j} f_i f_j}}{1 + e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}} \dots \text{ec.(3.9)}$$

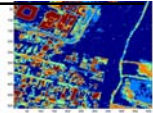
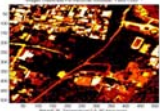
Cuando la distribución es homogénea, tenemos a $\alpha_i = \alpha$ y $\beta_{i,j} = \beta$, sin importar i ni j .

Un auto modelo es un auto modelo binomial si f_i toma valores en $\{0, 1, \dots, M - 1\}$ y cada f_i tiene una distribución condicional binomial con M pruebas y probabilidad de q éxitos.

$$P(f_i | N_i) = \binom{M-1}{f_i} q^{f_i} (1-q)^{M-1-f_i} \dots \text{ec.(3.10)}$$

Donde:

$$q = \frac{e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}}{1 + e^{\alpha_i + \sum_{j \in N_i} \beta_{i,j} f_j}} \dots \text{ec.(3.11)}$$



La correspondiente función de energía es de la forma

$$U(f) = \sum_{\{i\} \in C_1} \ln \binom{M-1}{f_i} - \sum_{\{i\} \in C_1} \alpha_i f_i - \sum_{\{i,j\} \in C_2} \beta_{i,j} f_i f_j \dots \text{ec. (3.12)}$$

Con lo que se reduce al modelo auto logístico cuando M = 1.

Un auto modelo es un auto modelo normal llamado campo aleatorio de Markov gaussiano [RC85], si el conjunto de etiquetas L es \mathfrak{R} y la distribución de probabilidad conjunta es una normal multivariante.

Su función de probabilidad condicional esta dada por la siguiente ecuación:

$$P(f_i | f_{N_i}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} \left[f_i - \mu_i - \sum_{j \in N_i} \beta_{i,j} (f_j - \mu_j) \right]^2} \dots \text{ec. (3.13)}$$

Que es una distribución normal con media condicional.

$$E(f_i | f_{N_i}) = \mu_i - \sum_{j \in N_i} \beta_{i,j} (f_j - \mu_j) \dots \text{ec. (3.14)}$$

Con varianza condicional dada por:

$$\text{var}(f_i | f_{N_i}) = \sigma^2 \dots \text{ec. (3.15)}$$

La probabilidad conjunta es una distribución de Gibbs

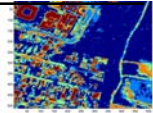
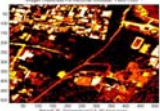
$$p(f) = \frac{\sqrt{\det(B)}}{\sqrt{(2\pi\sigma^2)^m}} e^{-\frac{1}{2\sigma^2} (f - \mu)^T B (f - \mu)} \dots \text{ec. (3.16)}$$

Donde f se considera un vector, μ es el vector $m \times 1$ de las medias condicionales, y $B = [b_{i,j}]$ es la matriz de interacción $m \times n$ cuyos elementos de la diagonal son 1 y el resto de elementos (i, j) es: $-\beta_{i,j}$, es decir $b_{i,j} = \delta_{i,j} - \beta_{i,j}$ con $\beta_{i,j} = 0$ y $(f - \mu)^T$ es la transpuesta de $(f - \mu)$

Por lo tanto, las funciones clique potenciales de un solo estado y de dos estados, para el modelo auto normal son:

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

$$V_1(f_i, f_j) = \beta_{i,j} (f_i - \mu_i) (f_j - \mu_j) / 2\sigma^2 \text{ (dos estados)} \dots \text{ec. (3.17)}$$



Un campo con ruido gaussiano independiente es un caso especial de campo aleatorio de Markov cuya energía de Gibbs consiste únicamente en funciones clique potenciales de un solo estado.

Como las funciones clique potenciales de todos los órdenes superiores son cero, no hay interacción del contexto en el ruido Gaussiano independiente. B está relacionado con la matriz de covarianza Σ ya que $B = \Sigma^{-1}$. La condición necesaria y suficiente para que sea una verdadera función de probabilidad es que B sea simétrica y definida positiva.

3.4.5 Modelo logístico multinivel

El modelo auto logístico se puede generalizar al modelo logístico multinivel (MLL), también llamado proceso de Strauss [DJS77] y modelo generalizado de Ising [SGDG84]. Hay $M (> 2)$ etiquetas discretas en el conjunto de etiquetas $L = \{1, \dots, M\}$.

Un clique potencial depende del tipo c (tamaño, dimensión y orientación posible) del clique y de la configuración local.

$$f_c \triangleq \{f_i / i \in c\} \dots\dots\dots ec.(3.18)$$

Para los cliques que contienen más de un estado ($c > 1$), definimos las funciones clique potenciales MLL como:

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

Donde ζ_c es el potencial para las cliques del tipo c ; para cliques de un solo estado, depende de la etiqueta asignada al estado.

$$V_c(f) = V_c(f_i) = \alpha_I \quad \text{si } f_i = I \in L_d \dots\dots\dots ec.(3.19)$$

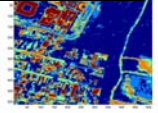
Donde, α_I es el potencial para la etiqueta I .

Asumiremos que un modelo MLL es de segundo orden, si y sólo si a (para cliques de un estado) y b (para cliques de dos estados) son parámetros distintos de cero.

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

$$P(f_i = I / f_{N_i}) = \frac{e^{-\alpha_I - \beta_{n_i}(I)}}{\sum_{I=1}^M e^{-\alpha_I - \beta_{n_i}(I)}} \dots\dots\dots ec.(3.20)$$

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



3.4.6 Ejemplo de segmentación por CAM y el algoritmo de optimización de recocido simulado (RS)

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

Estableciendo la estimación probabilística, la ley condicional de observaciones puede ser condicionada. Esto es denotado por P(X/Y), donde X es la imagen original en niveles de gris y Y la imagen resultante de la segmentación.

Los algoritmos a utilizar para la segmentación, se basan en el siguiente contexto.

Considerando la regla de Bayes

$$p(Y_i / X) = \frac{p(X / Y_i)P(Y_i)}{p(x)} \dots\dots\dots ec.(3.21)$$

Donde:

- x es la imagen original y $Y_i =$ son las clases, con $i = 0, 1, 2,3,4,$ etc.
- $P (Y_i / X)$ representa la probabilidad a posteriori.
- $P (Y_i)$ es la probabilidad a priori de la clase Y_i .
- $P(X / Y_i)$ es la probabilidad conjunta de X dado Y_i .
- $P(X)$ es la probabilidad total de X.

Aproximando las variables conjuntas del numerador de la regla de Bayes mediante las funciones de energía de la figura 3.4, las cuales indican la probabilidad condicional de que un píxel dado pertenezca a la clase Y_i , y empleando el modelo de Potts se calculara la energía a priori de clase, para lo cual se emplearon los siguientes parámetros T=2.5, b=0.35 y N-iteraciones = n= 60.

El estimador MAP consiste en encontrar la configuración, que maximiza la distribución a priori y un proceso equivalente es obtenido minimizando una función de energía, es decir, empleando la aproximación de la regla de Bayes ec. (3.22) para calcular a energía a posteriori ec. (3.23)

$$P(Y_i|X) \approx P(X|Y_i) + P(Y_i) \dots\dots\dots ec.(3.22)$$

$$U(Y_i|X) \approx U(X|Y_i) + U(Y_i) \dots\dots\dots ec.(3.23)$$

Para realizar la segmentación nos apoyaremos en las siguientes funciones de energía:

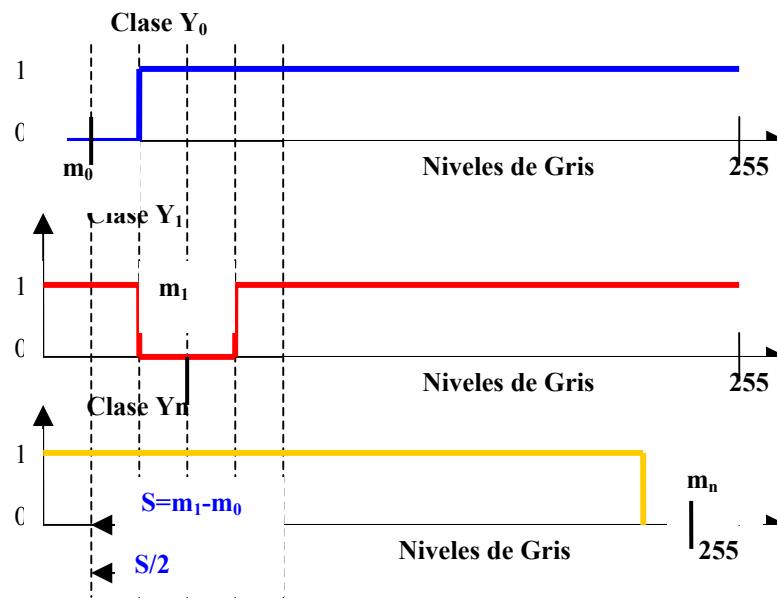
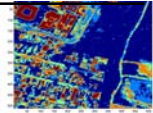
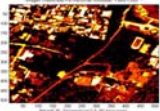


Figura 3.4 Grafica de energías

Para este caso, la segmentación de nuestra imagen será en cinco clases distintas, se tiene que el nivel de gris de un sitio analizado depende únicamente de los valores de gris de sus sitios vecinos, lo anterior en términos estadísticos. Como ya se mencionó anteriormente, los CAM están regidos además por una ley de probabilidad exponencial con temperatura (Ley de Gibbs), la cual representa a la energía global de la imagen.

La energía local de la imagen en cada sitio (site) se rige por el modelo de Potts. El algoritmo utilizado en la segmentación de la imagen se basa en un método de relajación óptimo, denominado recocido simulado (RS) en el cual se analiza y clasifica píxel por píxel de la imagen.

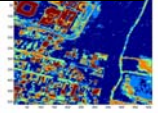
3.4.6.1

Algoritmo basado en CAM y optimizado mediante recocido simulado

Para la inicialización se creó una matriz (imagen) aleatoria $Y^{(0)}$ (figura 3.6), donde cada clase de los píxeles se determinó mediante un generador de números aleatorios con distribución uniforme $\xi \in \Omega$, con un intervalo igual al número de clases.

El proceso consistió en hacer un barrido secuencial de la imagen en cada iteración, para lo cual se aplicó el modelo de Potts y el criterio de Metrópolis el cual se describe a continuación:

En el criterio de Metrópolis se realizan barridos sobre la imagen y en cada site se tiene la posibilidad de efectuar un cambio en el estado del píxel. En el modelo de Potts se tiene una configuración de vecindad, en 8-conectividad, lo que permite obtener la energía local $U(Y)$ del site (energía a priori).



Se escoge un sitio: $r=s^{(n)}$, donde su valor es $Y_r^{(n)}$ y su configuración de vecindad es $V_r^{(n)}$, después se efectuará el sorteo de la variable aleatoria, la cual es la clase a elegir o clase elección, con distribución uniforme.

Por ejemplo, si se tienen las clases 1, 2, 3, 4, 5 (figura 3.5), entonces la clase actual sería 1 (píxel centro) y la clase elección 3 (aleatorio), con lo cual procedemos a hacer un análisis en conectividad 8.

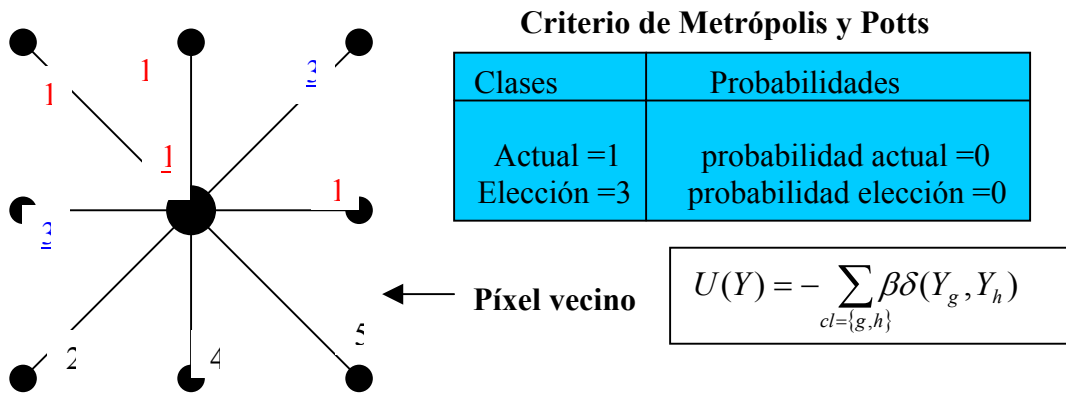


Figura 3.5 Análisis 8-conectividad

Si $X =$ actual, entonces:

La probabilidad actual = Probabilidad actual - β (actual)ec.(3.24)

Y por lo cual en referencia al ejemplo anterior (figura 3.5) tendríamos que: probabilidad actual = -3β , entonces si $X =$ elección la probabilidad elección = probabilidad elección - β (elección).

Por lo tanto sería: probabilidad elección = -2β

Ahora para el cálculo de la variación de la energía local asociada al sitio r , se tiene:

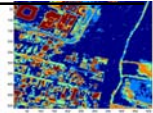
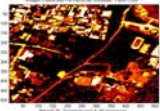
Probabilidad actual = probabilidad actual + potencial [actual] [i] [j]

$U(Y_{actual}|X) \approx U(X|Y_{actual})$ ec.(3.25)

Y para la probabilidad elección = probabilidad elección + potencial [elección] [i] [j]

$U(Y_{elección}|X) \approx U(X|Y_{elección})$ ec.(3.26)

Donde el potencial representa a la energía condicional $U(X/Y)$ e [i] [j] representan a las coordenadas del píxel a analizar.



Para realizar este procedimiento se tendrá que seguir con las siguientes condiciones:

Si la probabilidad elección \leq probabilidad actual, entonces $Y [i] [j] =$ elección, y en caso contrario se realizará el siguiente proceso:

$$p = \exp\left[-\frac{U(Y_{elección}|X) - U(Y_{actual}|X)}{T^{(n)}}\right] \dots\dots\dots ec.(3.27)$$

Es decir $P = \exp -[\text{probabilidad elección} - \text{probabilidad actual}] / T^{(n)}$

Sorteo de una variable aleatoria α , uniformemente distribuida en el intervalo [0,1]

Si $\alpha < P$, entonces $Y [i] [j] =$ elección, así sucesivamente hasta obtener un barrido y después se decrementará la temperatura de la siguiente forma:

$$T(n) = T(n-1) * 0.95 \dots\dots\dots ec.(3.28)$$

Pero para entender mejor el algoritmo se muestra a continuación un diagrama del método con el cual se efectúa la segmentación de una imagen.

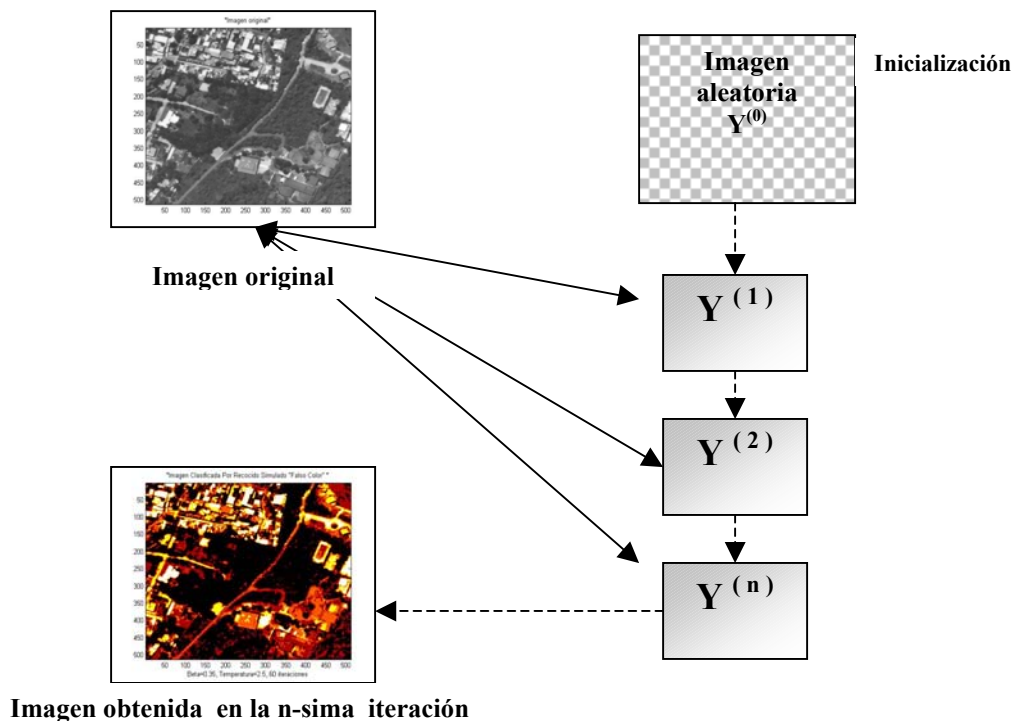
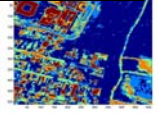


Figura 3.6 Segmentación mediante el algoritmo de optimización RS



3.5 Segmentación bayesiana

Un proceso de segmentación bayesiana puede constar de 2 etapas, la presegmentación y la segmentación.

3.5.1 Presegmentación bayesiana

Para realizar la presegmentación se seguirán los siguientes puntos.

- Sobre la imagen a trabajar, se definen n ventanas que identifiquen a n-regiones típicas: vegetación, caminos, zonas urbanas blancas, zonas urbanas opacas, etc.
- Sobre cada ventana se calcula la media m.
- Se aproximan las probabilidades conjuntas del numerador de la regla de Bayes mediante las funciones mostradas en la figura 3.7, las cuales indican la probabilidad condicional de que un píxel dado pertenezca a la clase Y_n .

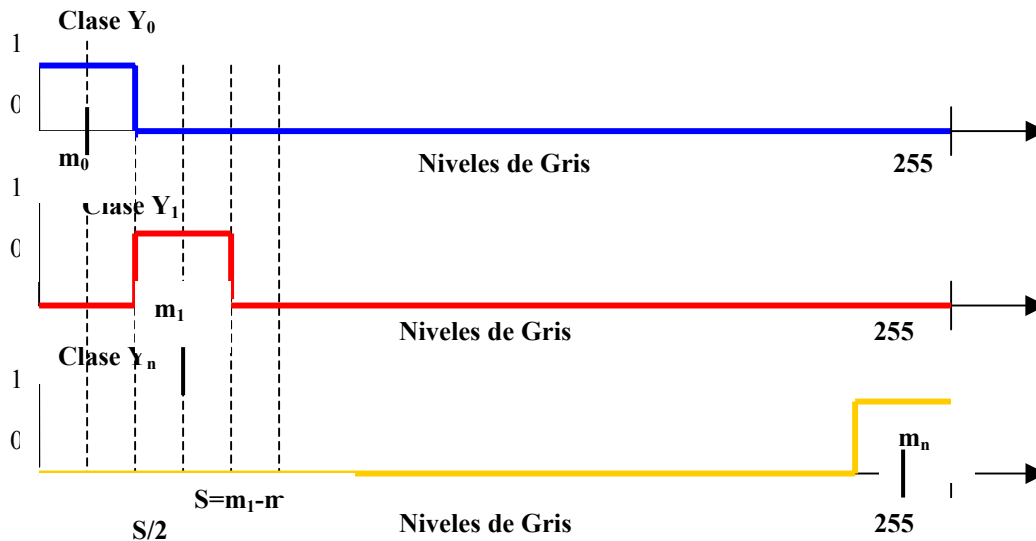
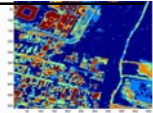
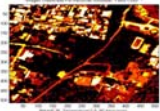


Figura 3.7 Graficas de probabilidad condicional

Después se emplea la función discriminante de Bayes dada por la siguiente ecuación:

$$d_n(x) = p(x | Y_n)p(Y_n) \dots\dots\dots ec.(3.29)$$



Donde:

- x Es la imagen original.
- Y_n Las clases con $n=0,1,2,\dots$
- $p(Y_n)$ Es la probabilidad a priori de la clase Y_n .
- $p(x | Y_n)$ Es la probabilidad conjunta de x dado Y_n y
- $d_n(x) = p(Y_n | x)$ Representa la probabilidad a posteriori.

Como no se conoce la probabilidad a priori $p(Y_n)$, se consideran equiprobables a las probabilidades de ocurrencia de las clases Y_n . Después se analizan cada uno de los píxeles de la imagen x , clasificando a cada píxel de acuerdo con los valores máximos de la probabilidad.

$$\arg \max[p(Y | x)] \dots\dots\dots ec.(3.30)$$

3.5.2 Segmentación bayesiana

En esta etapa, se efectúa nuevamente una segmentación bayesiana, utilizando las probabilidades a priori $p(Y_n)$ actualizadas. Esta actualización se lleva a cabo dividiendo el número de píxeles que pertenecen a una clase entre el total de píxeles de la imagen.

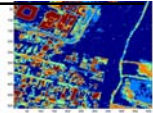
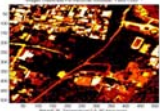
Después de realizar la actualización de probabilidades, se clasifica cada uno de los píxeles de la imagen utilizando la misma función bayesiana de la etapa anterior. Durante el proceso de segmentación, el nivel de gris de cada píxel se proyecta sobre cada una de las funciones para determinar la probabilidad condicional de que el píxel pertenezca a la clase Y_i .

El número de funciones está determinado por el número de clases.

Como se puede observar en la figura 3.7, si el nivel de gris del píxel analizado, esta en el intervalo de $0 \leq Y_0 \leq m_0$, la probabilidad condicional para la clase Y_0 es de 1. Para las demás clases, la probabilidad será cero. Pero si el nivel de gris del píxel analizado está en el intervalo $m_0+s/2$ y $m_1-s/2$ entonces existen 2 valores de probabilidad condicional para este píxel, ya que como se observa en la figura, este píxel puede pertenecer a la clase Y_0 o a la clase Y_1 .

Para obtener la probabilidad condicional de que el píxel analizado pertenezca a la clase Y_0 , se utiliza la siguiente ecuación:

$$Y_0 = \frac{-2 * x}{S} + \frac{2m_0}{S} + \frac{3}{2} \dots\dots\dots ec.(3.31)$$



Donde:

- Y_0 es la probabilidad condicional de la clase Y_0
- x es el píxel analizado.
- m_0 es la media de Y_0
- $S = m_1 - m_0$

De forma similar, para obtener la probabilidad condicional de que el píxel (nivel de gris) analizado pertenezca a la clase Y_1 , se utiliza la ecuación (3.31) para Y_1 y de esta forma el píxel analizado pertenecerá a la clase que tenga una probabilidad condicional mayor.

Para la clase Y_1 , si el nivel de gris del píxel a analizar es menor a $m_1-s/2$ (lado izquierdo de la figura para la clase Y_1) y mayor a $m_1+s/2$ (lado derecho de la figura para la clase Y_1) la probabilidad condicional para Y_1 es de cero.

Si el nivel de gris del píxel analizado es mayor a $m_1-s/2$ y menor a $m_1+s/2$, entonces la probabilidad condicional será de 1. Para las n-regiones, se realiza el mismo procedimiento.

Después de obtener todas las probabilidades condicionales $P(x/Y_n)$, se multiplica cada una de ellas por la probabilidad de ocurrencia, que en la presegmentación se consideran como equiprobables, es decir:

$$p(Y_n) = \frac{1}{n} \dots\dots\dots ec.(3.32)$$

Donde n es el número de clases. En la etapa de segmentación dichas probabilidades son actualizadas, de tal forma que el píxel analizado pertenecerá a la clase cuya probabilidad a posteriori sea mayor.

En las siguientes figuras se muestra el resultado de una segmentación bayesiana para 5 clases de la siguiente imagen. En cambio en el capítulo de resultados se podrán observar más ejemplos de segmentación.

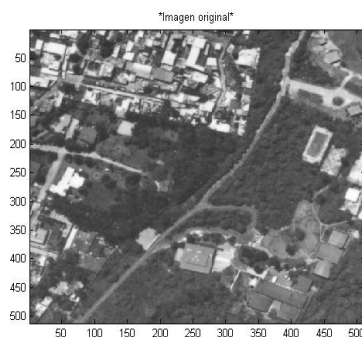


Imagen original

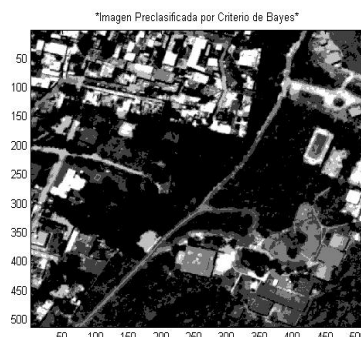


Imagen presegmentada

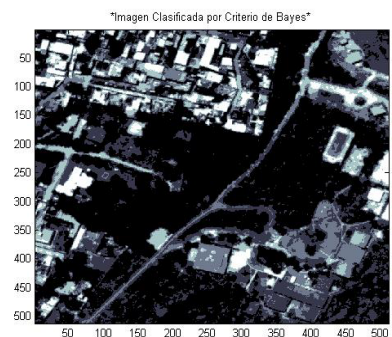


Imagen segmentada

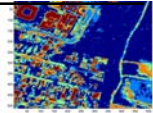
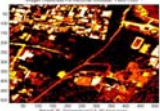


Figura 3.8 Resultado de una segmentación por medio de Bayes

3.5.3 Función discriminante y función óptima de Bayes

La función discriminante de Bayes se utiliza en la clasificación de imágenes [BTZ]. Para poder definirla debemos definir la pérdida condicional promedio o el riesgo condicional promedio como:

$$r_k(x) = \sum_{i=1}^M L_{ik} p(y_i | x) \dots\dots\dots ec.(3.33)$$

Donde L_{ij} es la función de pérdida, costo o penalización por decidir si x pertenece a y_i . La siguiente ecuación muestra el teorema de Bayes.

$$p(y_i|x) = \frac{p(y_i \cap x)}{p(x)} = \frac{p(y_i)p(x|y_i)}{p(y_1)p(x|y_1) + p(y_2)p(x|y_2) + \dots + p(y_n)p(x|y_n)} =$$

$$p(y_i | x) = \frac{p(y_i)p(x | y_i)}{p(x)} \dots\dots\dots ec.(3.34)$$

Si sustituimos el teorema de Bayes en la ecuación (3.33), obtendremos:

$$r_k(x) = \frac{1}{p(x)} \sum_{i=1}^M L_{ik} p(x | y_i) p(y_i) \dots\dots\dots ec.(3.35)$$

Como $p(x)$ en la ecuación anterior es común a toda $r_j(x), j = 1, \dots, M$ podemos eliminarla de la pérdida condicional promedio y solo buscar el mínimo, es decir:

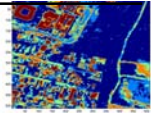
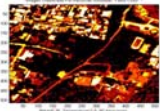
$$\min_k r_k(x) = \min_k \sum_{i=1}^M L_{ik} p(x | y_i) p(y_i) \dots\dots\dots ec.(3.36)$$

Para tomar la mejor decisión, podemos decir que:

$$d_k(x) = -r_k(x) \dots\dots\dots ec.(3.37)$$

La cual es llamada función discriminante de Bayes.

Para el caso en el que se tuvieran dos clases, el planteamiento se realizaría de la siguiente manera:



$$r_1(x) = L_{11}p(x|y_1)p(y_1) + L_{21}p(x|y_2)p(y_2) \quad y$$

$$r_2(x) = L_{12}p(x|y_2)p(y_2) + L_{22}p(x|y_1)p(y_1)$$

Se dice que $x \in y_1$ cuando:

$$L_{11}p(x|y_1)p(y_1) + L_{21}p(x|y_2)p(y_2) < L_{12}p(x|y_1)p(y_1) + L_{22}p(x|y_2)p(y_2) \dots\dots\dots ec.(3.38)$$

De otra forma:

$$\frac{p(x|y_1)}{p(x|y_2)} > \frac{(L_{21} - L_{22})p(y_2)}{(L_{12} - L_{11})p(y_1)} \dots\dots\dots ec.(3.39)$$

A la que se le conoce como la regla de máxima verosimilitud.

Ahora si generalizamos varias clases, tendremos:

$$L_{ki} = \frac{p(x|y_k)}{p(x|y_i)} > \frac{p(y_i)}{p(y_k)} \dots\dots\dots ec.(3.40)$$

Donde el criterio de asignación es el siguiente:

$$x \in \omega_k \quad \text{Si } L_{ki} > 1 \quad \acute{o} \\ p(x|y_k)p(y_k) > p(x|y_i)p(y_i) \dots\dots\dots ec.(3.41)$$

Entonces la función discriminante estará representada de la siguiente forma:

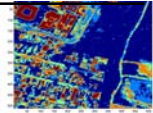
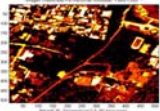
$$d_k(x) = p(x|y_k)p(y_k) \dots\dots\dots ec.(3.42)$$

Otra forma alternativa podría ser la siguiente:

$$d'_k(x) = \log p(x|y_k) + \log p(y_k) \dots\dots\dots ec.(3.43)$$

Cuando se aplica esta función discriminante a una función de densidad normal multi-variada para una clase de m-patrones se tiene que:

$$p(x | \omega_k) = \frac{1}{(2\pi)^{n/2} |C_k|^{1/2}} \exp \left[-\frac{1}{2} (x - m_k)^T C_k^{-1} (x - m_k) \right] \dots\dots\dots ec.(3.44)$$



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

La función discriminante estará dada por:

$$d_k''(x) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|C_k| - \frac{1}{2} (x - m_k)^T C_k^{-1} (x - m_k) + \log p(y_k) \dots \text{ec.}(3.45)$$

Si el primer término del lado derecho es el mismo para toda k, este término puede ser eliminado. Por lo tanto, la función discriminante óptima se reduce a:

$$d_k'''(x) = -\frac{1}{2} (x - m_k)^T C_k^{-1} (x - m_k) + \log p(y_k) - \frac{1}{2} \log|C_k| \dots \text{ec.}(3.46)$$

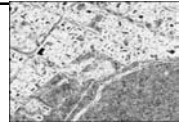
Esta es una función discriminante cuadrática y puede expresarse en forma más compacta como:

$$d_k^{IV}(x) = -\frac{1}{2} r^2 + f(k) \quad \forall x \in \omega_k \dots \text{ec.}(3.47)$$

Donde:

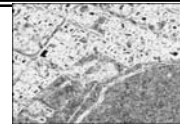
$$r^2 = (x - m_k)^T C_k^{-1} (x - m_k) \text{ es la distancia de Mahalanobis y } f(k) = \log p(y_k) - \frac{1}{2} \log|C_k|$$

Nota: dⁿ significa que se deriva de un función discriminante, mas no es una derivada.



4 *El modelo textural*





4.0

Introducción

Las texturas son empleadas para distinguir objetos y regiones en una imagen, así también, para inferir la orientación de superficies, las formas en escenas de tres dimensiones y para diferenciar entre varias regiones o clases.

Para nuestro contexto definiremos a la textura como una variación del contraste entre píxeles; es decir; la variación local de reflectancia dentro de una unidad que a mayor escala se puede considerar como homogénea.

La textura se puede describir como el compuesto de elementos de los primitivos de la textura, así para la textura, el término texel tiene gran importancia como elemento fundamental de la misma debido a que un texel es un primitivo visual con ciertas características invariantes que ocurre en varias ocasiones en diversas posiciones, deformaciones, y orientaciones dentro de un área dada.

Los primitivos de la textura pueden ser píxeles o agregados de píxeles tales como segmentos o regiones pequeñas [OSB].

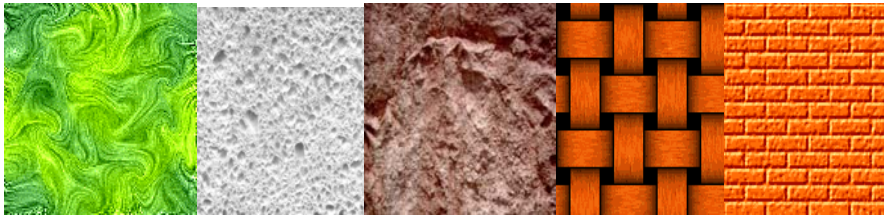


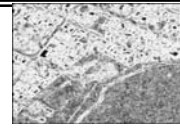
Figura 4.1
Ejemplos de texturas

Las texturas se caracterizan por su uniformidad, densidad, grosor, rugosidad, intensidad y regularidad de medidas discretas del tono y de sus relaciones espaciales, y para su análisis existen varias técnicas, entre las cuales se encuentran las técnicas espectrales, estructurales, basadas en modelos y estadísticas (sobre las cuales nos enfocaremos más detenidamente, así como en sus modelos).

4.1 Técnicas espectrales

Las técnicas espectrales se basan en las propiedades del espectro de Fourier, las cuales son utilizadas en la detección de periodicidad global de una imagen mediante la detección de picos estrechos de alta energía del espectro.

Estas técnicas tiene un desempeño bueno en ciertas aplicaciones, pero su implementación esta restringida por la poca información que se puede extraer cuando las texturas son heterogéneas o de carácter aleatorio.



La transformada de Fourier se puede utilizar para el análisis de la textura y la transformada de Fourier de $f(x, y)$ de la imagen es $F(u, v)$ su módulo será:

$$P(u, v) = |F(u, v)|^2 \dots\dots\dots ec.(4.1)$$

Si se utilizan las coordenadas polares, la distribución puede ser obtenida de la forma:

$$P(r) = 2 \sum_{\theta=0}^{\pi} P(r, \theta) \text{ y } P(\theta) = 2 \sum_{r=0}^{n/2} P(r, \theta) \dots\dots\dots ec.(4.2)$$

4.2 Técnicas estructurales

En cuanto a las técnicas estructurales, su desarrollo se basa en la obtención de patrones o primitivos de la textura mas reglas de organización, para generar una descripción de como se agrupan entre sí para formar la textura.

Los modelos estructurales ven a las primitivas como formación de un patrón de repetición y describen tales patrones en términos de reglas, en las cuales se describe cómo generar dichos patrones aplicando reglas que describan a un número pequeño de símbolos. No hay regla única para una textura dada, existen incontables opciones para las reglas y los símbolos.

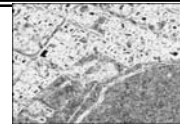
4.3 Técnicas estadísticas

Los modelos estadísticos describen a la textura mediante reglas estadísticas, las cuales gobiernan la distribución y la relación de los niveles de gris.

El análisis estadístico de la textura implica analizar la distribución de ciertos valores de las características (nivel de gris, media, desviación, dispersión, entropía, etc.) para cada píxel de la imagen; así, acordando el número de puntos se define la estadística de una textura en primera, segunda y la de orden superior.

Mediante el uso del histograma y el empleo de estas técnicas se pueden obtener medidas que nos proporcionen características de textura tales como suavidad, granulosidad y rugosidad.

Debido a que el histograma proporciona información sobre los niveles de gris de una región, pero no indica la distribución en el espacio, es mejor la aplicación de la matriz de co-ocurrencia GLCM (estadísticas de 2º orden), ya que esta valora (mide) la cantidad de píxeles que se encuentran a cierta distancia y que tienen el mismo valor o nivel de gris.



4.3.1 Estadísticas de primer orden

Si el histograma de la imagen es dividido por el número total de píxeles, el resultado representa la probabilidad de que cierto nivel gris aparezca en la imagen. Algunas características pueden ser estimadas de la siguiente manera:

- Promedio $\mu = \frac{1}{NM} \sum_{i=1}^n h(i)$ ec.(4.3)

- Tercer momento o coeficiente de asimetría $\mu^3 = \frac{1}{\sigma^3} \sum_{i=1}^n (i - \mu)^3 h(i)$ ec.(4.4)

- Entropía $-\sum_{i=1}^n h(i) \log h(i)$ ec.(4.5)

4.3.2 Estadística de segundo orden (G LCM)

Las estadísticas de 2º orden o momentos estadísticos son empleados en el cálculo de la extracción de información contenida en la matriz de co-ocurrencia de una imagen.

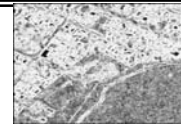
En la matriz de co-ocurrencia, cada elemento representa la probabilidad de ocurrencia de dos valores de la escala de gris separados por una distancia d y un ángulo θ en determinada dirección, donde el ángulo puede ser de 0°, 45°, 90° y 135°.

Los pares co-ocurrentes orientados a 0° y 180° generan una matriz simétrica, de igual forma los pares orientados a 45°, 135°, 90° y 270° también forman matrices simétricas.

Aunque en la clasificación de texturas los elementos individuales de la matriz casi no se utilizan, la matriz de co-ocurrencia es muy importante debido a que a partir de esta se pueden extraer ciertas características, que están determinadas por los descriptores de Haralick, entre los cuales se encuentran la entropía, energía, contraste, homogeneidad, etc.

Para construir una matriz de co-ocurrencia se debe de tomar en cuenta los siguientes puntos:

- **Condición.-** Es la regla que debe cumplir un determinado píxel para poder ser cuantificado con relación a un vecino; generalmente esta condición es la ubicación del píxel de determinada intensidad que se encuentra en la dirección del vecino.
- **Dirección.-** Se refiere a la orientación espacial en la cual se evalúa la condición, como casos típicos se utilizan direcciones como 0°, 45°, 90° y 135°.
- **Distancia.-** Esta condición se refiere al número de píxeles a los cuales se encuentran el par de píxeles evaluados en la condición, las distancias típicas son 1, 3 y 5 píxeles.



Para un mejor entendimiento de la construcción de la matriz de co-ocurrencia se realizará el siguiente ejemplo: Donde se tiene una matriz (imagen) de 4x4 ,con 4 niveles de gris representados por los números 0,1,2 y3.

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

Para formar la matriz de co-ocurrencia con parámetros $d=1$ y $\theta=0^\circ$, inicialmente obtenemos los pares de niveles de gris mediante una matriz auxiliar, el primer par de niveles de gris de esta matriz estará formado por el elemento del renglón 1 columna 1 de la imagen, en este caso el primer elemento es el cero.

El segundo elemento será el que está a una $d=1$ en dirección horizontal ya que $\theta=0^\circ$. Ahora siguiendo la flecha, este elemento también es cero, de esta manera el primer par (i,j) a tomar es $(0,0)$

después se toma como referencia al segundo elemento y se busca al elemento que cumpla con esta condición para que se forme el segundo par que en este caso es $(0,1)$, y el tercero será $(1,1)$, de esta manera se realiza la construcción renglón por renglón.

Tabla 4.1 Matriz 4x4

Una vez que se han obtenido todos los elementos del primer renglón se obtienen los elementos del segundo renglón de la misma manera, hasta obtener todos los pares (i,j) . Entonces la matriz auxiliar estará dada por la tabla 4.2.

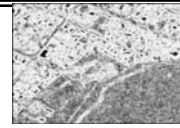
	i	j
Comb.	→	
1	0	0
2	0	1
3	1	1
4	0	0
5	0	1
6	1	1
7	0	2
8	2	2
9	2	2
10	2	2
11	2	3
12	3	3

Una vez que se ha creado esta tabla se procede a llenar la matriz de co-ocurrencia, donde se indicará la frecuencia con que ocurrieron los pares de niveles de gris, es decir; el par $(0,0)$ se repite 2 veces en la tabla 4.2, por lo tanto en $(i=0,j=0)$ de la matriz se coloca el número de repeticiones, en este caso 2, y así sucesivamente como se muestra en la tabla 4.3.

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

Tabla 4.3 Matriz de co-ocurrencia

Tabla 4.2 Matriz auxiliar



Para normalizar la matriz de co-ocurrencia se divide a cada uno de sus elementos entre el número de combinaciones resultantes de la matriz original (tabla 4.2), que en este caso es de 12, por lo tanto, la matriz resultante será:

i/j	0	1	2	3
0	0.1666	0.1666	0.0833	0
1	0	0.1666	0	0
2	0	0	0.25	0.0833
3	0	0	0	0.0833

Para obtener la matriz de co-ocurrencia simétrica, se obtienen los pares de niveles de gris para $d=1, \theta=0$ y $d=1, \theta=180^\circ$, lo que equivale a tomar las probabilidades anteriores más las probabilidades de ocurrencia de las parejas anteriores (en la dirección de la flecha amarilla) o tomar los pares de niveles de gris como (j,i) en lugar de tomarlos como (i,j).

Tabla 4.4 Matriz de co-ocurrencia normalizada

Por lo tanto, la matriz auxiliar completa estará dada por la tabla 4.5:

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

Tabla 4.5 Matriz de co-ocurrencia simétrica

De acuerdo a la frecuencia con la que se repiten los pares de niveles de gris en la tabla 4.6, la matriz de co-ocurrencia es la siguiente:

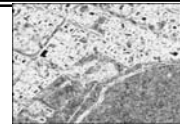
i/j	0	1	2	3
0	4	2	1	0
1	2	4	0	0
2	1	0	6	1
3	0	0	1	2

Tabla 4.7 Matriz de co-ocurrencia

Para normalizar la matriz de co-ocurrencia, dividimos a cada uno de sus elementos entre el número de combinaciones resultantes de la matriz original (Tabla 4.6) que ahora es 24, es decir

	i	j	i	j
	→		←	
1	0	0	13	0
2	0	1	14	1
3	1	1	15	1
4	0	0	16	0
5	0	1	17	1
6	1	1	18	1
7	0	2	19	2
8	2	2	20	2
9	2	2	21	2
10	2	2	22	2
11	2	3	23	3
12	3	3	24	3

Tabla 4.6 Matriz Auxiliar



i/j	0	1	2	3
0	0.1666	0.0833	0.0416	0
1	0.0833	0.1666	0	0
2	0.0416	0	0.25	0.0416
3	0	0	0.0416	0.0833

Tabla 4.8 Matriz de co-ocurrencia normalizada

4.4 Descriptores de textura

Los descriptores de textura nos sirven para extraer cierta información o características de la textura de una imagen, contenida en la matriz de co-ocurrencia [RMHS73]. Dichos descriptores muestran información que es muy útil para la obtención de características texturales, las cuales pueden complementarse mutuamente, ya que cada descriptor muestra características diferentes en una determinada imagen.

Los descriptores de textura se clasifican en 4 clases, las cuales se explican a continuación:

- **Clase 1:** Manifiestan los rasgos visuales de las texturas: Homogeneidad, contraste y correlación
- **Clase 2:** Se basan en estadísticas: Varianza, diferencia inversa, suma de promedios, suma de varianzas y diferencia de varianzas.
- **Clase 3:** Se basan en la teoría de la información: entropía, suma de entropías y diferencias de entropías.
- **Clase 4:** Se basan en las medidas de la información de correlación: coeficiente de correlación máxima.

Algunos descriptores de textura son detallados a continuación.

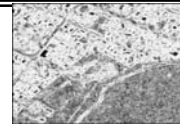
Correlación: La correlación de la textura es una medida de probabilidad que mide la relación entre las diferentes intensidades de los colores o tonos, y esta representada por la siguiente ecuación:

$$C_{rr} = \frac{1}{\sigma_i \sigma_j} \left| \sum_{i=1}^N \sum_{j=1}^N (i-\mu_x)(j-\mu_y) C_{ij} \right| \dots\dots\dots ec.(4.6)$$

Donde:

$$\mu_x = \sum_{i=1}^N \sum_{j=1}^N i C_{ij} \quad y \quad \mu_y = \sum_{i=1}^N \sum_{j=1}^N j C_{ij} ,$$

$$\sigma_i = \sum_{i=1}^N \sum_{j=1}^N (i-\mu_i)^2 C_{ij} \quad y \quad \sigma_j = \sum_{i=1}^N \sum_{j=1}^N (j-\mu_j)^2 C_{ij}$$



Debido al carácter estadístico que posee este descriptor su valor está ligado al valor de sus medidas estadísticas y éste aumenta cuando la distancia de los elementos a la media es grande y cuando la varianza es baja.

Contraste: El contraste de una textura proporciona información acerca de las variaciones bruscas de color en la imagen y esta dado por la siguiente ecuación:

$$\text{Contraste} = \sum_{i=1}^N \sum_{j=1}^N (i-j)^2 C_{ij} \dots\dots\dots \text{ec.}(4.7)$$

Donde N representa los niveles de gris cuantizados.

La ecuación anterior muestra que el valor del contraste aumentará si existen más elementos de la matriz de co-ocurrencia alejados de la diagonal principal, por lo cual, el efecto de las componentes de la diagonal principal será prácticamente despreciado, es decir, que en una textura de características suaves y uniformes su contraste será bajo y en caso contrario su valor será alto.

Homogeneidad local: Este descriptor proporciona información sobre la regularidad local de la textura y está representado por la siguiente ecuación:

$$\text{Homol} = \sum_{i=1}^N \sum_{j=1}^N C_{ij} / (1 + (i-j)^2) \dots\dots\dots \text{ec.}(4.8)$$

Este descriptor aumenta cuando la distancia i-j es mínima.

Direccionalidad: Este descriptor extrae información sobre patrones de dirección en la textura y esta representado por la siguiente ecuación:

$$\text{Dir} = \sum_{i=1}^N C_{ij} \dots\dots\dots \text{ec.}(4.9)$$

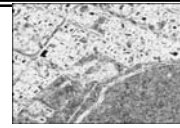
A este descriptor también se le conoce como descriptor de probabilidad de diferencia de niveles de gris; su valor aumentará si en la matriz de co-ocurrencia se notan picos de intensidad del patrón de ésta. Este descriptor sólo analiza la diagonal principal de la matriz de co-ocurrencia, debido a que sólo los elementos de colores iguales descansan sobre la diagonal principal.

Entropía: La entropía es una medida de la aleatoriedad contenida en la matriz de co-ocurrencia y se obtiene mediante la siguiente ecuación:

$$\text{Entropía} = - \sum_{i=1}^N \sum_{j=1}^N C_{ij} \log C_{ij} \dots\dots\dots \text{ec.}(4.10)$$

Donde Cij es cada elemento de la matriz y n es el número de niveles de gris.

El valor más alto de la entropía se obtiene cuando todos los elementos de la matriz son iguales.



En el caso de una imagen digital de textura, la información se encuentra principalmente concentrada cerca de la diagonal principal de la matriz de co-ocurrencia [AJJ02], a medida que se tengan mayores probabilidades de ocurrencia el valor de la entropía desciende debido a que la textura es más uniforme.

Momento inverso de diferencia de orden K: Este momento refleja la definición y agrupación de los elementos de la matriz de co-ocurrencia y se define por la siguiente ecuación:

$$\text{Minv} = \frac{\sum_{i=1}^N \sum_{j=1}^N C_{ij}}{\sum_{i \neq j} |i-j|^k} \dots \text{ec.}(4.11)$$

Este descriptor proporciona información sobre que tan compacta esta la distribución de la matriz, dando la idea de que la textura posee una variación suave de colores.

Máxima probabilidad: Este descriptor suministra información sobre la respuesta más fuerte de un par de píxeles en la matriz de co-ocurrencia y está dado por:

$$\text{Maxprob} = \max_{ij} (C_{ij}) \dots \text{ec.}(4.12)$$

Este descriptor es útil para percibir la mayor ocurrencia o la predominación de una tonalidad en la textura.

Cluster shade: Este descriptor aumenta su valor a medida que la media de los valores está más cercana al valor mínimo y tiene la tendencia a oscilar entre valores positivos y negativos, debido a que la influencia de la media no siempre se compara con el valor de los elementos alejados, y esta definido por la siguiente ecuación:

$$\text{Clush} = \sum_{i=1}^N \sum_{j=1}^N (i + j - \mu_x - \mu_y)^3 C_{ij} \dots \text{ec.}(4.13)$$

Varianza: La varianza esta definida por la siguiente ecuación:

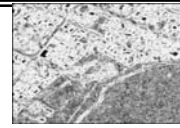
$$\text{Var} = \sum_{i=1}^N \sum_{j=1}^N (i - \mu)^2 C_{ij} \dots \text{ec.}(4.14)$$

La varianza indica la dispersión que tiene un conjunto de muestras con respecto a un valor medio, mientras los valores estén más alejados de éste, mayor será la varianza.

El descriptor de varianza tendrá un mayor valor para aquellas texturas en las cuales la matriz de co-ocurrencia sea de forma alargada o dispersa con respecto del máximo pico en la diagonal principal.

Cluster prominente: Este descriptor está dado por la siguiente ecuación:

$$\text{Cluprom} = \sum_{i=1}^N \sum_{j=1}^N (i + j - \mu_x - \mu_y)^4 C_{ij} \dots \text{ec.}(4.15)$$



Uniformidad: El descriptor de uniformidad está dado por la siguiente ecuación:

$$\text{Unif} = \sqrt{\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N C_{ij}^2} \dots\dots\dots \text{ec.}(4.16)$$

Cuando todos los valores de C_{ij} son semejantes, el valor de la uniformidad es menor y por lo contrario, si ocurre que en la diagonal principal se dan mayores picos de intensidad, el valor del descriptor aumentará. La propiedad de uniformidad da una idea de la suavidad de la textura, y esto se ve reflejado en la ubicación de sus probabilidades en la matriz de co-ocurrencia.

Mediante los descriptores mostrados anteriormente se obtuvieron algunos resultados para 2 imágenes diferentes con el objeto de detallar ciertas áreas texturales (dichos resultados serán mostrados detalladamente en el capítulo de resultados), los cuales nos servirán posteriormente en el análisis principal de este trabajo.

4.5 Modelo textural empleado

4.5.1 Algoritmo rápido para el cálculo de la matriz de co-ocurrencia

Cuando se emplea la matriz de co-ocurrencia para determinar ciertas características de textura en una imagen se desperdicia mucho tiempo calculando el resultado, por lo cual se han desarrollado algoritmos más rápidos para el cálculo de esta matriz. Para obtener la matriz de co-ocurrencia en este trabajo se basa en el método desarrollado por David A. Clausi y Yongping Zhao [DAC] llamado “Rapid Determination of Co-ocurrence Textura Features”.

Para obtener la matriz de co-ocurrencia por medio de este método se emplea una ventana deslizante, la cual comienza en la esquina izquierda superior de la imagen. Las probabilidades de co-ocurrencia se determinan para esta ventana y entonces se calculan los descriptores de esta matriz. Después de que se calculen los descriptores, la ventana se desliza una columna hacia la derecha, en este punto, la mayoría de las probabilidades que co-ocurren siguen siendo iguales, excepto por las probabilidades que se forman con la nueva columna que ha sido incluida a la ventana y por la columna que ha salido de ella. Por lo tanto, a la matriz de co-ocurrencia original se le suman las probabilidades que se forman con la columna que ha sido incluida y se le restan las probabilidades que se formaban con la columna que ha salido de la ventana.

Cuando se llega al extremo derecho de la matriz la ventana se desliza una fila hacia abajo. Para las filas pares, la ventana se desliza de izquierda a derecha y para las filas impares de derecha a izquierda. La ventana se mueve eficientemente en este patrón del zig-zag hasta que se cubre la imagen entera, como se muestra en la figura 4.2.

Este método fue empleado en este trabajo debido a los buenos resultados mostrados en trabajos anteriores, los cuales demuestran su eficacia y buen desempeño para el cálculo de la matriz de co-ocurrencia y la obtención de ciertas características texturales (descriptores de textura).

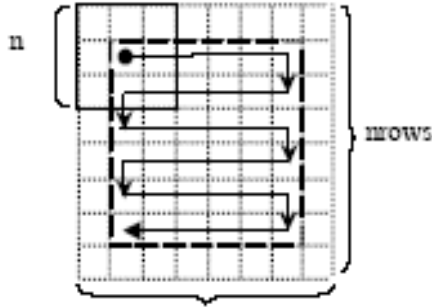
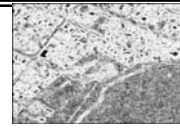


Figura 4.2 Deslizamiento de ventana en zig-zag para determinar las características de textura en una imagen.

Para entender mejor este método se realizará el siguiente ejemplo, tomando como referencia una imagen representada por una matriz de 7x8 (la tabla 4.9) y se realizará un análisis del método:

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

Tabla 4.9 Matriz de 7x8 (imagen ejemplo)

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

i/j	0	1	2	3
0	6	3	0	0
1	3	2	5	3
2	0	5	6	2
3	0	3	2	0

Entonces nuestra matriz auxiliar de co-ocurrencia para una $d=1$, $\theta=0^\circ$ y $\theta=180^\circ$ estará dada por la tabla 4.10.

Ahora la matriz de co-ocurrencia sin normalizar estará representada por la tabla 4.11.

Tabla 4.11 Matriz de co-ocurrencia sin normalizar

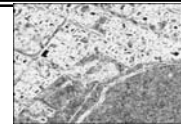
0	0	0	1	1	3	2	2
0	0	1	2	2	2	1	3
0	1	2	2	1	1	2	1
3	2	1	2	2	3	3	2
2	3	1	3	1	1	1	3
2	2	3	1	1	2	2	1
0	2	2	2	2	2	3	2

Tabla 4.10 Matriz auxiliar de co-ocurrencia

Deslizando la ventana una columna hacia la derecha obtenemos la tabla 4.12.

Para ya no realizar el cálculo de la matriz de co-ocurrencia que se forma con esta ventana, simplemente tomamos la matriz de co-ocurrencia anterior y le restamos las ocurrencias que se forman con las columnas 1 y 2, y le sumamos las probabilidades que se forman con las columnas 5 y 6, es decir:

Tabla 4.12 Deslizamiento de la matriz



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

Como se puede observar, para el par de píxeles (0,0) existen 4 probabilidades de ocurrencia, uno para el par (0,1) y (1,0), dos para el par (2,3) y (3,2).

Tabla 4.13 Matriz con probabilidades de las columnas 1 y 2

Entonces se restan estas probabilidades a la matriz de ocurrencia ya calculada y se obtiene la siguiente matriz (tabla 4.14):

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

Tabla 4.14 Matriz de ocurrencias restadas

i/j	0	1	2	3
0	2	2	0	0
1	2	2	5	3
2	0	5	6	0
3	0	3	0	0

Tabla 4.15 Matriz con probabilidades de las columnas 5 y 6

Como se observa para el par (1,3) y para el par (3,1) existe una probabilidad de co-ocurrencia respectivamente. Para el par (1,1) existen 4, para los pares (2,3) y (3,2) también existe una probabilidad de co-ocurrencia. Entonces sumamos estas nuevas probabilidades a la matriz anterior y de esta manera obtenemos una nueva matriz de co-ocurrencia (tabla 4.16).

i/j	0	1	2	3
0	2	2	0	0
1	2	6	5	4
2	0	5	8	1
3	0	4	1	0

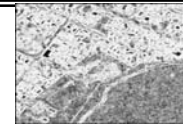
Este procedimiento se realiza hasta obtener un barrido por completo al extremo derecho de la tabla, después la ventana se deslizada una fila hacia abajo y las probabilidades de co-ocurrencia son calculadas de la misma manera. En las siguientes tablas se observa el barrido que realiza la ventana.

Tabla 4.16 Matriz de ocurrencia con probabilidades sumadas

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

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

Tabla 4.17 y 4.18 Barrido de la ventana sobre la matriz de co-ocurrencia



i/j	0	1	2	3
0	0	0	0	0
1	0	8	4	4
2	0	4	8	3
3	0	4	3	2

Para este ejemplo, la matriz resultante obtenida al llegar al extremo derecho de la imagen estará dada por la matriz de la tabla 4.19.

Debido a que la ventana se deslizó un renglón abajo, ahora los términos que se restan a la matriz de co-ocurrencia serán las probabilidades de las parejas que forman el renglón que fue excluido de la tabla 4.18, es decir:

1	1	3	2	2
---	---	---	---	---

Tabla 4.19 Matriz resultante del primer barrido

Y sus parejas están representadas en la tabla 4.20.

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

Entonces, la matriz anterior menos las probabilidades anteriores estará dada por la tabla 4.21.

Tabla 4.20 Parejas resultantes del renglón excluido de la Tabla 5.18

i/j	0	1	2	3
0	0	0	0	0
1	0	6	4	3
2	0	4	6	2
3	0	3	2	2

Tabla 4.21

Por otro lado también se suman nuevas probabilidades que se obtienen del renglón que se ha incluido en la ventana que se desliza, en este caso el renglón es:

1	1	2	2	1
---	---	---	---	---

Los pares de niveles de gris están dados por la tabla 4.22.

i	j	j	i
1	1	1	1
1	2	2	1
2	2	2	2
2	1	1	2

Entonces la matriz de co-ocurrencia resultante estará dada por la tabla 4.23.

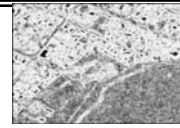
Este proceso de cálculo de la matriz de co-ocurrencia sigue el mismo procedimiento (en zig-zag) hasta completar la matriz de co-ocurrencia.

i/j	0	1	2	3
0	0	0	0	0
1	0	8	6	3
2	0	6	8	2
3	0	3	2	2

Tabla 4.22 Pares de los niveles de gris

Tabla 4.23 Matriz de co-ocurrencia resultante (2º barrido)

Este método reduce el tiempo para el cálculo de la matriz co-ocurrencia en imágenes aéreas de 512*512 píxeles, pero para agilizar este método antes de calcular la matriz se redujo el número de niveles de gris de la imagen, con lo cual se obtiene un buen resultado y tiempo de respuesta mucho mejor sin afectar demasiado los resultados deseados.



Para este trabajo, las imágenes utilizadas fueron reducidas a 32 niveles de gris y después se cálculo su matriz de co-ocurrencia normalizada para cada una de las ventanas utilizadas y se obtuvieron los descriptores de textura para cada una de las imágenes.

En la siguiente figura se muestran algunos ejemplos de los descriptores de textura obtenidos mediante el cálculo de la matriz de co-ocurrencia, no obstante en el capítulo 7 se verán más ampliamente los resultados obtenidos mediante este método.

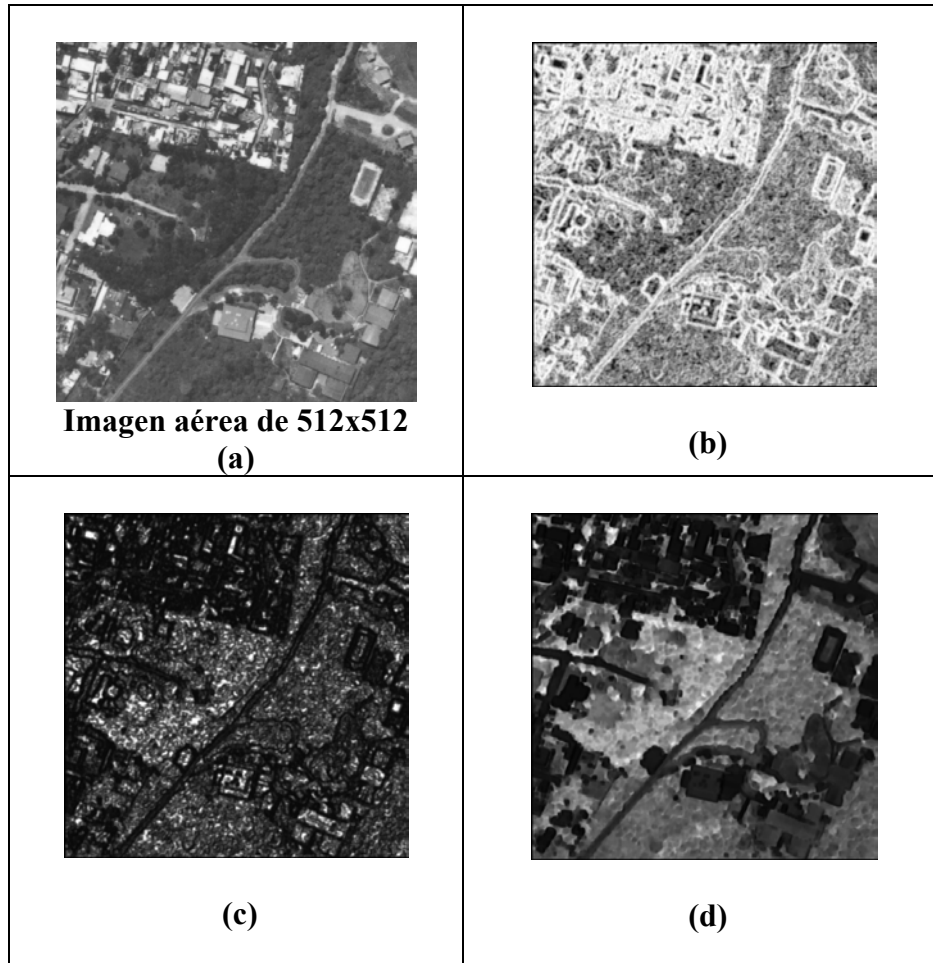
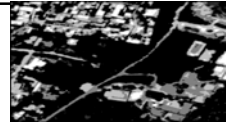


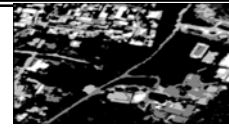
Figura 4.3 Ejemplo de descriptores obtenidos mediante la matriz de co-ocurrencia a una imagen de fotografía aérea de 512 x 512 píxeles con 256 niveles de gris, donde:

- (a) Sección de una imagen aérea del DF. de 512x512 y 256 niveles de gris (Imagen original)
- (b) Descriptores de entropía de la imagen (a)
- (c) Descriptores de energía de la imagen (a)
- (d) Descriptores de correlación de la imagen (a)



5 *Método de binarización y homogeneización de imágenes*





5.0 Introducción

Los métodos de binarización y homogeneización de una imagen consisten en el proceso de reducción de información de la misma, en la que sólo persisten dos (para el caso binario verdadero o falso) y n-valores respectivamente [LSS05]. En una imagen digital, estos valores pueden representarse por 255, 192, 128 o 0 niveles de gris, es decir, blanco, gris claro, gris oscuro y negro respectivamente para imágenes multiclases en el método de homogeneización, para el caso de binarización solamente se necesitan 2 valores (usualmente se emplean el blanco y negro).

En el proceso y análisis de imágenes, estos métodos se emplean para separar o preclasificar regiones u objetos de interés en una imagen del resto. Las imágenes binarias se usan en operaciones booleanas o lógicas para identificar individualmente objetos de interés o para crear máscaras sobre regiones.

En muchos casos, una imagen binaria o multinivel homogeneizada es el resultado de una segmentación por niveles de gris o de una segmentación por selección de un rango de color determinado. En otros casos, una imagen homogeneizada en n-niveles es simplemente el resultado de una selección interactiva de regiones de interés, las cuales se utilizarán como mascararas de comparación o referencia.

En este trabajo se utilizarán estos métodos para realizar una mejor preclasificación de las imágenes obtenidas mediante los descriptores de textura, con el objeto de tener mejores resultados y resaltar de una forma más eficaz las características de las texturas obtenidas, así también se reducirá el número de clases no importantes con el fin de destacar ciertas áreas de interés.

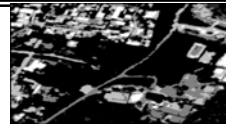
5.1 Método de binarización óptima

Se supone una imagen, que contiene solamente 2 regiones principales de niveles de gris.

El histograma de este tipo de imagen puede considerarse como una estimación de la función de densidad de probabilidad del nivel de gris, es decir, $p(z)$. Esta función de densidad total es la suma o mezcla de 2 densidades unimodales, una para las regiones claras de la imagen y otra para las regiones oscuras [RCG96]. Además los parámetros de la mezcla son proporcionales a las superficies de cada región de n-gris (nivel de gris). Es posible obtener el umbral óptimo para la segmentación de una imagen en dos regiones de brillo sí se conoce la forma de las densidades (en términos del error mínimo).

Suponiendo que una imagen contiene dos modas combinados con ruido aditivo gaussiano, la función de densidad de probabilidad de la mezcla esta dada por la siguiente ecuación:

$$P(z) = P_1p_1(z) + P_2p_2(z) \dots\dots\dots ec(5.1)$$



Para el caso gaussiano, es:

$$P(z) = \frac{P_1}{\sqrt{2\pi\sigma_1}} e^{[-(z - \mu_1)^2 / 2\sigma_1^2]} + \frac{P_2}{\sqrt{2\pi\sigma_2}} e^{[-(z - \mu_2)^2 / 2\sigma_2^2]} \dots\dots\dots ec.(5.2)$$

Donde:

- μ_1 y μ_2 son las medias de los 2 niveles de brillo.
- Las σ son las desviaciones estándar respecto a las medias.
- P_1 y P_2 son las probabilidades a priori de los niveles, en donde se debe cumplir con la condición $P_1 + P_2 = 1$.

Por lo tanto la densidad de la mezcla tendrá cinco parámetros desconocidos, y en el caso en que se conocieran todos los parámetros, la umbralaje puede ser determinada con facilidad.

Suponiendo que las regiones oscuras correspondan al fondo y que las claras a los objetos, se tendría que $\mu_1 < \mu_2$, entonces se puede definir un umbral T de forma que todos los píxeles con un nivel superior a T se consideren puntos objeto. La probabilidad de clasificar erróneamente un punto objeto del fondo esta dado por:

$$E_1(T) = \int_{-\infty}^T p_2(z) dz \dots\dots\dots ec.(5.3)$$

Y de la misma forma la probabilidad de clasificar erróneamente el fondo de un punto objeto esta dado por:

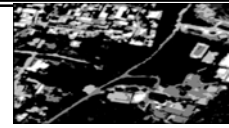
$$E_2(T) = \int_T^{\infty} p_1(z) dz \dots\dots\dots ec.(5.4)$$

Por lo tanto el error total es:

$$E(T) = \int_{-\infty}^T p_2(z) dz + \int_T^{\infty} p_1(z) dz \dots\dots\dots ec.(5.5)$$

Para encontrar el valor de umbral que hace que el error sea mínimo es necesario derivar E(T) con respecto a T utilizando la regla de Leibnitz e igualar el resultado a cero, es decir:

$$P_1 p_1(T) = P_2 p_2(T) \dots\dots\dots ec.(5.6)$$



Aplicando este resultado a la densidad gaussiana, tomando logaritmos y simplificando, se obtiene la siguiente ecuación cuadrática:

$$AT^2 + BT + C = 0 \dots\dots\dots ec.(5.7)$$

Donde:

- $A = (\sigma_1^2 - \sigma_2^2)$
- $B = 2(\sigma_2^2 \mu_1 - \sigma_1^2 \mu_2)$
- $C = \sigma_1^2 \mu_2^2 - \sigma_2^2 \mu_1^2 + 2\sigma_1^2 \sigma_2^2 \ln \left[\frac{\sigma_2^2 P_1}{\sigma_1^2 P_2} \right]$

La posibilidad de tener dos soluciones indica que se pueden necesitar dos valores de umbral para obtener una solución óptima. Si las varianzas son iguales, es decir: $\sigma^2 = \sigma_2^2 = \sigma_1^2$ es suficiente con un umbral óptimo T.

$$T = \frac{(\mu_1 + \mu_2)}{2} + \frac{\sigma^2}{(\mu_1 - \mu_2)} \ln \left(\frac{P_2}{P_1} \right) \dots\dots\dots ec.(5.8)$$

Si las probabilidades P_1 y P_2 son iguales, el umbral óptimo es el promedio de las medias, lo cual es también válido para $\sigma = 0$. La determinación del umbral óptimo se puede realizar de forma similar para otras densidades de un solo modo, de la forma Raleigh y la normal.

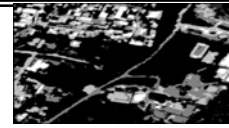
Se puede utilizar un método de error medio por mínimos cuadrados para estimar los parámetros de una imagen a partir de un histograma. Por ejemplo, el error cuadrático medio entre la densidad de la mezcla $p(z)$ y el histograma experimental $h(z)$ es:

$$e_{ms} = \frac{1}{n} \sum_{i=1}^n [p(z_i) - h(z_i)]^2 \dots\dots\dots ec.(5.9)$$

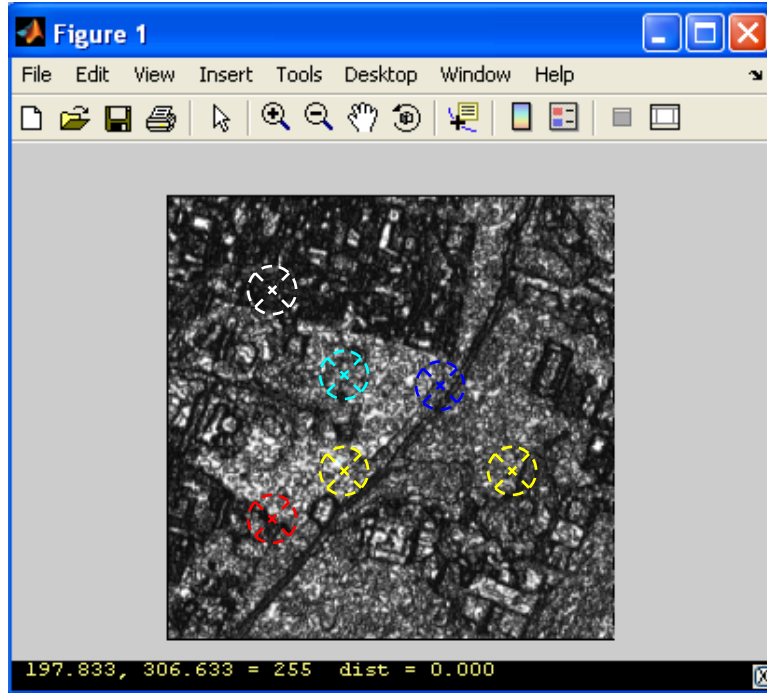
Donde se ha supuesto un histograma de n-puntos.

5.2 Método de binarización y homogeneización subjetiva propuesto

Mediante este método se determinan n-umbrales T para cada imagen, los cuales se obtienen por criterios personales, con ayuda de métodos de visualización de la imagen y de sus niveles, para así poder formar un criterio y poder emplear un método basado en la homogeneización óptima.



Con ayuda de las siguientes figuras se muestra un ejemplo de visualización y obtención de criterios de umbralaje, los cuales fueron empleados para la homogeneización en n-niveles de los descriptores de textura, basados en el método de homogeneización óptima.



Para esta imagen (figura 5.1) (descriptor de máxima probabilidad) se obtuvieron los siguientes criterios de umbralaje, los cuales permiten realizar el proceso de binarización de la imagen (homogeneización en 2 clases).

*Zonas urbanas
 $0 \leq T \leq 50$ (áreas blancas)

*Zonas no urbanas = 51-255
 $50 < T \leq 255$ (áreas oscuras)

Figura 5.1 Visualización de niveles de gris y posición un determinado píxel en una imagen.



En la figura 5.1 se muestra el procedimiento de obtención de niveles de gris (mediante la visualización del nivel de gris empleando matlab7) para el descriptor de entropía, cuyos resultados son mostrados a continuación:

```
244.265, 197.206 = 100 dist = 1.882
165.206, 131.324 = 236 dist = 0.000
345.912, 157.676 = 73 dist = 0.000
163.324, 451.324 = 89 dist = 1.882
```

Figura 5.2 Imagen binarizada.

En la figura 5.2 se muestra el resultado del proceso de binarización utilizando el descriptor de máxima probabilidad.

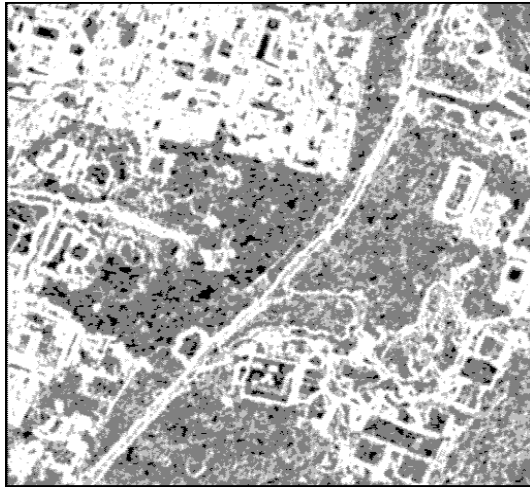
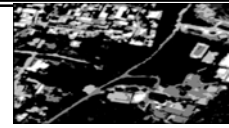


Figura 5.3 Imagen resultante del proceso de homogeneización en 4 clases.

Para realizar la homogeneización del descriptor de entropía se emplearon los siguientes umbrales, con el fin de obtener las 4 clases de mayor interés.

- $T_1 > 190 = 255$ blanco.
- $T_2 \geq 149$ y $T_2 \leq 190 = 192$ gris claro.
- $66 \leq T_3 \leq 148 = 128$ gris oscuro.
- $T_4 < 65 = 0$ negro.

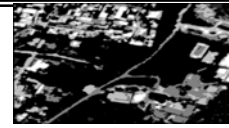
Algunos de los umbrales utilizados en el proceso de homogeneización en 2 (binarización) y 4 clases en los descriptores de textura, son los mostrados en las siguientes tablas:

Imagen	Umbrales (niveles de gris)	Imagen	Umbrales
Autocorrelación	Zonas urbanas (blancas) = 50-250 No urbanas (oscuras) = 0-50	Cluster shade	Urbanas = 50-225 No urbanas = 0-40
Cluster prominence	Zonas urbanas (blancas) = 20-221 No urbanas (oscuras) = 0-20	Varianza	Urbanas = 50-250 No urbanas = 0-50
Contraste	Zonas urbanas (blancas) = 30-225 No urbanas (oscuras) = 0-30	Correlación	Urbanas = 0-30 No urbanas = 30-200
Energía	Zonas urbanas (blancas) = 120-255 No urbanas (oscuras) = 0-120	Entropía	Urbanas = 0-150 No Urbanas = 150-220

Tabla 5.1 Umbrales binarios

Imagen	Umbrales (niveles de gris)	Imagen	Umbrales (niveles de gris)
Autocorrelación	$T_1 > 172 \Rightarrow 255$ ng (blanco) $90 \leq T_2 \leq 170 \Rightarrow 192$ gris claro $89 \geq T_3 \geq 50 \Rightarrow 128$ gris oscuro $T_4 < 40 \Rightarrow 0$ negro	Varianza	$T_1 > 90 \Rightarrow 255$ blanco $T_2 \geq 50$ y $T_2 \leq 90 \Rightarrow 192$ gris claro $T_3 \geq 21$ y $T_3 \leq 52 \Rightarrow 128$ gris oscuro $T_4 < 21 \Rightarrow 0$ negro
Cluster shade	$T_1 > 120$ $T_2 \geq 80$ y $T_2 \leq 120$ $T_3 \geq 45$ y $T_3 \leq 79$ $T_4 < 45$	Dissimilarity	$T_1 > 140$ $T_2 \geq 80$ y $T_2 = 139$ $T_3 \geq 50$ y $T_3 \leq 79$ $T_4 < 50$
Energía	$T_1 > 172 = 255$ $T_2 \geq 90$ y $T_2 \leq 170$ $T_3 \geq 41$ y $T_3 \leq 89$ $T_4 < 40$	Entropía	$T_1 > 190$ $T_2 \geq 149$ y $T_2 \leq 190$ $T_3 \geq 66$ y $T_3 \leq 148$ $T_4 < 65$

Tabla 5.2 Umbrales para la homogeneización en 4 clases



Mediante los métodos descritos anteriormente se realizó el proceso de homogeneización de las imágenes descriptoras de textura en n-clases.

Este método nos garantiza un resultado muy bueno en la homogeneización, ya que se realiza de forma supervisada, aunque es un proceso laborioso y tardado debido a que se tiene que realizar un número indeterminado de pruebas hasta encontrar los umbrales óptimos o que arrojen la información deseada, esto se ve compensado en los resultados obtenidos.

Debido a la obtención de diferentes umbrales para cada imagen esto nos minimiza la pérdida de información, la cual es muy importante para un proceso posterior, en cambio con un método de umbrales fijos, no se garantiza que exista una pérdida excesiva de información importante.

Los diferentes niveles de gris nos dan una clasificación de diversas texturas entre las cuales se encuentra:

Para el proceso de binarización:

El color blanco (nivel de gris 255) nos representa las áreas urbanas de nuestra escena, es decir: calles, edificios con diversa tonalidades, senderos asfaltados, etc.

El color negro (nivel de gris 0) representa las áreas no urbanas de nuestra escena, es decir: parques, bosques, arbustos, senderos de terracería, etc.

Para el proceso de homogeneización en 4 clases:

El color negro (nivel de gris 0) representa áreas no urbanizadas de nuestra escena, es decir: parques, bosques, arbustos, etc.

El color gris oscuro (nivel de gris 128) representa áreas poco urbanas de nuestra escena, es decir: senderos de terracería y áreas no del todo pavimentadas, etc.

El color gris claro (nivel de gris 192) representa áreas urbanas de nuestra escena, es decir: edificaciones con texturas no muy clara, techos, calles, etc.

El color blanco (nivel de gris 255) representa áreas muy urbanizadas de nuestra escena, es decir: edificaciones con texturas muy claras, techos, etc.

En el capítulo de resultados se puede realizar una comparación entre diferentes métodos de homogeneización realizados para 2 imágenes, entre los cuales se encuentran los métodos mostrados anteriormente y los que se encuentran en Matlab 7 denominados como `gray2ind(P,2)` y `im2bw(I,0.5)`. A continuación se muestra un ejemplo de los resultados obtenidos en la homogeneización de dos clases, en el que se podrán observar las diferencias entre los métodos empleados, así como sus ventajas y desventajas.

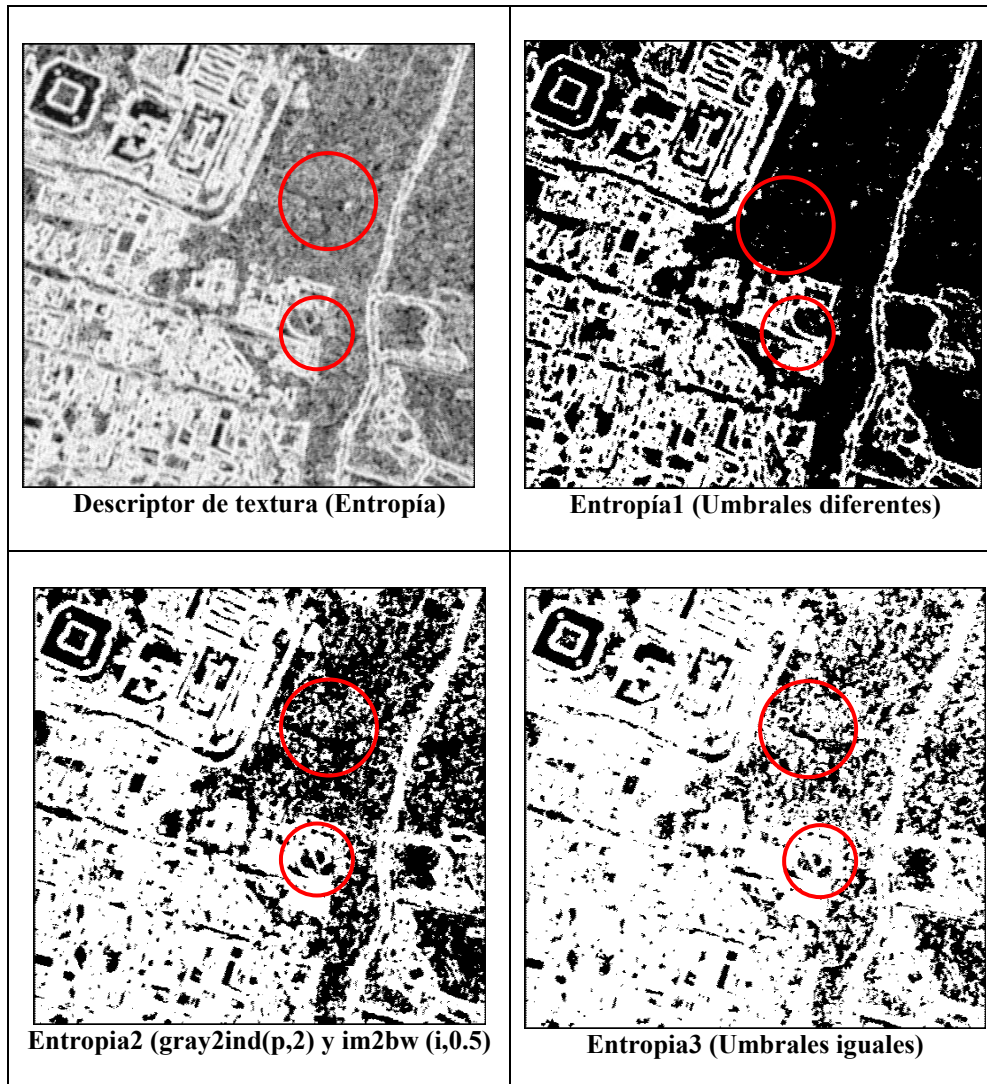
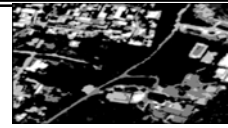
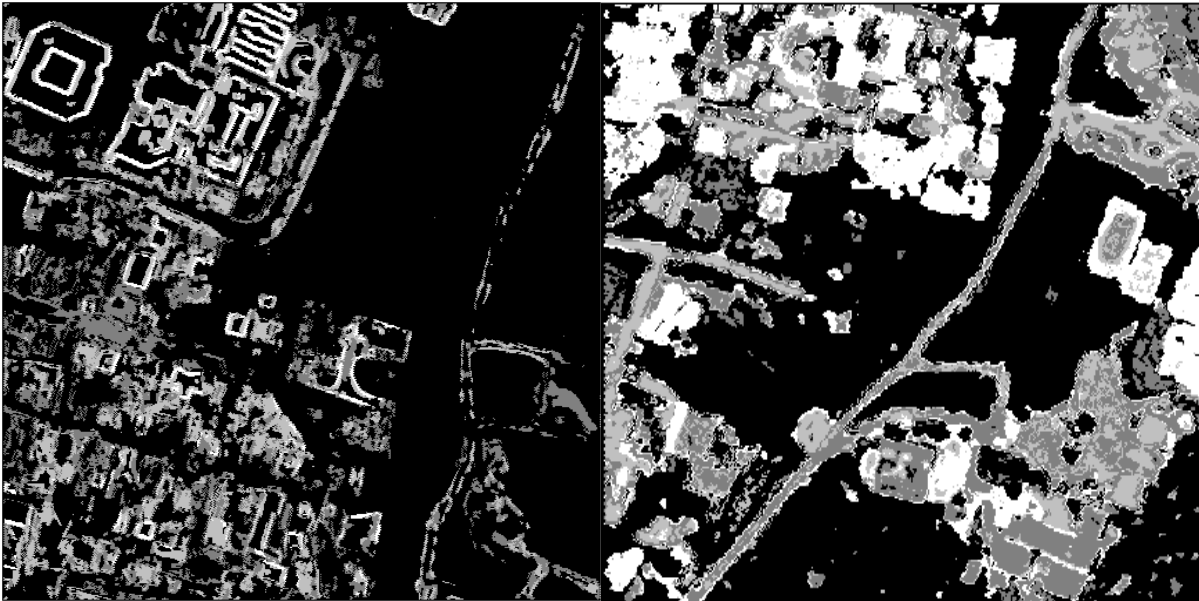


Figura 5.4 Resultados de homogeneización en 2 clases.

De la comparación de los resultados, nos podemos dar cuenta de las diferencias entre los métodos empleados; para el 1^{er} caso se emplearon distintos umbrales en cada imagen, en el 2^o se utilizaron algoritmos clásicos (Matlab7) y por último umbrales fijos para todas las imágenes. En este caso tratamos de diferenciar las zonas urbanas y no urbanas (apoyándonos en la imagen original), por lo cual optaremos por el método de diferentes umbrales, el cual nos ofrece mejores resultados, debido a que es un método supervisado. Por lo tanto para el análisis multiclasas se empleará este método.

Algunas de las diferencias que se pueden observar en las imágenes son indicadas con círculos rojos, en las cuales se puede identificar la confusión entre clases.

6 *Modelo de fusión*





6.0 Introducción

La fusión de datos es un conjunto de diversas técnicas, cuyo objetivo es el de mejorar ciertas características de los datos, como pueden ser la resolución, cobertura, etc. a partir de una serie de conjuntos de datos que por sí solos pueden carecer de importancia o relevancia.

Algunos métodos que se engloban bajo el epígrafe de fusión de datos son la integración de información vectorial y ráster, de información procedente de diferentes sensores, de datos GPS con cartografía catastral, o bases de datos tradicionales.

En este trabajo nos enfocaremos en el método propuesto por Shan Yu [YS04], el cual estima la confiabilidad del resultado, del análisis de una imagen digital, calculando sus medidas de error y basándose en la información de un mapa burdo de la imagen original [YSMB04]; se propone la modificación de este método para realizar una fusión multiclasas (4 clases) basándose en el método fundamental de Shan Yu.

También se propone obtener una mejoría en el proceso de fusión, con un pre-procesamiento de las imágenes aéreas; es decir, como se vio en capítulos anteriores, a las imágenes originales se les obtienen sus descriptores de textura mediante GLCM, con los cuales resaltamos ciertas características de textura; después, a estos descriptores, los procesamos mediante un algoritmo de homogeneización en n-clases, el cual nos resalta con más claridad la información que más nos interesa y desprecia aquella información o clases que no tienen gran relevancia en nuestro proceso, y finalmente se crea un criterio de fusión con el objeto de obtener mejores resultados.

6.1 Modelo de fusión utilizado en este trabajo

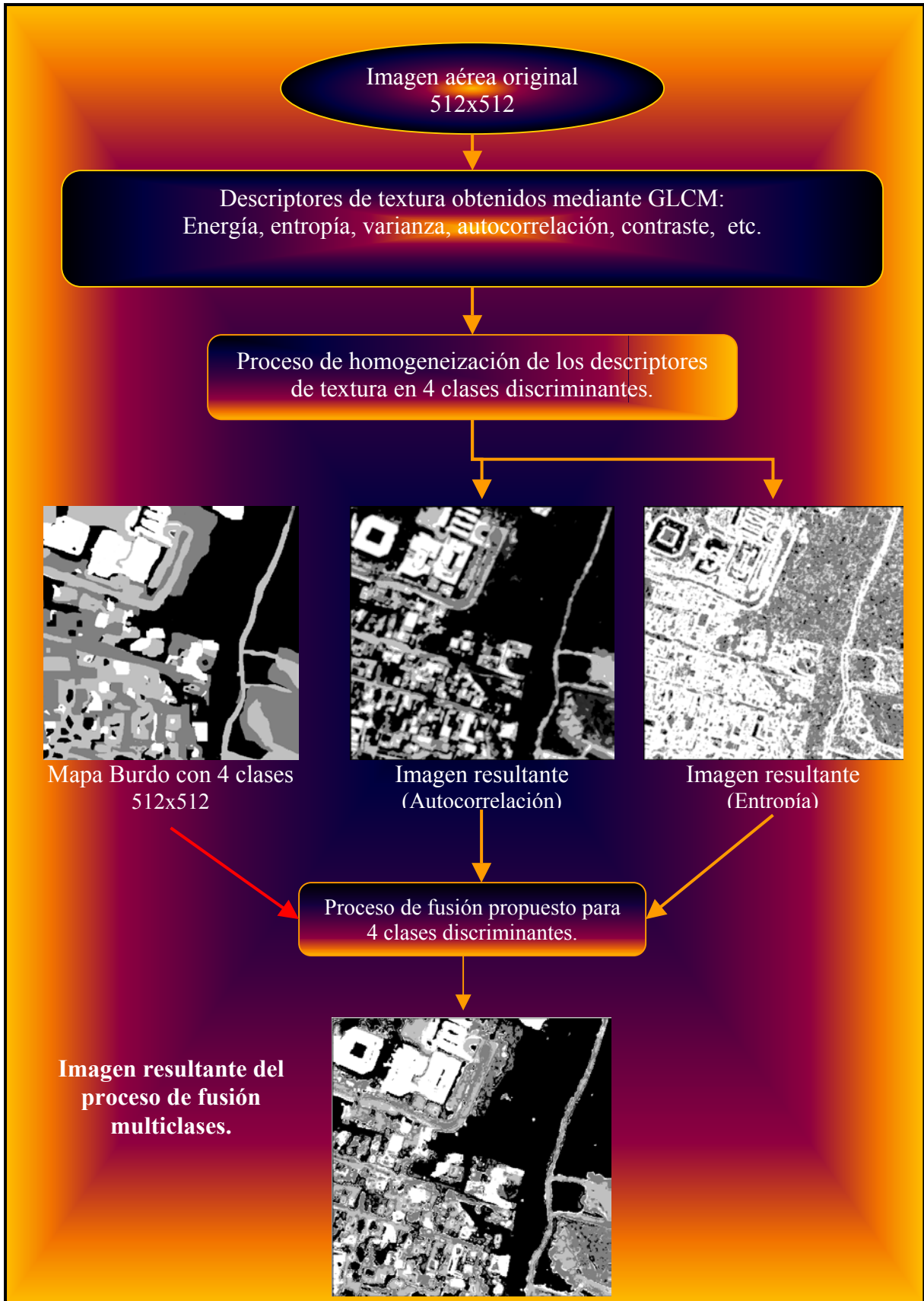
Debido a que las imágenes obtenidas remotamente son difíciles de analizar utilizando solamente un algoritmo, se considera la posibilidad de fusionar los resultados obtenidos mediante diferentes algoritmos, los cuales realizan una tarea similar sobre la misma imagen, sea esta satelital, aérea, etc. con el objetivo de obtener mejores resultados.

Para obtener mejores resultados lo primero que se debe de tomar en cuenta es la estimación de la confiabilidad de los resultados debidos a los diferentes algoritmos, después se considera el problema de la fusión para incrementar la confiabilidad, para así obtener una fusión confiable.

A continuación se describe el método propuesto por Shan Yu [YS04] para la estimación de la confiabilidad del análisis de una imagen, calculando sus errores y utilizando un mapa burdo de la imagen a procesar como referencia en el análisis. Después se realiza el proceso de fusión de las medidas de error. Este Proceso de fusión se realizó para el caso binario y multiclasas:

Para entender mejor esto, a continuación se muestra un diagrama del proceso.

Figura 6.1 MÉTODO PROPUESTO





6.1.1 Caso binario

En este caso solamente se utilizan dos clases las cuales son delimitadas por diferentes criterios, los cuales se enfocaron en la textura, clase y ubicación, por medio de los que se llegó a una clasificación general de áreas urbanas y no urbanas con niveles de gris 0 y 255 para definir a la imagen. Previamente se les realizó un proceso de homogeneización a las imágenes de textura, con el objeto de obtener mejores resultados en la fusión.

6.1.2 Caso multiclases

En este caso se propone la modificación y mejora del algoritmo utilizado en el caso de la fusión binaria (algoritmo Shan Yu), para realizar una fusión multiclases con imágenes de textura homogeneizadas de 5 a 4 clases, y obtener mejores resultados en el proceso de fusión.

6.1.3 Evaluación de confiabilidad

Para estimar la confiabilidad resultante del análisis de una imagen se calcularán dos tipos de error, un error local para cada píxel del resultado y un error global para el resultado completo.

El error global es la acumulación del error local sobre todas las regiones de la imagen; El error global determina si el resultado es lo suficientemente bueno para ser usado en el proceso de fusión, mientras que el error local determina como debe tomarse en cuenta cada píxel dentro del proceso de fusión; de esta manera, un píxel con un pequeño error local tiene muy alta confiabilidad y por lo tanto tiene mayor influencia en el proceso de fusión.

La forma ideal de medir la calidad del resultado del análisis de una imagen es usando el campo de escena real, lo cual no siempre es posible, por lo tanto se puede utilizar la información extraída de mapas geográficos y en nuestro caso la forma mas fácil es la comparación con resultados previamente obtenidos por medio de otros modelos y criterios de fusión.

La información previamente simplificada y clasificada, obtenida de un mapa es muy fácil de usar, sin embargo, sólo arroja un modelo aproximado de la escena y, por lo tanto, es incierta e imprecisa, no obstante este mapa puede ayudarnos a enfatizar o despreciar cierta información en la imagen.

Para realizar la clasificación de una imagen, asumamos que el mapa se ha registrado con la imagen y que:

1. s es un lugar o una región de la imagen.
2. S denota a la imagen completa.
3. $L = \{l_i, 1 \leq i \leq N\}$ denota un conjunto de clases o de regiones a segmentar.



- 4. $M(s)$ es la clase que pertenece a la región s de acuerdo al mapa.
- 5. $I(s)$ es la clase de s obtenida por el resultado del análisis de imagen.

Asumiendo que $M(s) = l_i$ y que $I(s) = l_j$. El cambio de la región de clase l_i (en el mapa burdo) a clase l_j (en el resultado del análisis de la imagen) provoca un error llamado riesgo:

$$X_{l_i, l_j} \geq 0 \quad (X_{l_i, l_j} = 0 \text{ Si } i=j)$$

Este riesgo se estima usando la información del mapa y el resultado del análisis de la imagen. Cada una de las regiones de la imagen debe de relacionarse con su vecindario. Suponiendo que s' es una región en la vecindad V_s de la región s , y $d(s, s')$ la distancia entre s y s' , entonces se calculan las siguientes medidas [YS04], mediante las ecuaciones:

$$a(s) = \sum_{s' \in V_s} \frac{1_{\{M(s')=l_i\}}}{d(s, s')} \dots\dots\dots \text{ec. (6.1)}$$

$$b(s) = \sum_{s' \in V_s} \frac{1_{\{I(s')=l_j\}}}{d(s, s')} \dots\dots\dots \text{ec. (6.2)}$$

Donde $1_{\{ \cdot \}}$ es el indicador de una función.

- $a(s)$ Es el número de regiones que existen en V_s , y que son clasificadas como l_j en el mapa, su valor es inversamente proporcional a la distancia que tiene con respecto a la región s .
- $b(s)$ Representa el peso que el lugar s obtiene del resultado del análisis de la imagen. El tamaño de la vecindad depende de la resolución de la imagen. Si la resolución es alta el tamaño es grande.

El error local $E(s)$ del sitio s esta definido como:

$$E(s) = \frac{X_{l_i, l_j}}{ra + (1 + b)}, r \geq 1 \dots\dots\dots \text{ec. (6.3)}$$

Como el resultado del análisis de la imagen y el mapa contienen imprecisiones y errores; se considera que el mapa contiene información más confiable que el resultado del análisis de la imagen (es decir, tiene mayor peso en el proceso) y esto se verá reflejado en el coeficiente r .

El error global es la acumulación de las medidas de errores locales a través de cada parte de la imagen y está definido por la siguiente ecuación:

$$GE = \sum_{s \in S} E(s) \dots\dots\dots \text{ec. (6.4)}$$



6.2 Modelo de fusión

Una vez conocido el valor del error global obtenido del análisis de la imagen, podemos decidir si este resultado es lo suficientemente bueno para ser usado en el proceso de fusión. Para hacerlo, simplemente tomamos los mejores resultados de todos los disponibles. Otra forma de hacer esto es utilizando un umbral T, es decir, solo los valores que cumplan con la condición $GE < T$ serán retenidos para el proceso de fusión, sin embargo no es tan sencilla la determinación de un umbral T.

Aunque un resultado seleccionado para ser usado en el proceso de fusión tiene una confiabilidad alta, todas las regiones de este resultado no tienen la misma confiabilidad.

Entonces, procedemos a calcular un coeficiente de confiabilidad para cada región del resultado basándonos en el valor del error local.

$$C(s) = 1 - \frac{E(s)}{\sum_{l_i, l_j \in L} X_{l_i l_j}} \dots\dots\dots ec.(6.5)$$

Donde el denominador es el valor normalizado de E(s), es decir; $0 \leq C(s) \leq 1$

La decisión de la clase final que le corresponderá a cada región depende de la acumulación del coeficiente de confiabilidad del mismo lugar a través de todos los resultados almacenados:

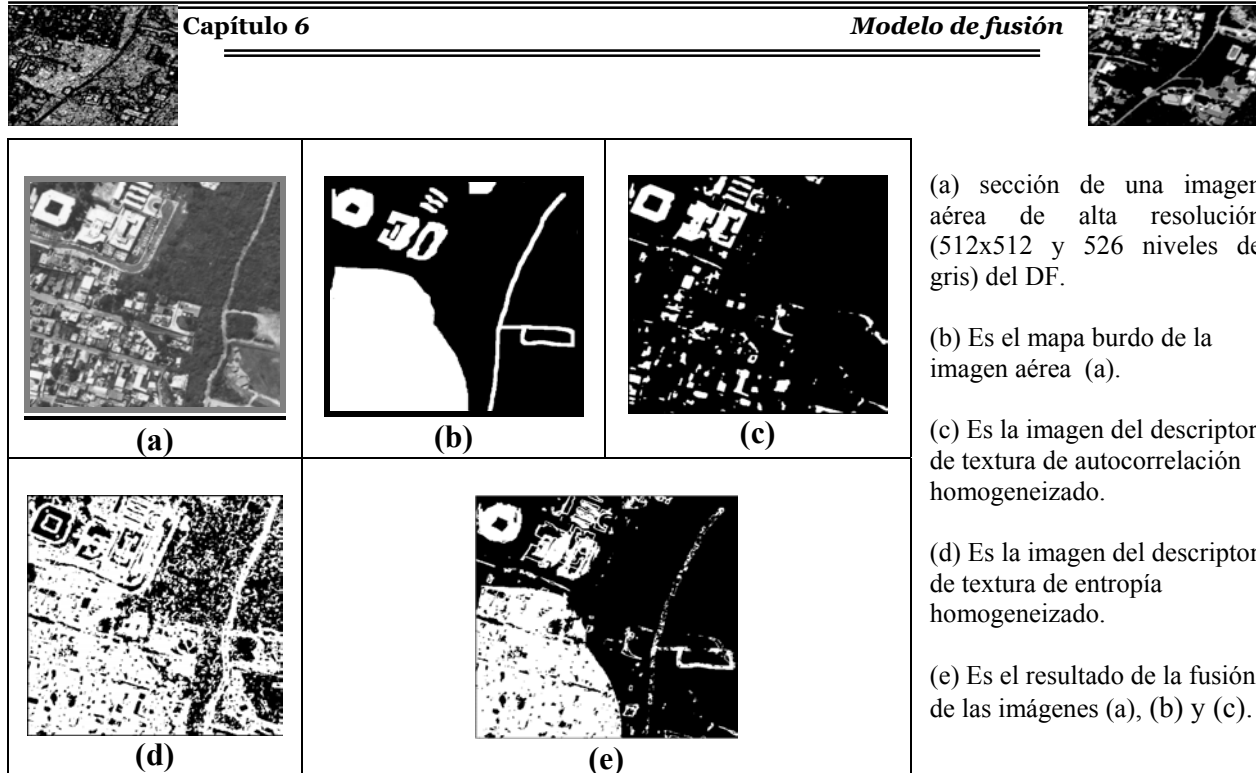
$$l(s) = \arg \max_{l_i \in L} \sum_{K=1}^p C_k(s) 1_{\{I_k(s)=l_i\}} \dots\dots\dots ec.(6.6)$$

En donde $I_k(s)$ significa que el resultado es usado como la k^{th} entrada al proceso de fusión y $C_k(s)$ es el coeficiente de confiabilidad de la región s en este resultado.

6.2.1 Ejemplo de fusión binaria

A continuación se muestra un ejemplo de la aplicación del método (Shan Yu) descrito anteriormente a un área urbana del Distrito Federal, en donde las imágenes a fusionar son los descriptores de textura homogeneizados, resultado de la GLCM y de la binarización de imágenes. En estas imágenes se utilizan solo dos clases para nombrar las diferentes regiones de la imagen: L_1 = área urbana (áreas de color blanco) y L_2 = área no urbana (áreas de color negro).

La figura 6.2 (a) muestra la sección de una imagen aérea del DF de alta resolución. La figura 6.2 (b) muestra el mapa burdo de la imagen aérea del DF. La figura 6.2 (c) muestra el descriptor de textura binarizado de autocorrelación. La figura 6.2 (d) es el descriptor de textura binarizado de entropía. La figura 6.2 (e) muestra el resultado de la fusión binaria.



(a) sección de una imagen aérea de alta resolución (512x512 y 526 niveles de gris) del DF.

(b) Es el mapa burdo de la imagen aérea (a).

(c) Es la imagen del descriptor de textura de autocorrelación homogeneizado.

(d) Es la imagen del descriptor de textura de entropía homogeneizado.

(e) Es el resultado de la fusión de las imágenes (a), (b) y (c).

Figura 6.2 Ejemplo de fusión binaria

En el ejemplo anterior, se utilizó el método anteriormente expuesto, cuyos parámetros usados fueron: número de imágenes =2, un mapa burdo de la imagen original debidamente delimitado en las zonas de interés y con el mismo número de clases que las imágenes a fusionar, $r=2$, riesgo $X = 0, 1, 20$, niveles de gris = 2 (0 (negro) y 255(blanco)), 2 umbrales de decisión para la evaluación de la confiabilidad estrechamente relacionados con el proceso de homogeneización, y una selección adecuada de imágenes con características especiales e información adecuada para nuestro objetivo principal, es decir, clasificar de una forma deseada las áreas urbanas y no urbanas de nuestra escenas del Distrito Federal.

6.2.2 Modelo de fusión multiclases propuesto

El modelo propuesto en este trabajo, se enfoca en la fusión de imágenes áreas con 4 clases, la cual basa sus fundamentos en el método de Shan Yu utilizado en la fusión binaria, al cual se le proponen una serie de modificaciones y criterios de pre-procesamiento, con el fin de poder obtener un mejor resultado en el proceso de fusión.

Para entender mejor como se lleva a cabo la fusión multiclases, con fundamentos en la fusión binaria se muestra a continuación un ejemplo del modelo propuesto, donde los valores de las variables propuestas que se utilizan, son los mismos valores que se utilizaron en el ejemplo del caso binario (sección 6.2.1), es decir:

1. $p=2$ = número de imágenes.
2. $r=2$ = constante.
3. $X_{l_i,l_j} = 20, X_{l_j,l_i} = 1, X_{l_i,l_i} = 0$



El punto más importante en este modelo radica en los criterios de umbralaje, los cuales deben seguir un acuerdo estrecho con los umbrales de homogeneización, es decir, el mapa burdo y imágenes utilizadas deben contener 4 clases diferentes con 0, 128, 192 y 255 niveles de gris, los cuales generaran 4 criterios diferentes de umbralas T. A consecuencia de realizar un método supervisado, este nos da la facilidad de experimentar con los umbrales hasta obtener un resultado satisfactorio.

Las imágenes (descriptores de textura previamente procesados) y los mapas burdos utilizados en este trabajo fueron clasificados en 4 clases diferentes, para las cuales se propusieron los niveles de gris 0,128, 192 y 255.

Para facilitar el análisis, la clase 1 siempre será el valor de nivel de gris más pequeño, es decir; 0; con esta consideración, se calculan las a(s) y b(s) correspondientes para cada imagen y después se obtiene X_{li} , X_{lj} , E(s), C(s) (el coeficiente de confiabilidad para cada imagen) y finalmente se analiza cada píxel de cada imagen; si el píxel de una de las imágenes pertenece a la clase uno, entonces se multiplica al coeficiente de confiabilidad resultante por uno o por cero (dependiendo del caso), de tal forma que los coeficientes de confiabilidad se acumulen en determinada clase; para que al final del análisis, el píxel respectivo pertenezca a la clase con mayor valor de acumulación del coeficiente de confiabilidad C(s).

255	255	255	128	255
255	255	255	255	255
255	255	255	255	255
255	255	255	0	0
128	255	255	0	0

Los datos con los que vamos a trabajar corresponden a una sección o ventanas de 5x5 de las imágenes del mapa burdo, descriptores de textura de autocorrelación (tabla 6.2 (a)), varianza (tabla 6.2 (b)) y entropía (tabla 6.2 (c)), con el propósito de simplificar el análisis.

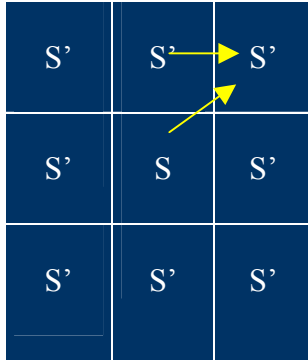
Tabla 6.1 Sección del Mapa Burdo

Imagen1 <i>Descriptor de autocorrelación</i>	Imagen2 <i>Descriptor de varianza</i>	Imagen3 (Opcional) <i>Descriptor de entropía</i>																																																																											
<table border="1"><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<table border="1"><tr><td>0</td><td>128</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table>	0	128	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	<table border="1"><tr><td>128</td><td>128</td><td>128</td><td>192</td><td>0</td></tr><tr><td>128</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>128</td><td>128</td><td>0</td><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>128</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table>	128	128	128	192	0	128	0	0	0	0	128	128	0	0	0	0	0	0	0	0	128	0	0	0	0
0	0	0	0	0																																																																									
0	0	0	0	0																																																																									
0	0	0	0	0																																																																									
0	0	0	0	0																																																																									
0	0	0	0	0																																																																									
0	128	0	0	0																																																																									
0	0	0	0	0																																																																									
0	0	0	0	0																																																																									
0	0	0	0	0																																																																									
0	0	0	0	0																																																																									
128	128	128	192	0																																																																									
128	0	0	0	0																																																																									
128	128	0	0	0																																																																									
0	0	0	0	0																																																																									
128	0	0	0	0																																																																									
Tabla 6.2(a)	Tabla 6.2(b)	Tabla 6.2(c)																																																																											

Tablas 6.2 Descriptores de textura homogeneizados



La distancia de una región s a la región s' se calcula mediante el siguiente método, haciendo uso de la ecuación 6.1:



Como se puede observar en la figura 6.3 si s' esta en dirección horizontal o vertical con respecto a s , la distancia es uno, si esta en dirección diagonal, la distancia es $\sqrt{2}$.

Como se puede observar en la tercera imagen, entre paréntesis aparece el texto opcional, esto es debido a que por resultados previamente obtenidos se comprobó que gracias al pre-procesamiento de las imágenes se obtuvieron mejores resultados mediante la fusión de solamente dos imágenes y un mapa burdo, lo cual simplifica y agiliza el análisis, obteniéndose buenos resultados.

Figura 6.3 Calculo de $a(s)$

Entonces para la 1^{er} imagen homogeneizada de textura (autocorrelación) se tiene que:

$$a(s) = 0 \quad Y \quad b(s) = 4 + \frac{4}{\sqrt{2}} = 6.8284$$

Para obtener el valor de X , consideraremos los siguientes aspectos:

Si la región analizada en el mapa pertenece a la clase 1 y la región analizada en la imagen es de clase 2, entonces $X_{i,j} = 0$. Ahora si la región analizada en el mapa pertenece a la misma clase que la región analizada en la segunda imagen $X_{i,j} = 0$, y, por último, si la región analizada en el mapa pertenece a la clase 2 y la región analizada en la imagen es de clase 1, $X_{i,j} = 1$.

El valor de $X_{i,j}$ es igual a 1 para la primera comparación, ya que la clase uno será el valor menor de las dos clases que están comparándose, como en este caso las clases que se están comparando son 0 y 128, por lo tanto, la clase menor es cero y es la primera clase.

Como se puede observar se compara por pares, hasta recorrer todas las clases y sus posibles comparaciones entre sí.

Entonces $X_{11} = 1$, debido a que la región analizada en el mapa pertenece a la clase 2, y la región analizada en la imagen pertenece a la clase 1.

Entonces:

$$E = \frac{X_{i,j}}{2a + b + 1} = \frac{1}{2(0) + 6.8284 + 1} = 0.1277 \quad Y \quad C = 1 - \frac{0.1277}{20 + 1} = 0.9939$$



Ahora para la 2ª imagen homogeneizada de textura (descriptor de varianza) se tiene que:

$$a(s) = 0 \quad b(s) = 4 + \frac{4}{\sqrt{2}} = 6.8284 \quad X_{21} = 1$$

$$E = \frac{X_{i,j}}{2a + b + 1} = \frac{1}{2(0) + 6.8284 + 1} = 0.1277 \quad Y \quad C = 1 - \frac{0.1277}{20 + 1} = 0.9939$$

Y para la 3ª imagen descriptora de textura (Entropía, imagen opcional) se tiene que:

$$a(s) = 0; \quad b(s) = 1 + \frac{1}{\sqrt{2}} = 1.7071; \quad X_{31} = 1$$

$$E = \frac{X_{i,j}}{2a + b + 1} = \frac{1}{2(0) + 1.7071 + 1} = 0.3693 \quad Y \quad C = 1 - \frac{0.3693}{20 + 1} = 0.9824$$

De forma similar al caso binario, después de obtener los coeficientes de confiabilidad para cada una de las imágenes, tomamos la región que analizamos y revisamos a qué clase pertenece; si pertenece a la 1ª clase se coloca su coeficiente de confiabilidad en $I_1(s)$, si pertenece a la 2ª clase lo colocamos en $I_2(s)$, si pertenece a la 3ª clase se coloca en $I_3(s)$, y por último si perteneciera a la 4ª clase lo colocamos en $I_4(s)$.

Después analizamos la región de la segunda imagen y realizamos el mismo procedimiento, si tanto la primera como la segunda región pertenecen a la misma clase sus coeficientes serán sumandos.

Para la región de la tercera imagen el procedimiento es el mismo, de tal forma que:

Para la clase 1:

$$I_1(s) = 0.9939 + 0.9939 = 1.9878$$

Como las 2 primeras regiones pertenecen a la clase uno, sus respectivos coeficientes de confiabilidad serán sumados.

Para la clase 2 tenemos que:

$$I_2(s) = 0.9824$$

Debido a que solamente la región analizada de la tercera imagen pertenece a esta clase.



Para la clase 3 y 4 tenemos que: $I_3(s)= I_4(s)=0$, debido a de que ninguna de las regiones analizadas para estas ventanas pertenece a estas clases.

a(s)

0	0	0	3.4142	5.1213
0	0	0	3.4142	6.1231
0	0	3.4142	5.1213	6.8284
0	0.7071	3.4142	6.1213	6.8284
0	1.7071	5.1213	6.8284	6.8284

Tabla6.3 (a)

b(s)

6.8284	6.8284	6.8284	6.8284	6.8284
6.8284	6.8284	6.8284	6.8284	6.8284
6.8284	6.8284	6.8284	6.8284	6.8284
6.8284	6.8284	6.8284	6.8284	6.8284
6.8284	6.8284	6.8284	6.8284	6.8284

Tabla6.3 (b)

Coefficiente 1

0.9939	0.9939	0.9939	0.9968	1.0000
0.9939	0.9939	0.9939	1.0000	1.0000
0.9939	0.9939	0.9948	1.0000	1.0000
0.9939	0.9939	0.9958	1.0000	1.0000
0.9939	0.9948	0.9968	1.0000	1.0000

Tabla6.3 (c)

Entonces se procede a realizar una comparación de los resultados obtenidos anteriormente; es decir, como $1.9878 > 0.9824$ entonces la región analizada pertenece a la clase cero, la cual puede visualizarse en la tabla resultante del proceso de fusión multiclasas que se encuentra al final de este capítulo.

Realizando el procedimiento anterior, obtenemos los siguientes resultados:

Para el caso de la 1^{er} imagen (descriptor de autocorrelación) tenemos los siguientes resultados de a(s), b(s) y c(s), respectivamente mostrados en las tablas 6.3 (a), 6.3 (b) y 6.3(c).

Tablas 6.3 Resultados del análisis de la imagen descriptora de autocorrelación.

Para la 2^a Imagen de textura (varianza) se obtuvieron los siguientes resultados de a(s), b(s) y c(s), respectivamente mostrados en las tabla 6.4 (a), 6.4 (b) y 6.4 (c).

a(s)					b(s)				
0	0	0	0.7071	3.4142	4.1213	0.7071	5.8284	6.8284	6.8284
0	0	0	1.7071	3.4142	5.1213	5.8284	6.1213	6.8284	6.8284
0	0	0.7071	3.4142	5.1213	6.1213	6.8284	6.8284	6.8284	6.8284
0	0	1.7071	3.4142	6.1213	6.8284	6.8284	6.8284	6.8284	6.8284
0	0.7071	3.4142	5.1213	6.8284	6.8284	6.8284	6.8284	6.8284	6.8284
Tabla 6.4 (a)					Tabla 6.4 (b)				



0.9907	0.9721	0.9930	0.9948	0.9968
0.9922	0.9930	0.9933	0.9958	1.0000
0.9933	0.9939	0.9948	0.9968	1.0000
0.9939	0.9939	0.9958	1.0000	1.0000
0.9939	0.9948	0.9968	1.0000	1.0000

Tabla 6.4 (c)

Tablas 6.4 Resultados del análisis de la imagen descriptora de varianza.

Para la 3^{er} imagen de textura analizada se tienen los siguientes resultados:

<table border="1" style="width: 100%; border-collapse: collapse;"> <caption style="text-align: center;">a(s)</caption> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>3.4142</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>1.7071</td><td>3.4142</td></tr> <tr><td>0</td><td>0</td><td>0.7071</td><td>3.4142</td><td>5.1213</td></tr> <tr><td>0</td><td>0.7071</td><td>1.7071</td><td>3.4142</td><td>6.1213</td></tr> <tr><td>0</td><td>0</td><td>3.4142</td><td>5.1213</td><td>6.8284</td></tr> </tbody> </table> <p style="text-align: center;">Tabla 6.5(a)</p>	0	0	0	0	3.4142	0	0	0	1.7071	3.4142	0	0	0.7071	3.4142	5.1213	0	0.7071	1.7071	3.4142	6.1213	0	0	3.4142	5.1213	6.8284	<table border="1" style="width: 100%; border-collapse: collapse;"> <caption style="text-align: center;">b(s)</caption> <tbody> <tr><td>6.1213</td><td>5.1213</td><td>3.7071</td><td>1.7071</td><td>6.8284</td></tr> <tr><td>5.8284</td><td>1.7071</td><td>3.7071</td><td>5.1213</td><td>6.1213</td></tr> <tr><td>4.4142</td><td>1.7071</td><td>5.8284</td><td>6.8284</td><td>6.8284</td></tr> <tr><td>1.7071</td><td>4.4142</td><td>6.1213</td><td>6.8284</td><td>6.8284</td></tr> <tr><td>1.7071</td><td>5.8284</td><td>6.8284</td><td>6.8284</td><td>6.1213</td></tr> </tbody> </table> <p style="text-align: center;">Tabla 6.5(b)</p>	6.1213	5.1213	3.7071	1.7071	6.8284	5.8284	1.7071	3.7071	5.1213	6.1213	4.4142	1.7071	5.8284	6.8284	6.8284	1.7071	4.4142	6.1213	6.8284	6.8284	1.7071	5.8284	6.8284	6.8284	6.1213
0	0	0	0	3.4142																																															
0	0	0	1.7071	3.4142																																															
0	0	0.7071	3.4142	5.1213																																															
0	0.7071	1.7071	3.4142	6.1213																																															
0	0	3.4142	5.1213	6.8284																																															
6.1213	5.1213	3.7071	1.7071	6.8284																																															
5.8284	1.7071	3.7071	5.1213	6.1213																																															
4.4142	1.7071	5.8284	6.8284	6.8284																																															
1.7071	4.4142	6.1213	6.8284	6.8284																																															
1.7071	5.8284	6.8284	6.8284	6.1213																																															
<table border="1" style="width: 100%; border-collapse: collapse;"> <caption style="text-align: center;">Coefficiente 3</caption> <tbody> <tr><td>0.9933</td><td>0.9922</td><td>0.9899</td><td>0.9824</td><td>0.9965</td></tr> <tr><td>0.9930</td><td>0.9824</td><td>0.9899</td><td>0.9950</td><td>1.0000</td></tr> <tr><td>0.9912</td><td>0.9824</td><td>0.9942</td><td>0.9968</td><td>1.0000</td></tr> <tr><td>0.9924</td><td>0.9912</td><td>0.9955</td><td>1.0000</td><td>1.0000</td></tr> <tr><td>0.9924</td><td>0.9942</td><td>0.9968</td><td>1.0000</td><td>1.0000</td></tr> </tbody> </table> <p style="text-align: center;">Tabla 6.5(c)</p>	0.9933	0.9922	0.9899	0.9824	0.9965	0.9930	0.9824	0.9899	0.9950	1.0000	0.9912	0.9824	0.9942	0.9968	1.0000	0.9924	0.9912	0.9955	1.0000	1.0000	0.9924	0.9942	0.9968	1.0000	1.0000	<table border="1" style="width: 100%; border-collapse: collapse;"> <caption style="text-align: center;">I₁(s)</caption> <tbody> <tr><td>1.9846</td><td>0.9939</td><td>1.9869</td><td>1.9897</td><td>2.9900</td></tr> <tr><td>1.9861</td><td>2.9693</td><td>2.9771</td><td>2.9875</td><td>3.0000</td></tr> <tr><td>1.9872</td><td>1.9878</td><td>2.9839</td><td>2.9903</td><td>3.0000</td></tr> <tr><td>2.9702</td><td>2.9790</td><td>2.9870</td><td>3.0000</td><td>3.0000</td></tr> <tr><td>1.9878</td><td>2.9839</td><td>2.9903</td><td>3.0000</td><td>3.0000</td></tr> </tbody> </table> <p style="text-align: center;">Tabla 6.5(d)</p>	1.9846	0.9939	1.9869	1.9897	2.9900	1.9861	2.9693	2.9771	2.9875	3.0000	1.9872	1.9878	2.9839	2.9903	3.0000	2.9702	2.9790	2.9870	3.0000	3.0000	1.9878	2.9839	2.9903	3.0000	3.0000
0.9933	0.9922	0.9899	0.9824	0.9965																																															
0.9930	0.9824	0.9899	0.9950	1.0000																																															
0.9912	0.9824	0.9942	0.9968	1.0000																																															
0.9924	0.9912	0.9955	1.0000	1.0000																																															
0.9924	0.9942	0.9968	1.0000	1.0000																																															
1.9846	0.9939	1.9869	1.9897	2.9900																																															
1.9861	2.9693	2.9771	2.9875	3.0000																																															
1.9872	1.9878	2.9839	2.9903	3.0000																																															
2.9702	2.9790	2.9870	3.0000	3.0000																																															
1.9878	2.9839	2.9903	3.0000	3.0000																																															
<table border="1" style="width: 100%; border-collapse: collapse;"> <caption style="text-align: center;">I₂(s)</caption> <tbody> <tr><td>0.9933</td><td>1.9643</td><td>0.9899</td><td>0.9824</td><td>0</td></tr> <tr><td>0.9930</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0.9912</td><td>0.9824</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0.9824</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table> <p style="text-align: center;">Tabla 6.5(e)</p>	0.9933	1.9643	0.9899	0.9824	0	0.9930	0	0	0	0	0.9912	0.9824	0	0	0	0	0	0	0	0	0.9824	0	0	0	0	<table border="1" style="width: 100%; border-collapse: collapse;"> <caption style="text-align: center;">I₃(s)</caption> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> </tbody> </table> <p style="text-align: center;">Tabla 6.5(f)</p>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.9933	1.9643	0.9899	0.9824	0																																															
0.9930	0	0	0	0																																															
0.9912	0.9824	0	0	0																																															
0	0	0	0	0																																															
0.9824	0	0	0	0																																															
0	0	0	0	0																																															
0	0	0	0	0																																															
0	0	0	0	0																																															
0	0	0	0	0																																															
0	0	0	0	0																																															

Tablas 6.5 Resultados del análisis de la imagen de entropía y de los valores obtenidos de cada clase I₁(s), I₂(s) y I₃(s).

En la tabla 6.5(a), (b) y (c) se muestran los valores obtenidos de a(s), b(s) y c(s) respectivamente, mientras que en las tablas (d), (e) y (f) se muestran los valores de I(s) para las clases 1, 2 y 3 respectivamente.

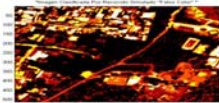


0	127	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

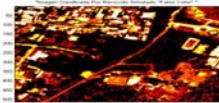
Tabla 6.6 Resultado de la fusión multiclases propuesta

En la tabla 6.6 se muestra el resultado del proceso de fusión multiclases. Donde el píxel seleccionado corresponde al nivel 0, debido a que la mayor información proveniente de las imágenes corresponde a este nivel.

Para el caso propuesto de fusión multiclases se debe dar gran importancia a los parámetros empleados para el cálculo de los coeficientes de confiabilidad (en especial en el riesgo X), debido al empleo de n -clases, y así también a los criterios de umbralaje empleados en el algoritmo de fusión (los cuales fueron previamente complementados en el proceso de homogeneización supervisada). Dependiendo de un buen criterio de estos parámetros se obtendrán buenos resultados en el proceso de fusión de datos para el caso de 4 clases. En el capítulo de resultados y conclusiones se podrá apreciar con más claridad el método propuesto, así como sus ventajas y desventajas.



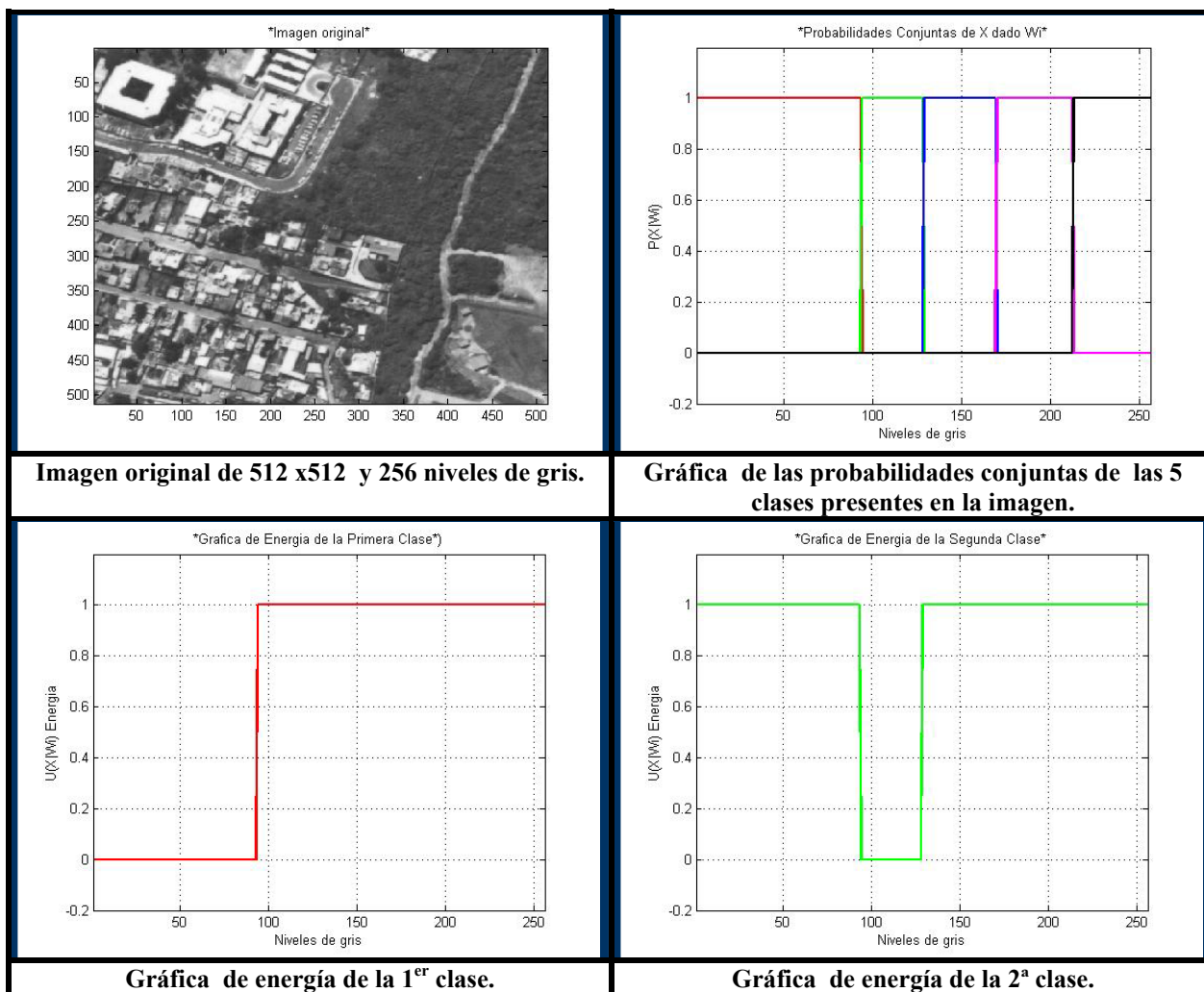
7 *Resultados*

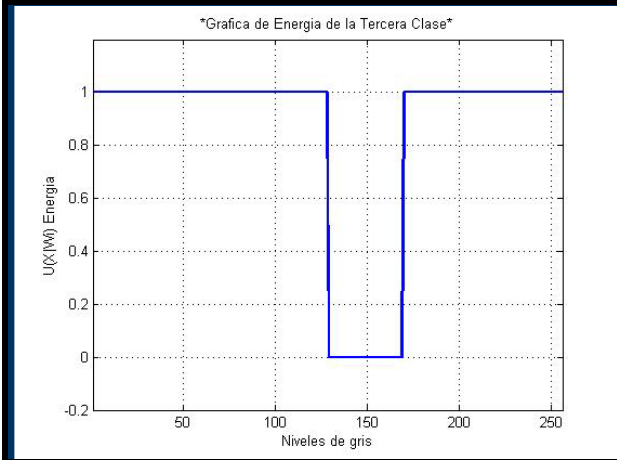
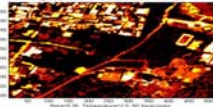


7.0 Resultados de segmentación bayesiana y el método contextual (Markov)

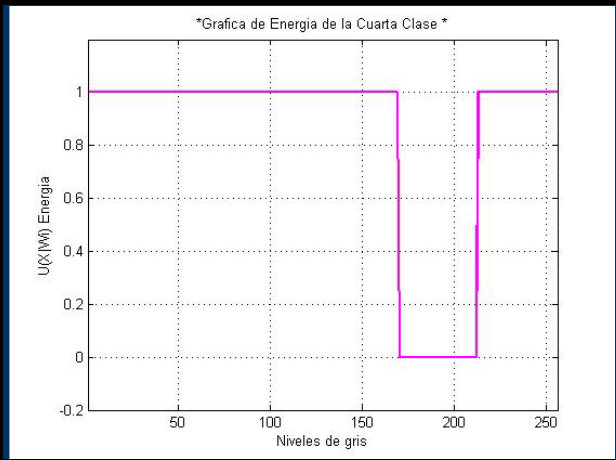
Las imágenes siguientes muestran el proceso de segmentación bayesiana y segmentación por el algoritmo basado en campos aleatorios de Markov (por medio de funciones de energía de tipo escalón), con el objetivo de comparar estos resultados con el proceso de fusión. Para el proceso de segmentación y clasificación de imágenes se utilizaron los siguientes parámetros:

1. Probabilidades iguales en la presegmentación.
2. Probabilidades actualizadas en la entapa de segmentación.
3. Funciones del tipo escalón definidas por las medias de cada clase.
4. $\beta = 0.35$
5. Temperatura = 2.5
6. Número de iteraciones para el algoritmo de recocido simulado = 60

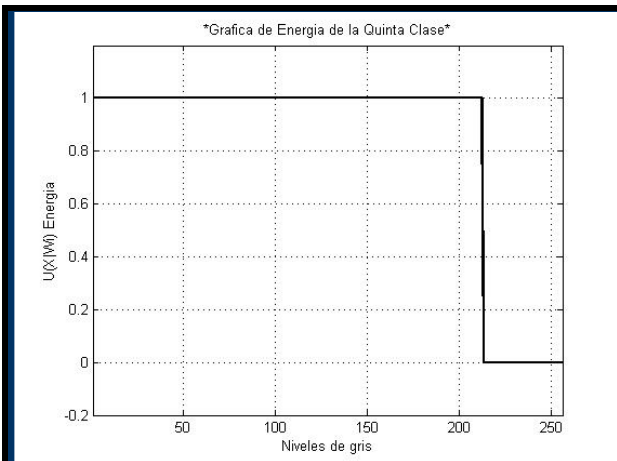




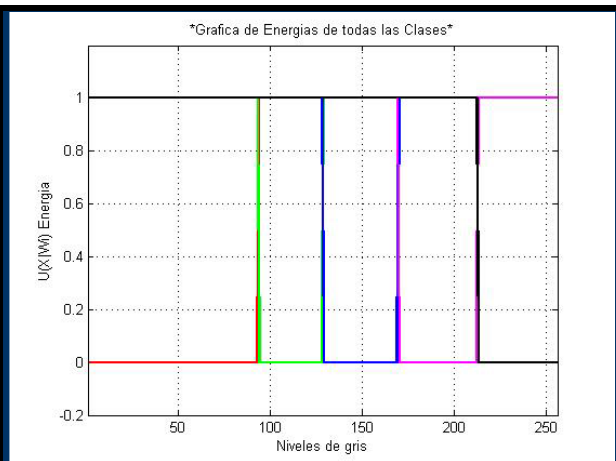
Gráfica de energía de la 3ª clase.



Gráfica de energía de la 4ª clase.



Gráfica de energía de la 5ª clase.



Gráfica de energía de todas las clases.

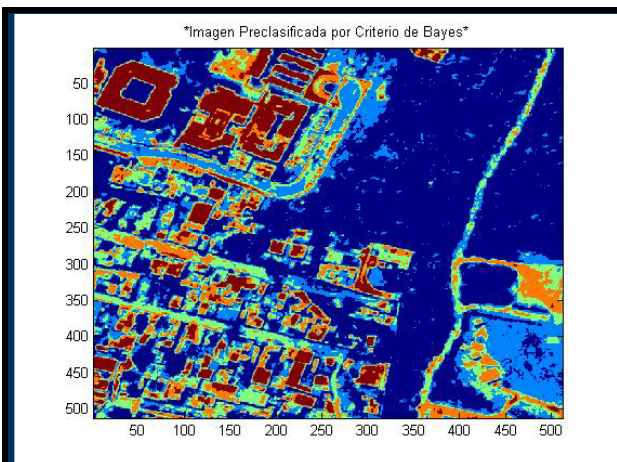


Imagen presegmentada por el método de Bayes.

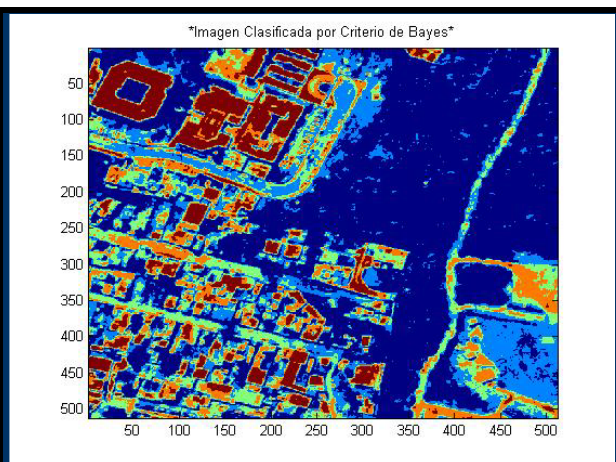


Imagen segmentada por el método de Bayes.

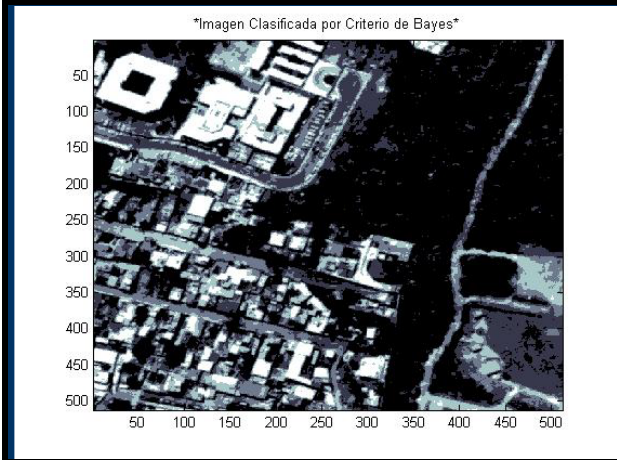
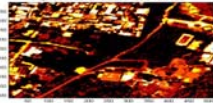


Imagen segmentada por el método de Bayes, en escala de grises.

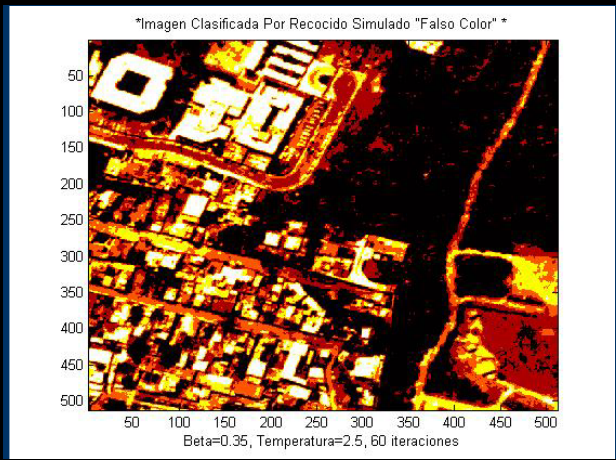


Imagen segmentada empleando el método de optimización de recocido simulado, en falso color.

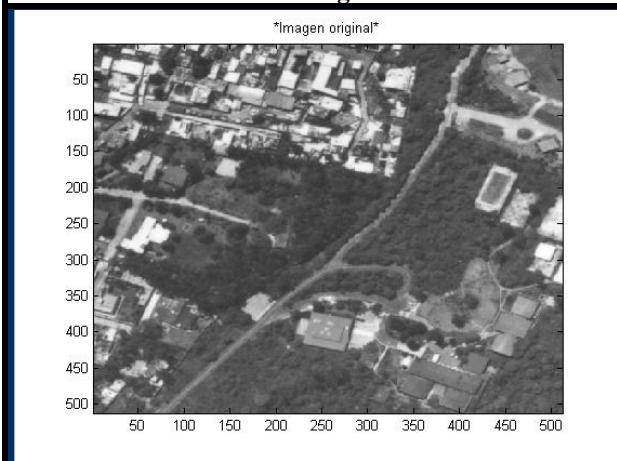
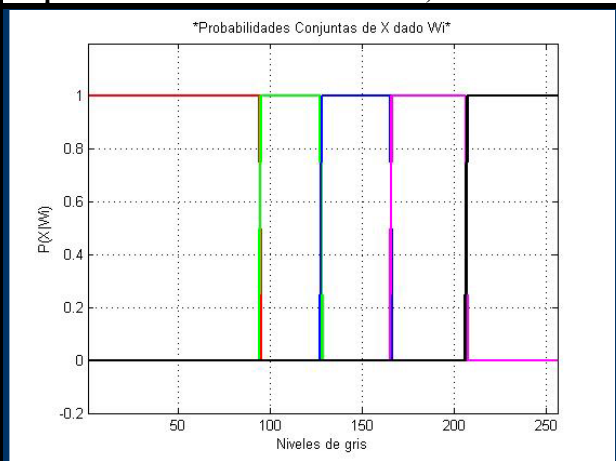
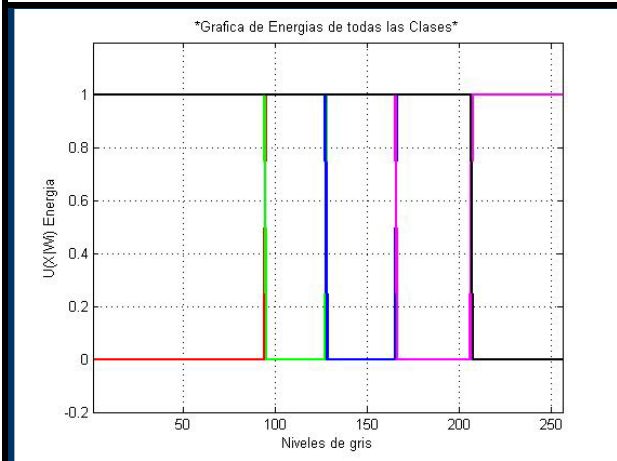


Imagen original de 512 x512 y 256 niveles de gris



Gráfica de las probabilidades conjuntas de las 5 clases presentes en la imagen.



Gráfica de energías de las 5 clases presentes en la imagen.

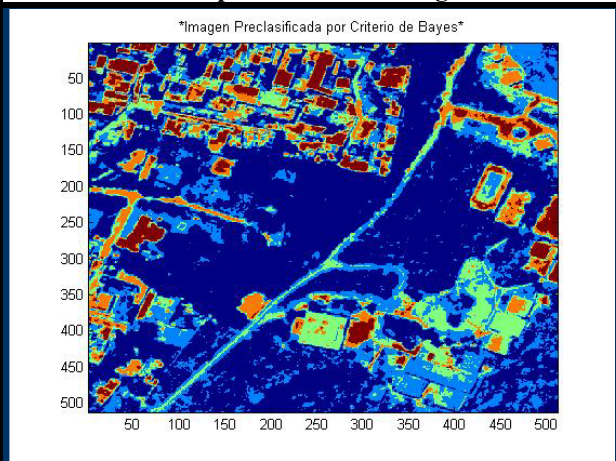


Imagen presegmentada por el método de Bayes.

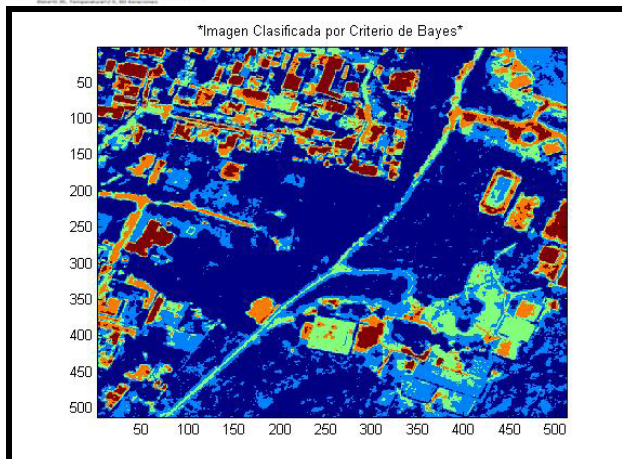


Imagen segmentada por el método Bayes.

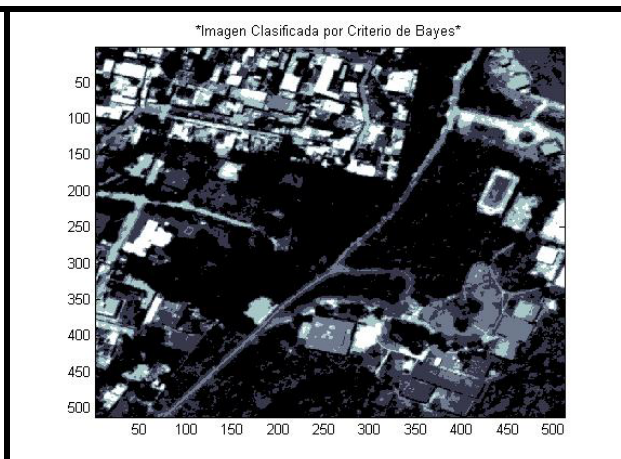


Imagen segmentada por el método de Bayes en escala de grises

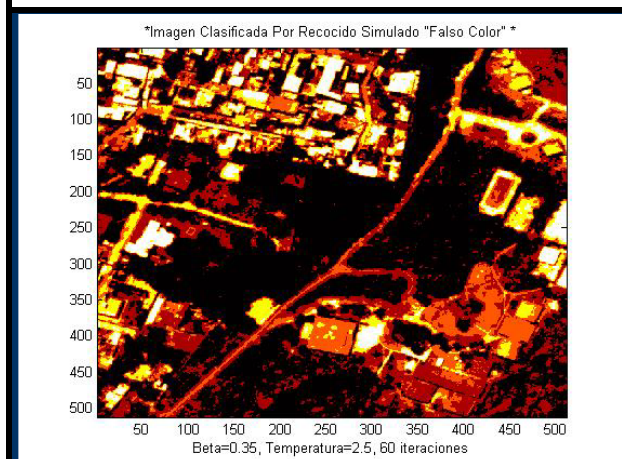


Imagen segmentada por el método de recocido simulado en falso color.

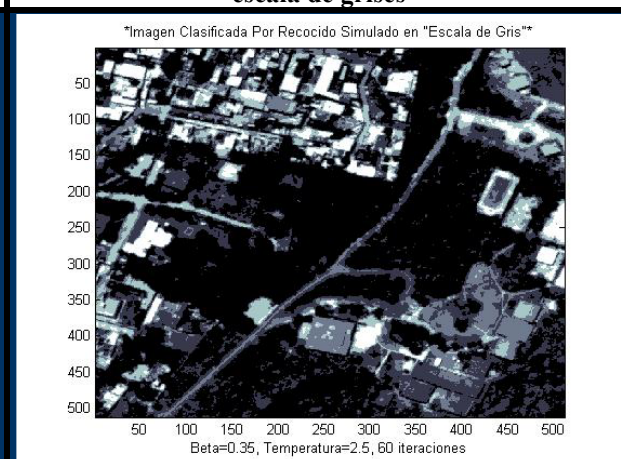
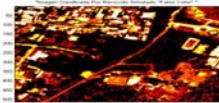


Imagen segmentada por el método de recocido simulado en escala de grises.

7.1 Descriptores de textura obtenidos mediante la matriz de co-ocurrencia

Mediante el algoritmo propuesto por Clausi llamado “Rapid Determination Of Co-Occurrence Textura Features” [DAC] (Descriptores de Textura).

Las imágenes que se presentan a continuación son los descriptores de textura obtenidos a través del cálculo de la matriz de co-ocurrencia para dos imágenes distintas, las cuales se utilizaron en el proceso de binarización y homogeneización de 4 clases, (que se presentarán posteriormente) para así poderlas utilizar en el proceso de fusión de imágenes propuesto en este trabajo. Estos descriptores fueron calculados con los valores que a continuación se mencionan:



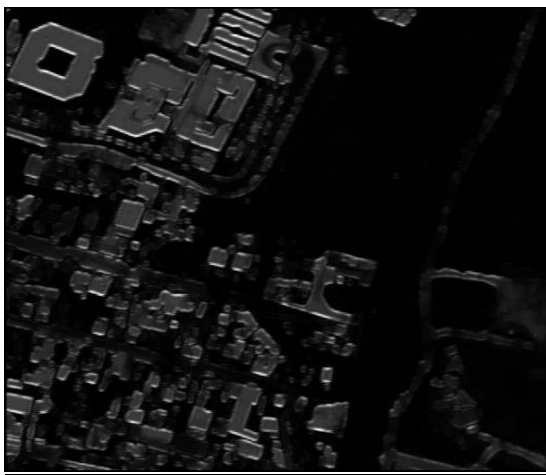
1. Reducción de 256 a 32 niveles de gris, evitando una pérdida excesiva de información.
2. Desplazamiento horizontal para la GLCM de 2.
3. Utilización de ventanas de 5x5
4. Número de clases presentes 5.



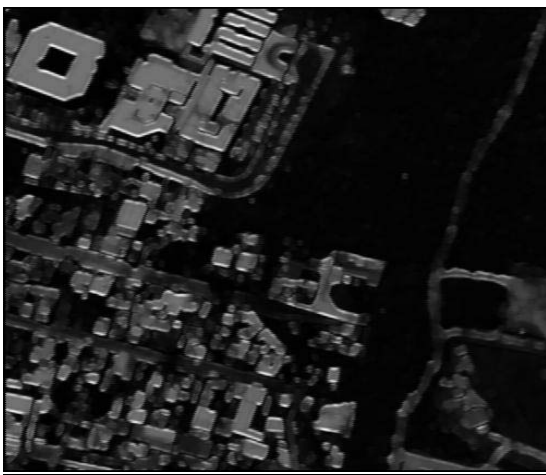
Imagen original



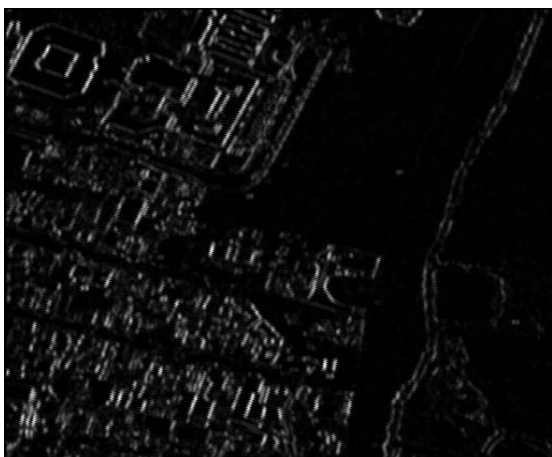
Autocorrelación



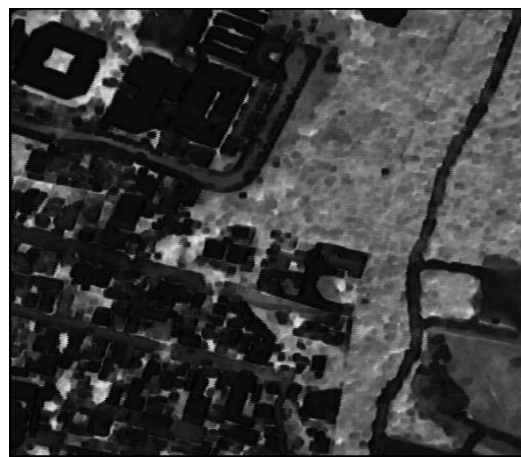
Cluster prominence



Cluster shade



Contraste



Correlación



Dissimilarity



Energía



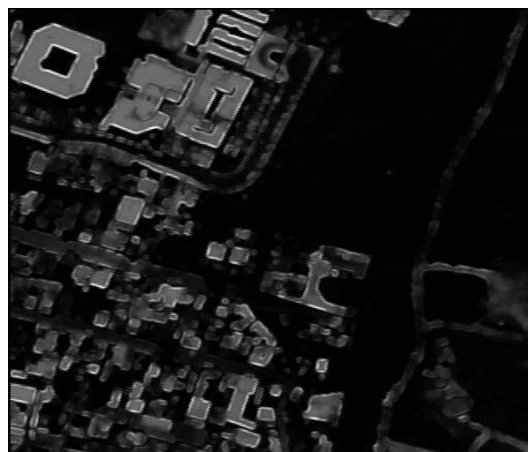
Entropía



Homogeneidad



Máxima probabilidad



Varianza



Imagen original "Villa olímpica"



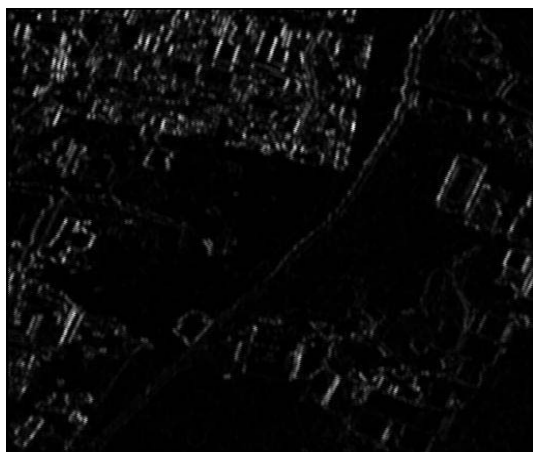
Autocorrelación



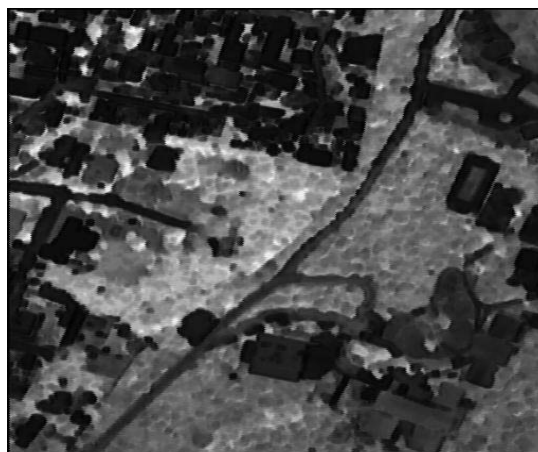
Cluster prominence



Cluster shade



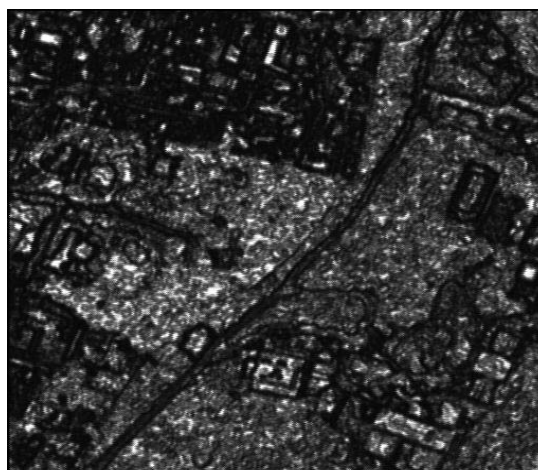
Contraste



Correlación



Dissimilarity



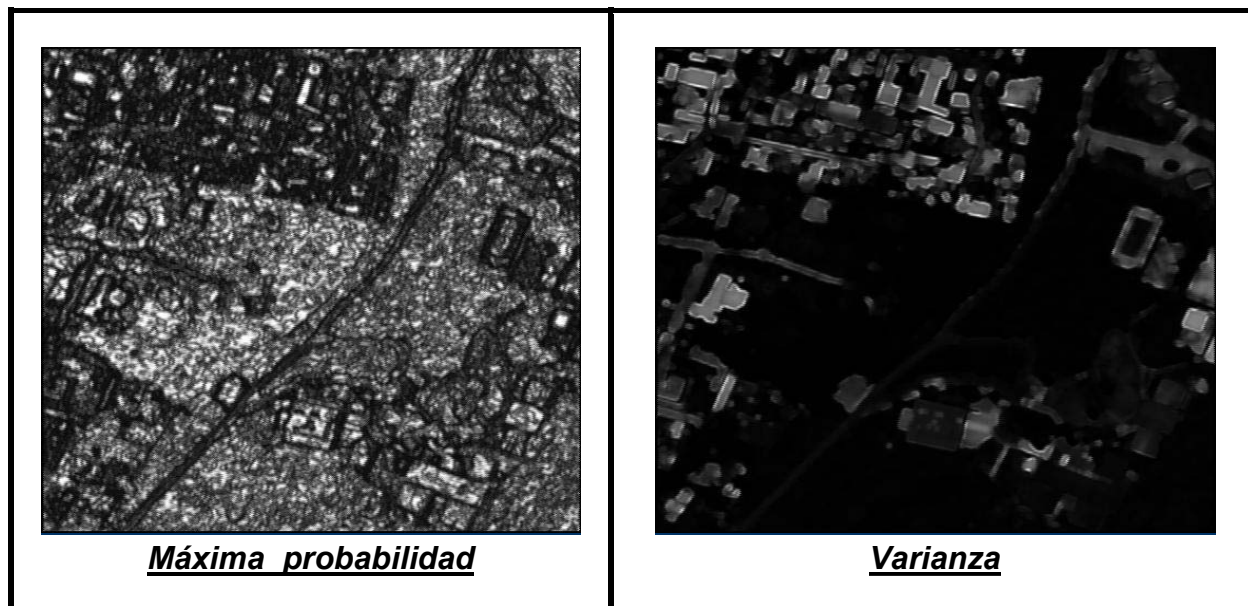
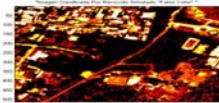
Energía



Entropía



Homogeneidad



7.2 Resultados del proceso de binarización y homogeneización de las imágenes de textura

(Descriptores de Textura) para imágenes diferentes, por medio de diferentes métodos y con parámetros diferentes.

Las imágenes que se presentan a continuación son los resultados obtenidos del proceso de homogeneización de los descriptores de textura calculados por medio de la matriz de co-ocurrencia, con tamaño de 512x 512 y 256 niveles de gris.

Estas imágenes tienen la característica de haber reducido sus niveles de gris en 2 y 4 clases con valores en la escala de gris de: 0, 255 y 0, 128, 192 y 255 para binarización y homogeneización 4 clases, respectivamente.

Se puede observar claramente que en algunas imágenes el método de binarización u homogeneización de n-clases utilizado, puede o no afectar significativamente en nuestro resultado, por lo tanto se probaron diferentes métodos, llegando a la conclusión de que el mejor método era el de una homogeneización supervisada con diferentes umbrales para cada imagen, debido a que es más fácil extraer la información necesaria y desechar aquella que no aporte un significado relevante de nuestro interés. A continuación se muestra 2 tablas de los umbrales empleados en los algoritmos de binarización y homogeneización en 4 clases, así como los resultados obtenidos en diferentes imágenes.



Imágenes obtenidas mediante la modificación de umbrales para cada imagen. Para obtener 2 clases con niveles de 0 y 255

Autocorrelación

Urbanas (blancas) = 50-250
no urbanas (oscuras)=0-50

Cluster prominence

Urbanas (blancas) = 20-221
no urbanas (oscuras)=0-20

Cluster shade

Urbanas (blancas) = 50-225
no urbanas (oscuras)=0-40

Contraste

Urbanas (blancas) = 30-225
no urbanas (oscuras)=0-30

Correlación

Urbanas (blancas) = 0-30
no urbanas(oscuras)=30-200

Energía

Urbanas (blancas) = 120-255
no urbanas (oscuras)=0-120

Entropía

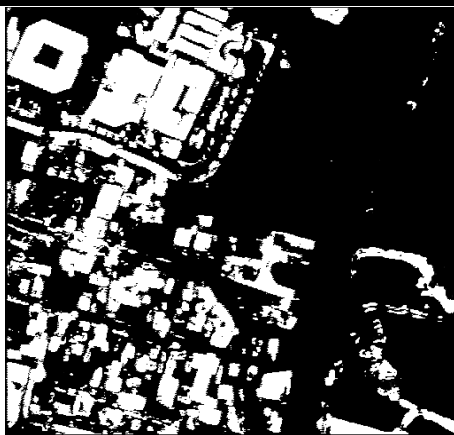
Urbanas (blancas) = 0-150
no urbanas (oscuras)=150-220

Varianza

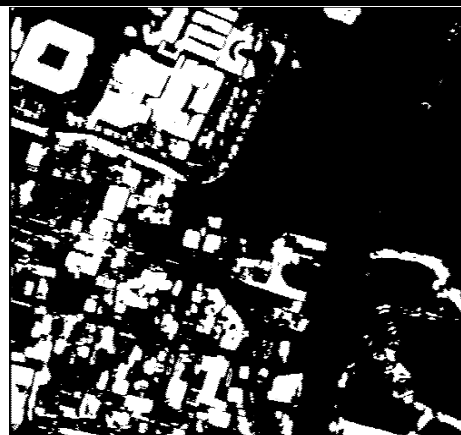
Urbanas (blancas) = 50-250
no urbanas (oscuras)=0-50

Máxima probabilidad

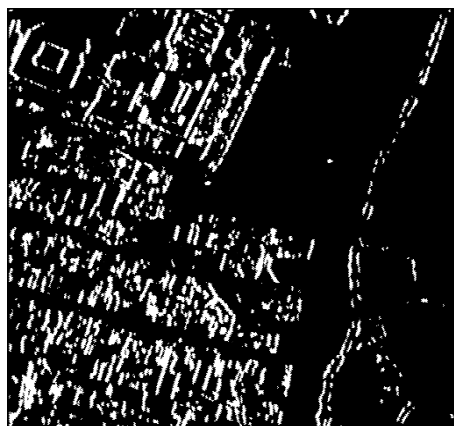
Urbanas (blancas) = 0-50
no urbanas (oscuras)=50-255



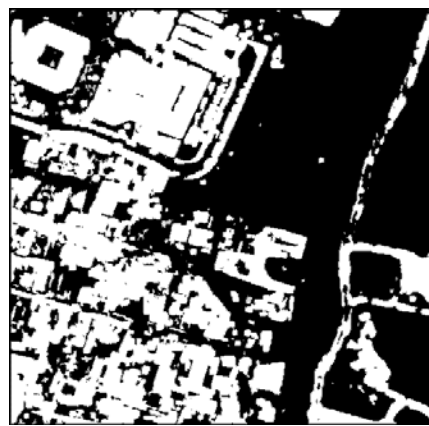
Cluster prominence1



Cluster shade1



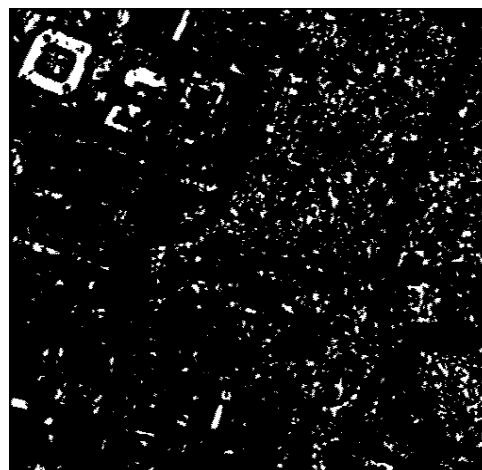
Contraste1



Correlación1



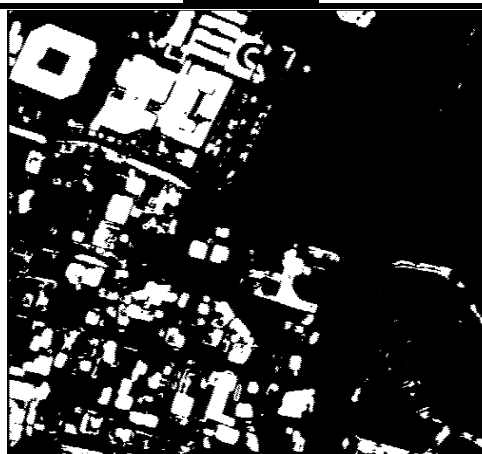
Entropía1



Energía1

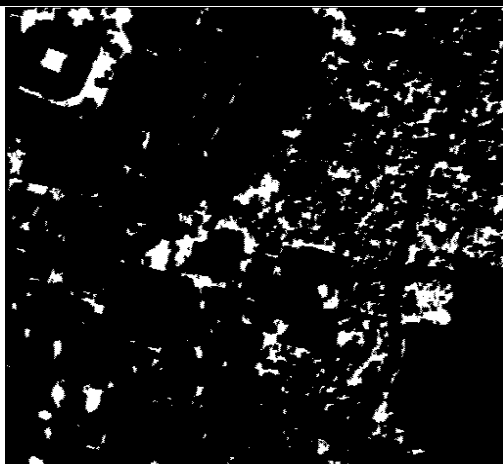


Máxima probabilidad1

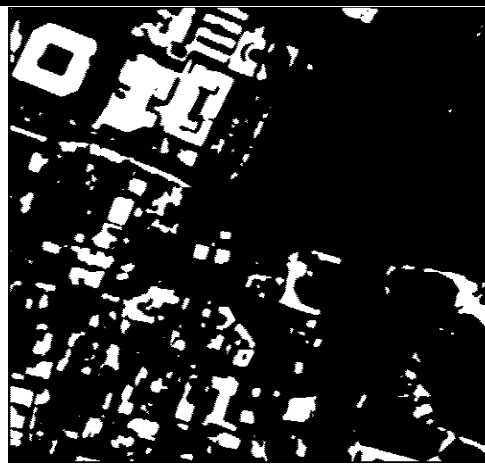


Varianza1

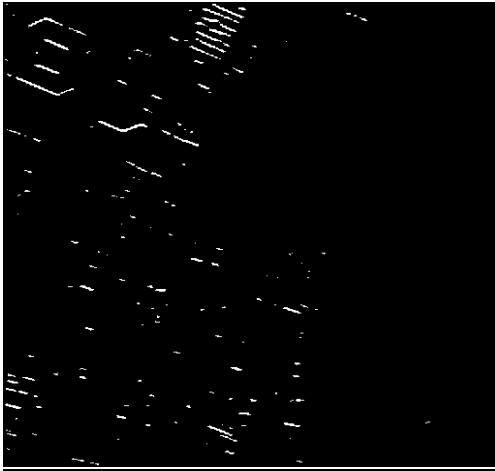
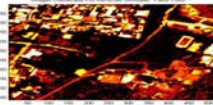
Imágenes obtenidas por algoritmos preestablecidos del programa Matlab versión 7
(gray2ind (P,2) y im2bw (I,0.5)) Para obtener clases 0 y 1



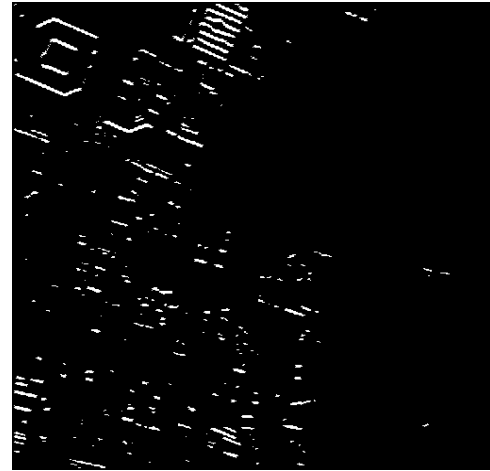
Correlación2



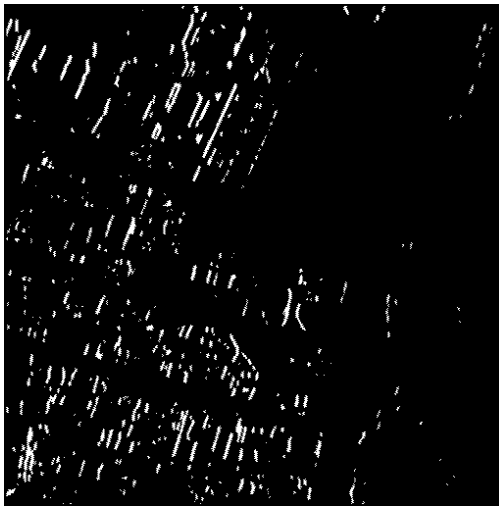
Autocorrelación2



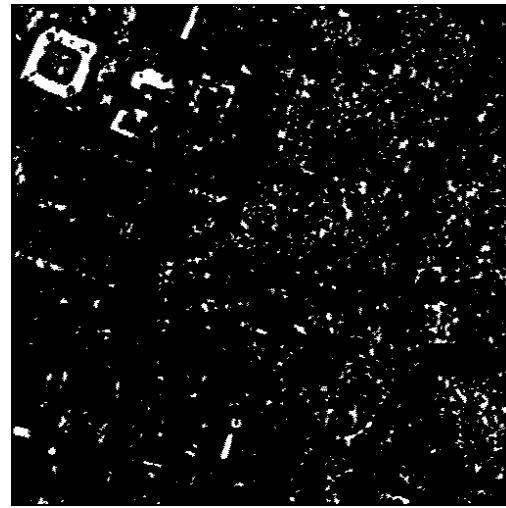
Cluster prominence2



Cluster shade2



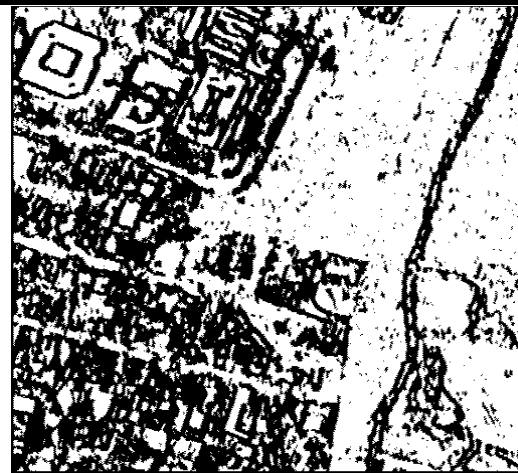
Dissimilarity2



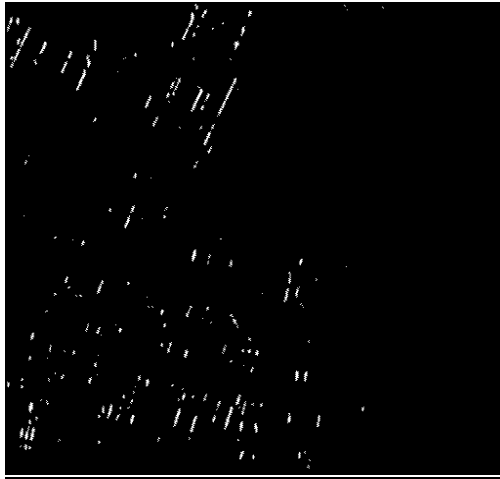
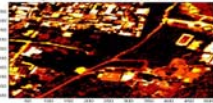
Energía2



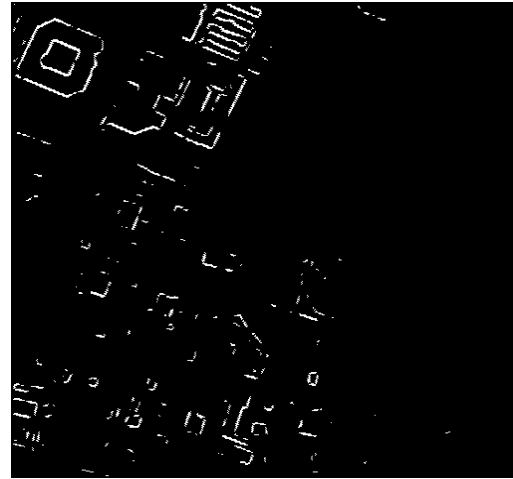
Entropía2



Homogeneidad2



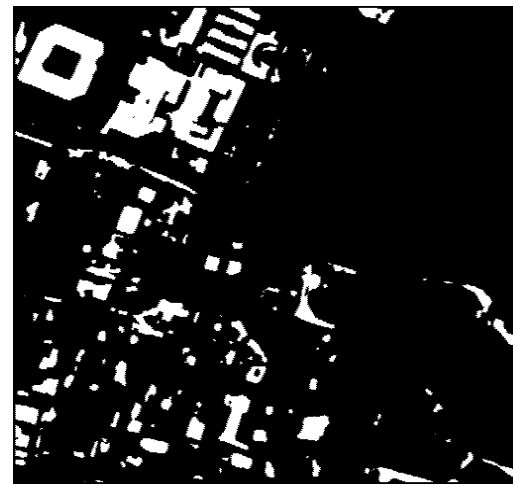
Contraste2



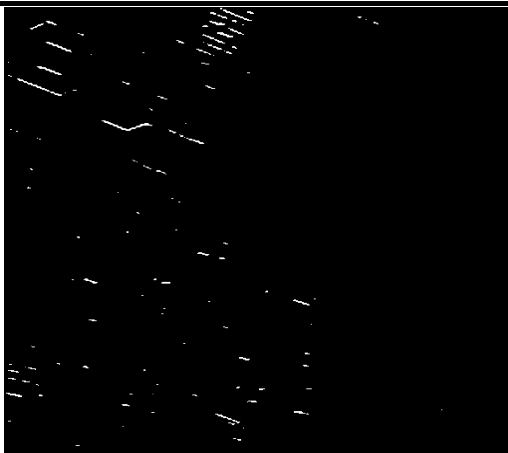
Varianza2

Imágenes obtenidas mediante umbrales fijos para todas las imágenes descriptoras de textura, para obtener 2 niveles de gris 0 y 255 (negro y blanco respectivamente) :

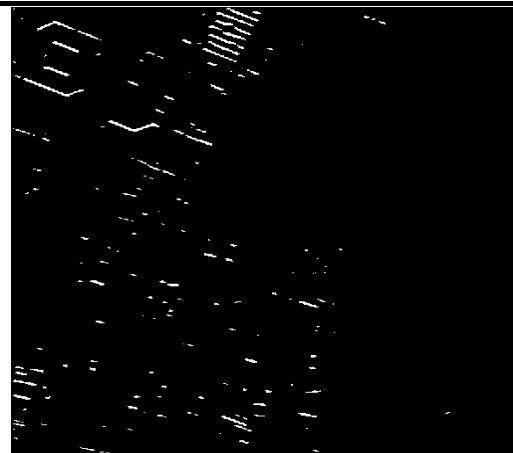
T = umbral empleado
T \leq 150 = nivel de gris=0
T > 151 y T < 256 = nivel de gris=256



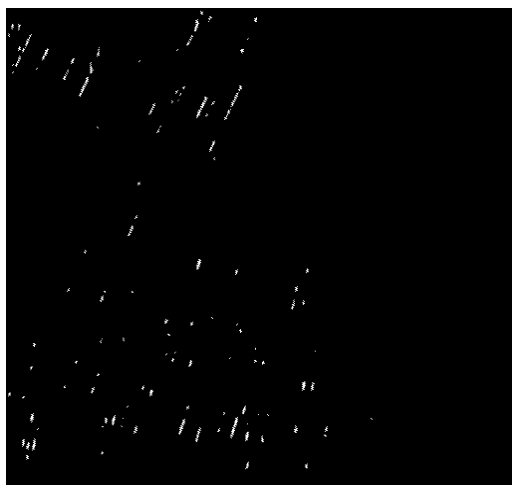
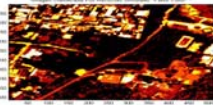
Autocorrelación3



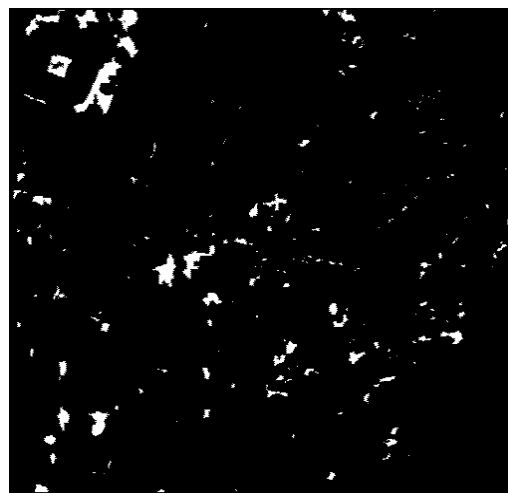
Cluster prominence3



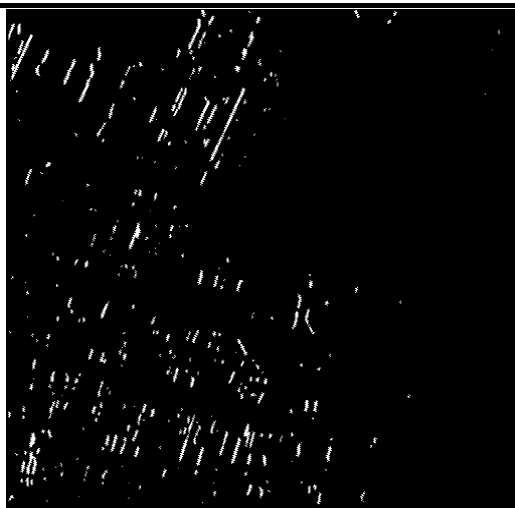
Cluster shade3



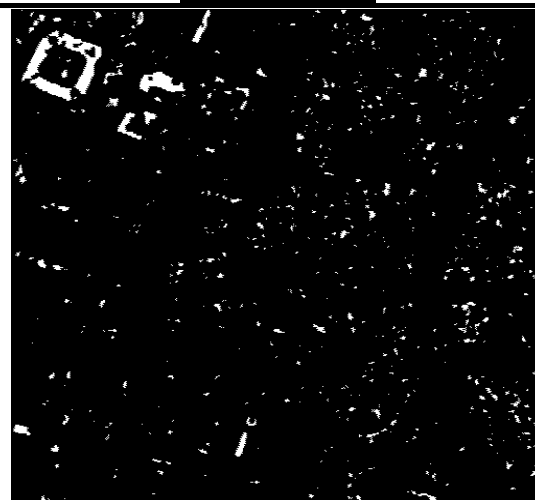
Contraste3



Correlación3



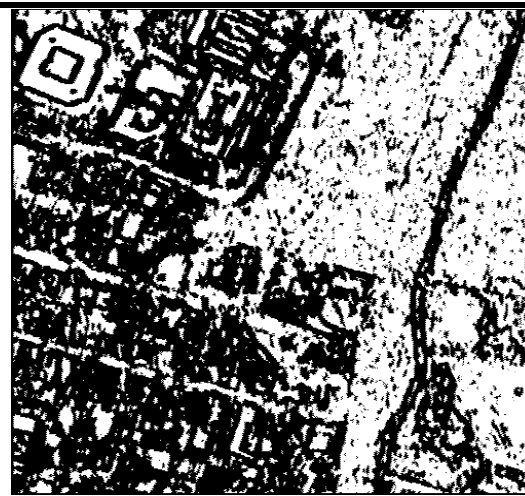
Dissimilarity3



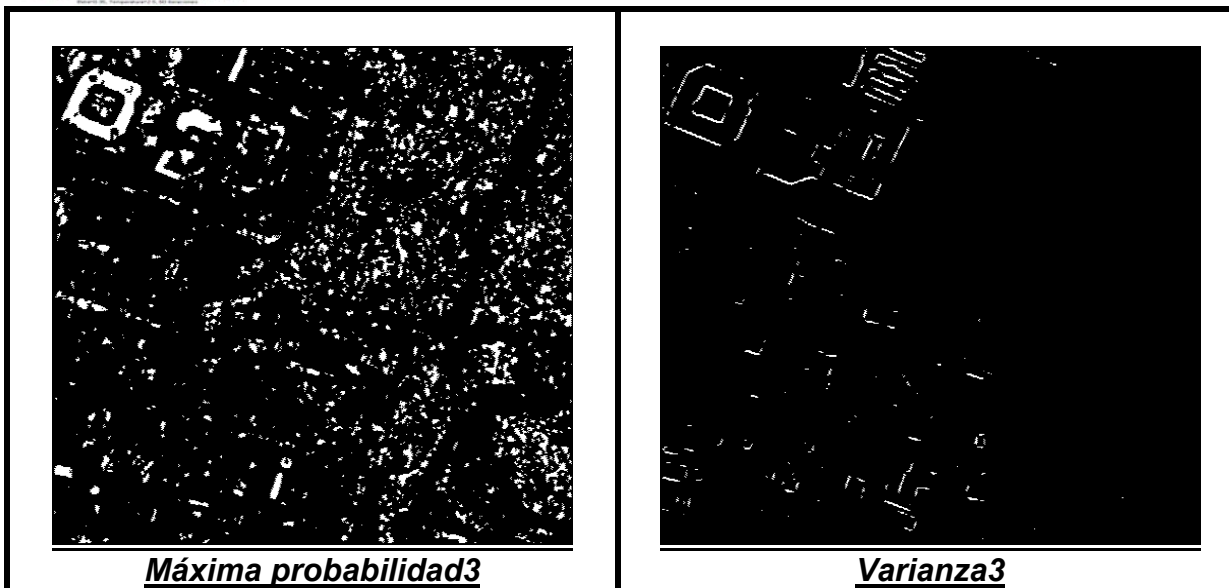
Energía3



Entropía3

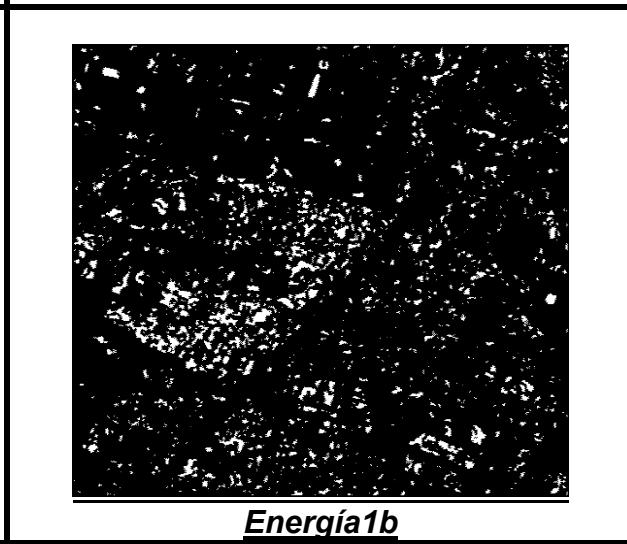
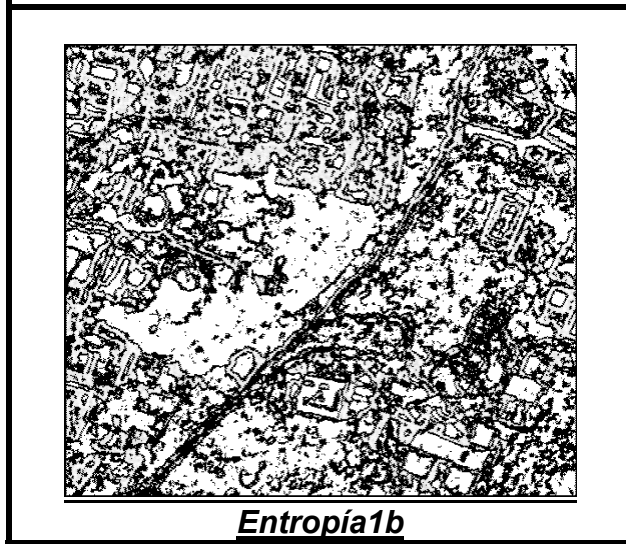
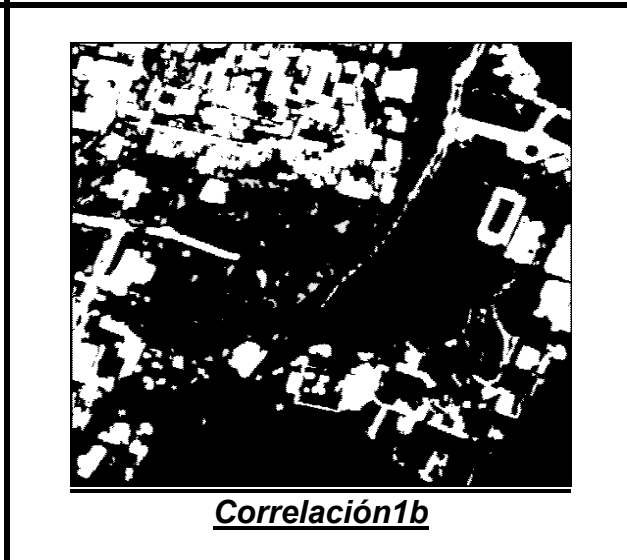
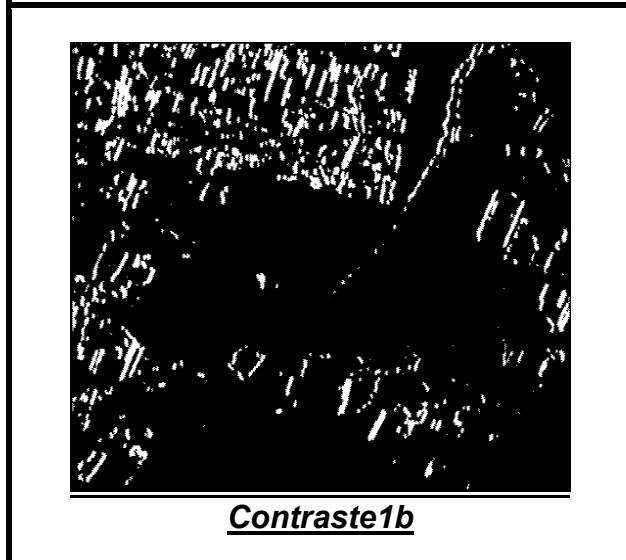
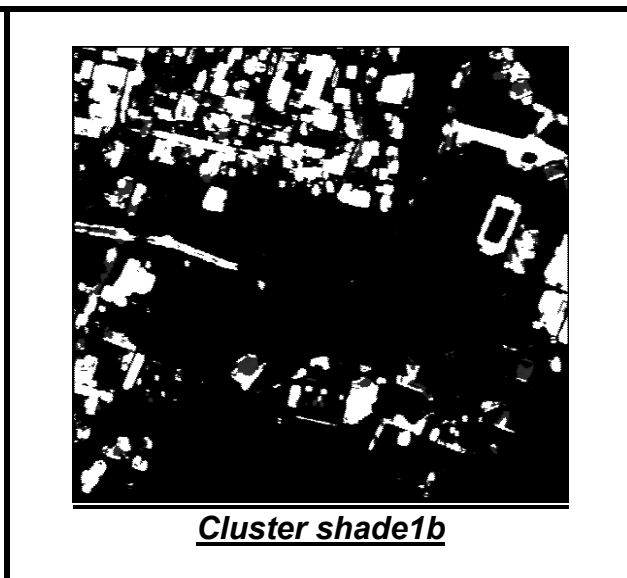
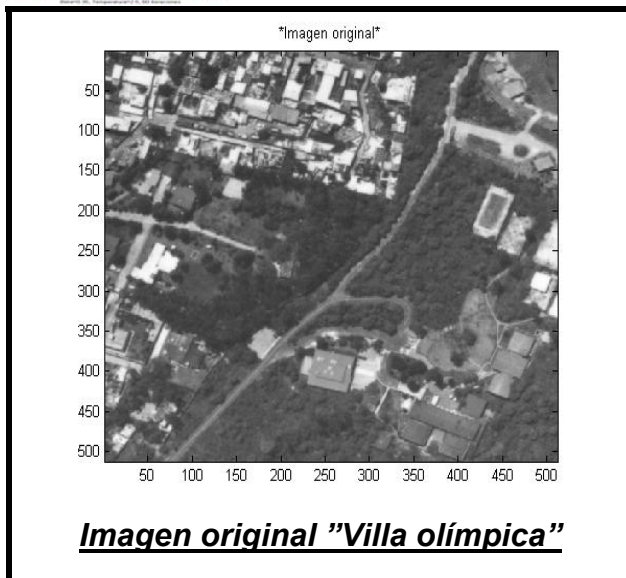


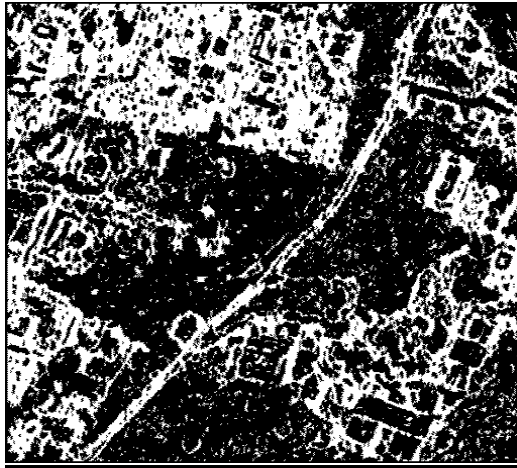
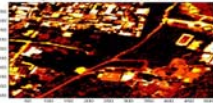
Homogeneidad3



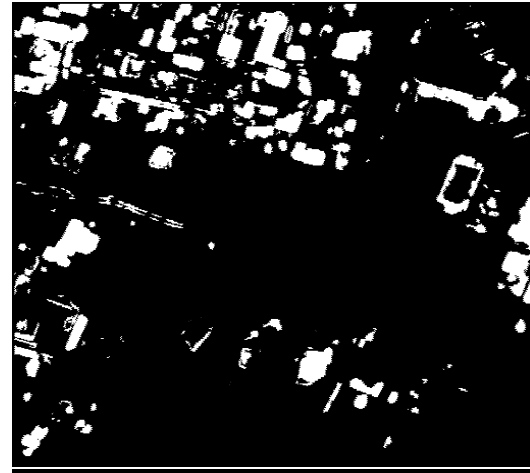
Para la obtención de las siguientes imágenes solamente se utilizó uno de los métodos que mejores resultados arrojó, por lo tanto, solamente se utilizó el procedimiento supervisado mediante diferentes umbrales para el proceso de la binarización en la siguiente imagen

<p>Imágenes obtenidas mediante la modificación de umbrales para cada descriptor, para obtener 2 clases con niveles de de gris 0 y 255</p> <p><u>Autocorrelación</u></p> <p>Zonas urbanas (blancas) = 50-250 Zonas no urbanas (oscuras)=0-50</p> <p><u>Cluster prominence</u></p> <p>Zonas urbanas (blancas) = 20-221 Zonas no urbanas (oscuras)=0-20</p> <p><u>Cluster shade</u></p> <p>Zonas urbanas (blancas) = 50-225 Zonas no urbanas (oscuras)=0-40</p> <p><u>Contraste</u></p> <p>Zonas urbanas (blancas) = 30-225 Zonas no urbanas (oscuras)=0-30</p>	<p><u>Correlación</u></p> <p>Zonas urbanas (blancas) = 0-30 Zonas no urbanas (oscuras)=30-200</p> <p><u>Energía</u></p> <p>Zonas urbanas (blancas) = 120-255 Zonas no urbanas (oscuras)=0-120</p> <p><u>Entropía</u></p> <p>Zonas urbanas (blancas) = 0-150 Zonas no urbanas (oscuras)=150-220</p> <p><u>Varianza</u></p> <p>Zonas urbanas (blancas) = 50-250 Zonas no urbanas (oscuras)=0-50</p> <p><u>Máxima probabilidad</u></p> <p>Zonas urbanas (blancas) = 0-50 Zonas no urbanas (oscuras)=50-255</p>
---	--

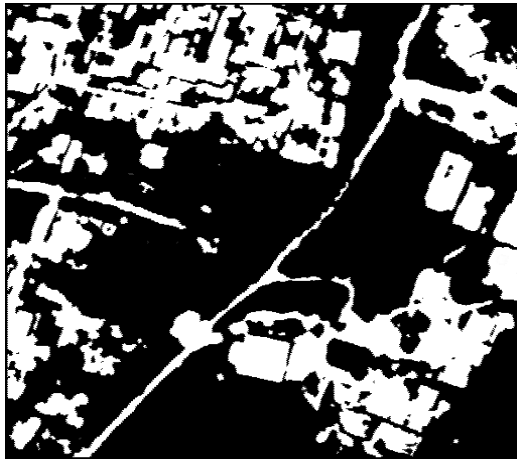




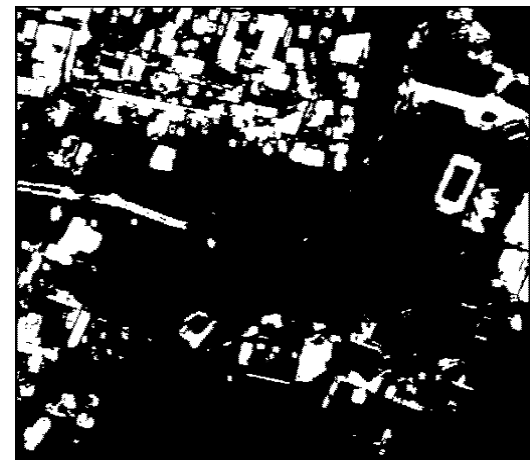
Máxima Probabilidad1b



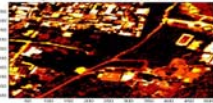
Varianza1b



Autocorrelación1b



Cluster prominenc1b



Homogenización de imágenes en 4 clases con diferentes umbrales para cada imagen.

Mapas Burdos :

0 negro
192 gris claro
128 gris fuerte
255 blanco

Autocorrelación4 y 5:

T>172 = 255 blanco
90 <=T<=170 = 192 gris claro
89 >=T>=50 y T<40 = 128 gris oscuro
= 0 negro

Varianza4 y 5:

T>90 = 255 blanco
T>=50 y T<=90 = 192 gris claro
T>=21 y T<=52 = 128 gris oscuro
T<21 = 0 negro

Cluster shade4 y 5:

T>120 = 255 blanco
120>=T>=80 = 192 gris claro
79>=T>=45 = 128 gris oscuro
T<45 = 0 negro

Dissimilarity4 y 5:

T>140 = 255 blanco
139>=T>=80 = 192 gris claro
T>=50 y T<=79 = 128 gris obscuro
T<50 = 0 negro

Energia4 y 5:

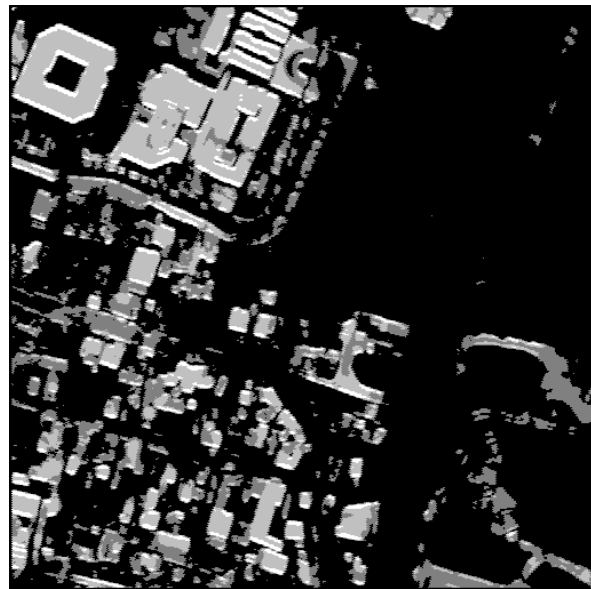
T>172 = 255 blanco
170>=T>=90 = 192 gris claro
89>=T>=41 = 128 gris oscuro
T<40 = 0 negro

Entropia4 y 5:

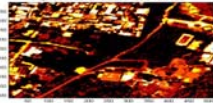
T>190 = 255 blanco
190>=T>=149 = 192 gris claro
148>=T>=66 = 128 gris oscuro
T<65 = 0 negro



Autocorrelacion4



Cluster shade4



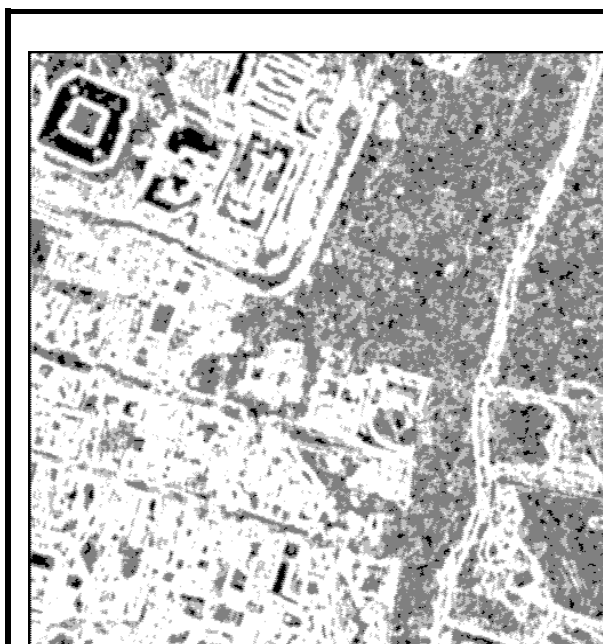
Las siguientes imágenes se presentan en un formato más grande con el fin de observar con mejor claridad las zonas, debido a su importancia en el proceso de fusión.



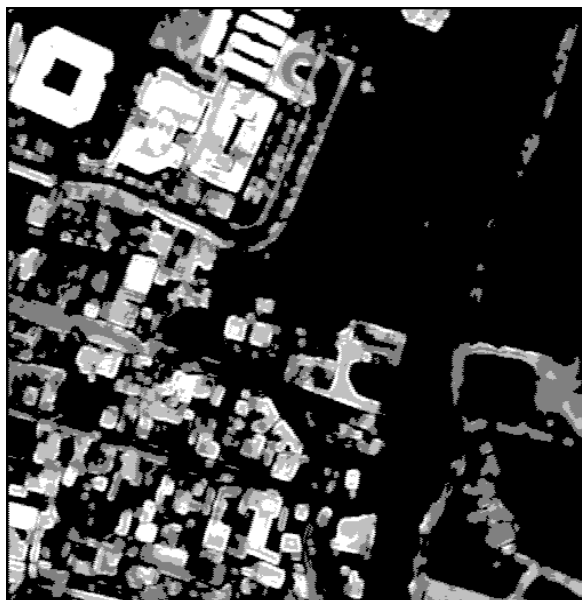
Dissimilarity4



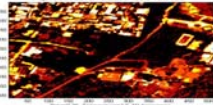
Energia4



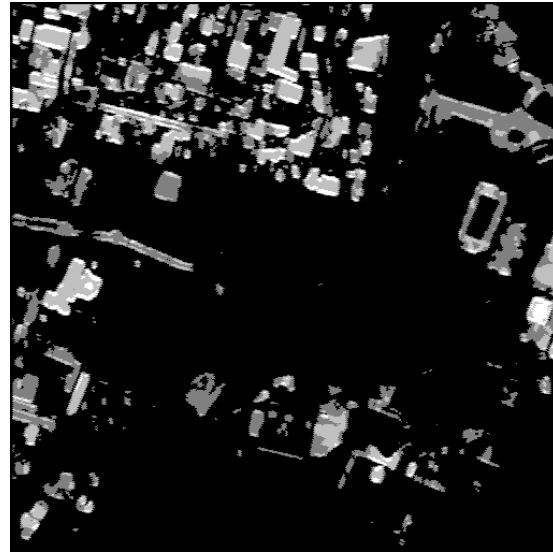
Entropia4



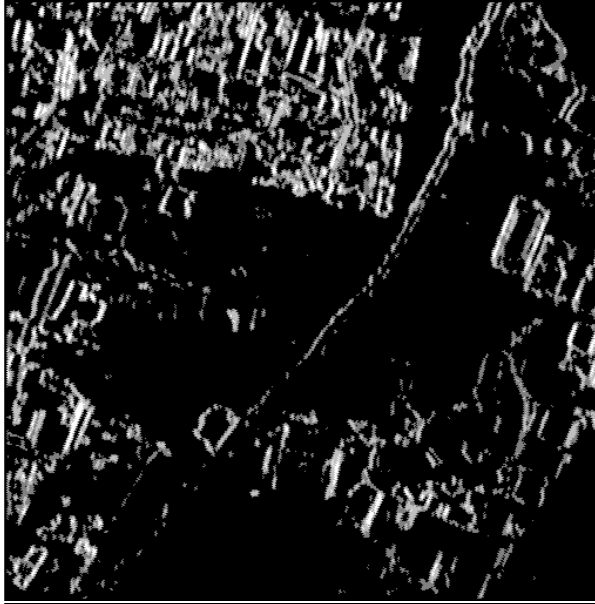
Varianza4



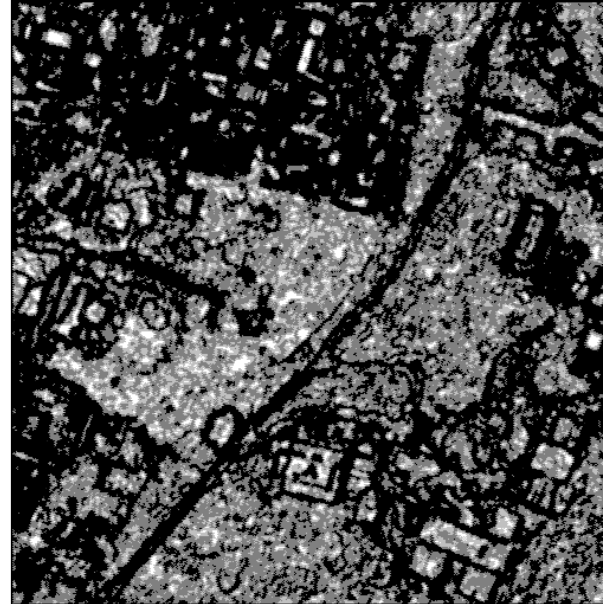
Autocorrelacion5



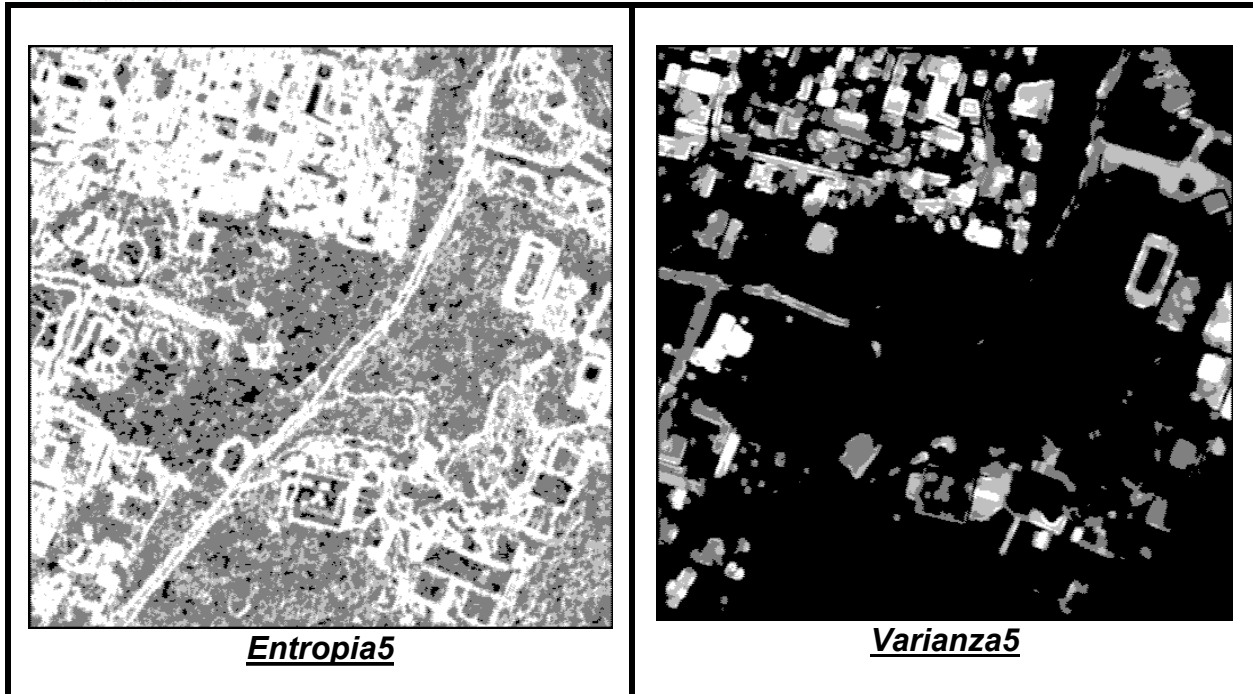
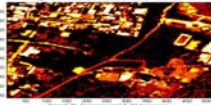
Cluster shade5



Dissimilarity5



Energia5

**7.3****Resultados del proceso de fusión binaria y multiclases**

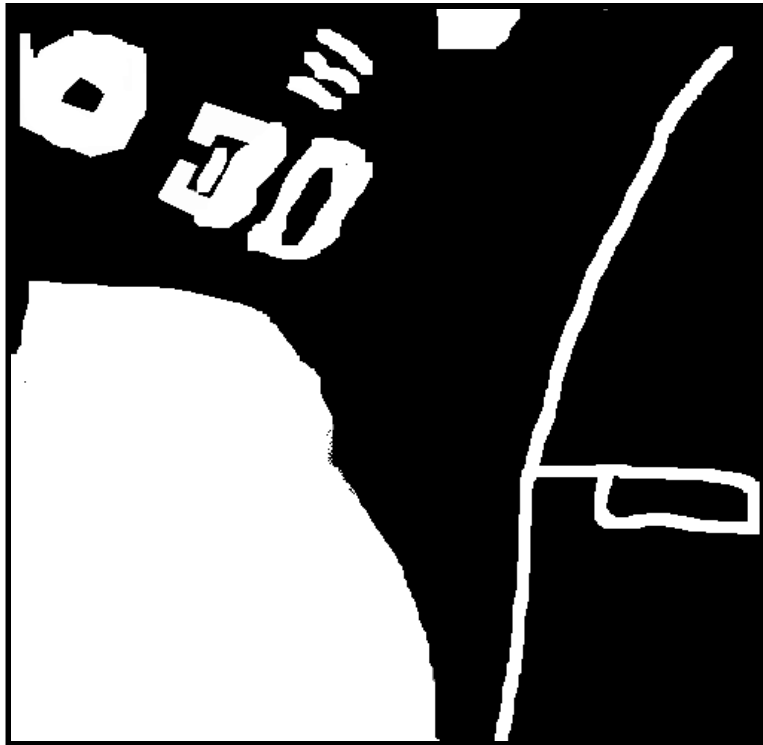
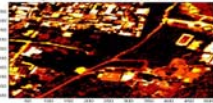
Las imágenes que a continuación se muestran presentan los resultados del modelo de fusión propuesto para el caso binario y multiclases.

En estas imágenes se pueden observar las diferencias obtenidas, procedentes de una selección adecuada de los descriptores homogéneos a fusionar, y de los mapas burdos seleccionados.

Como se puede observar tanto los criterios de umbralaje y los mapas burdos cumplen un papel sumamente importante en el resultado de la fusión, debido a que estos pueden limitar o ampliar los intervalos de fusión, por lo tanto, nos puede favorecer o entorpecer el proceso, es decir:

En la fusión binaria un mapa burdo muy delimitado nos puede favorecer en la clasificación, despreciando las zonas de menor importancia y enfatizando las zonas de interés.

En la fusión multiclases un mapa burdo muy delimitado nos puede perjudicar en la clasificación, ya que se puede perder información de suma importancia, debido al manejo de 4 clases con diferentes características entre sí, pero sí nuestro propósito es delimitar ciertas zonas y despreciar otras se pueden utilizar umbrales y mapas menos específicos.

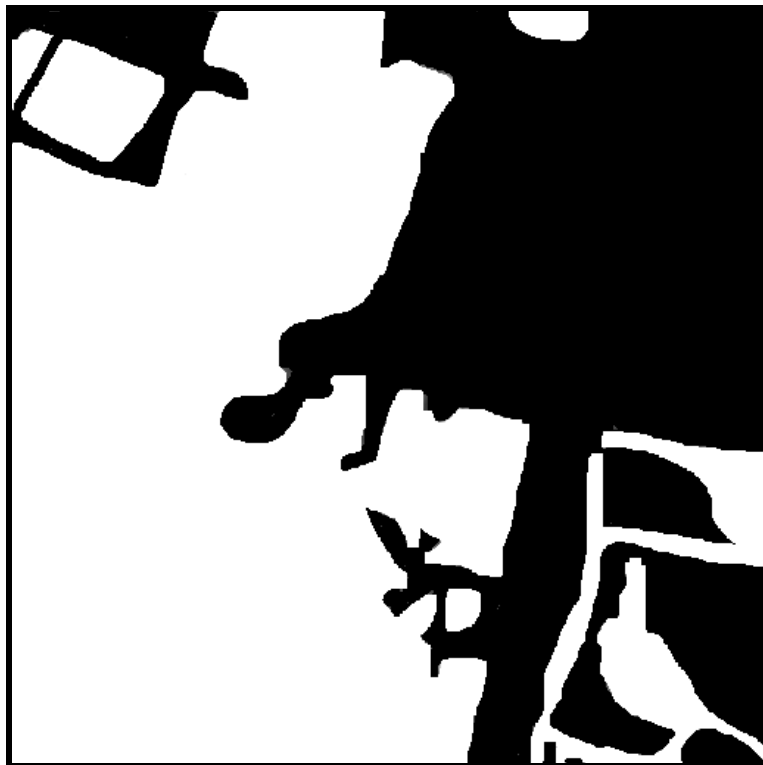


Mapa A

**Mapa burdo de
512x512 píxeles
Con niveles de gris
De 0 y 255**

**Zonas urbanas =
255 niveles de gris
Áreas de color blanco**

**Zonas no urbanas
(Áreas verdes, etc.) =
0 niveles de gris
Áreas de color negro**

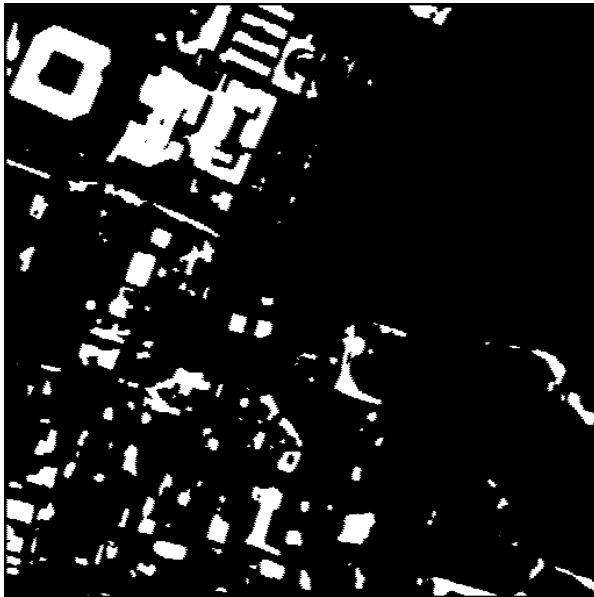
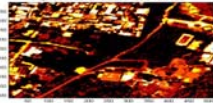


Mapa B

**Mapa burdo de
512x512 píxeles
Con niveles de gris
De 0 y 255**

**Zonas urbanas =
255 niveles de gris
Áreas de color blanco**

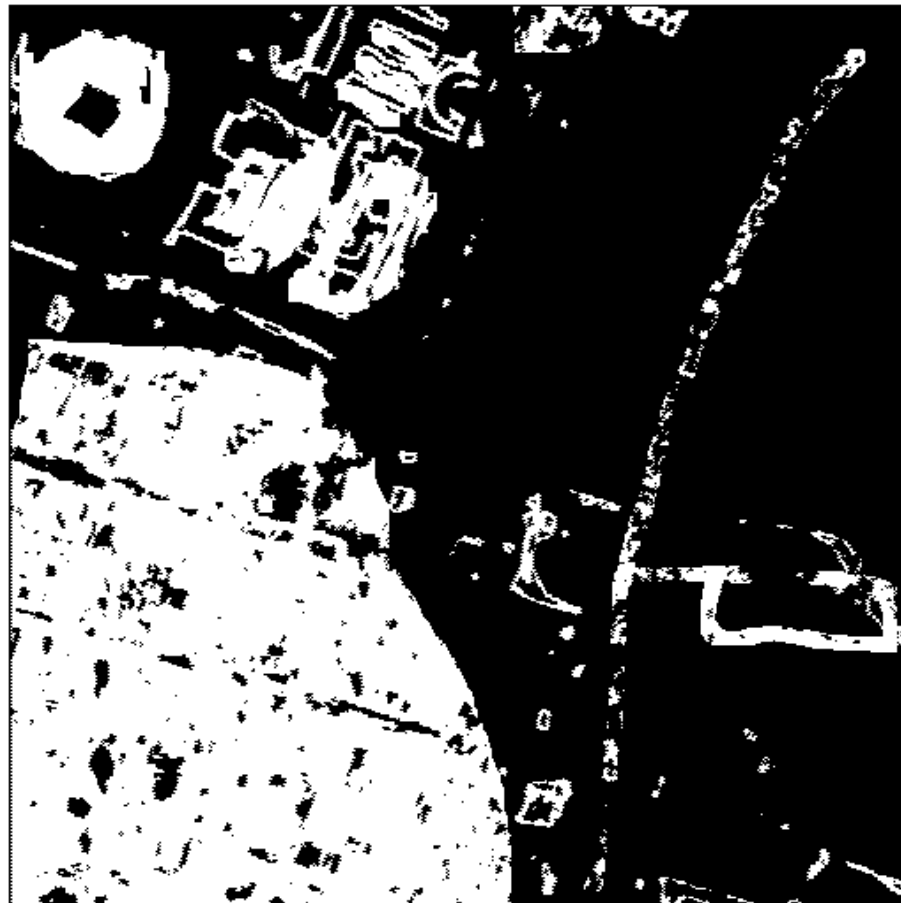
**Zonas no urbanas
(Áreas verdes, etc.) =
1 niveles de gris
Áreas de color negro**



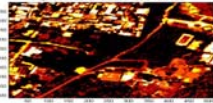
Autocorrelacion3



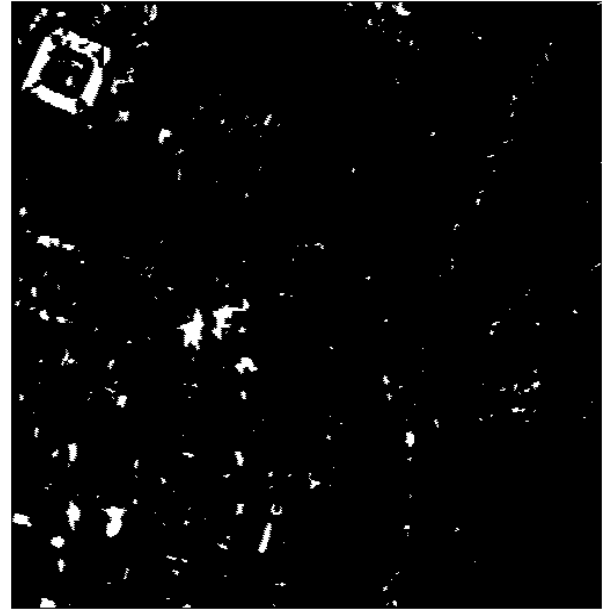
Entropia3



Resultado de la fusión entre las imágenes: mapa A -Entropia3 - Autocorrelacion3



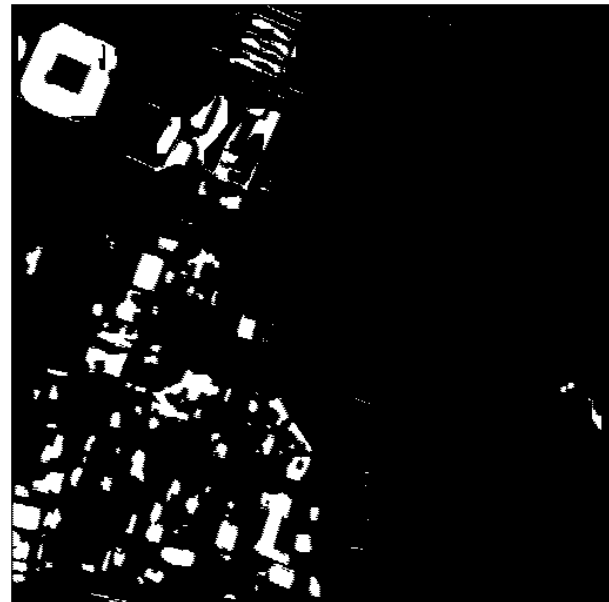
Resultado de la fusión entre las imágenes:
mapa A, Autocorrelación1 y Contraste1



Resultado de la fusión entre las imágenes:
mapa A, Energía3 y Correlacion3



Resultado de la fusión entre las imágenes:
mapa A, Entropia3 y Dissimilarity3



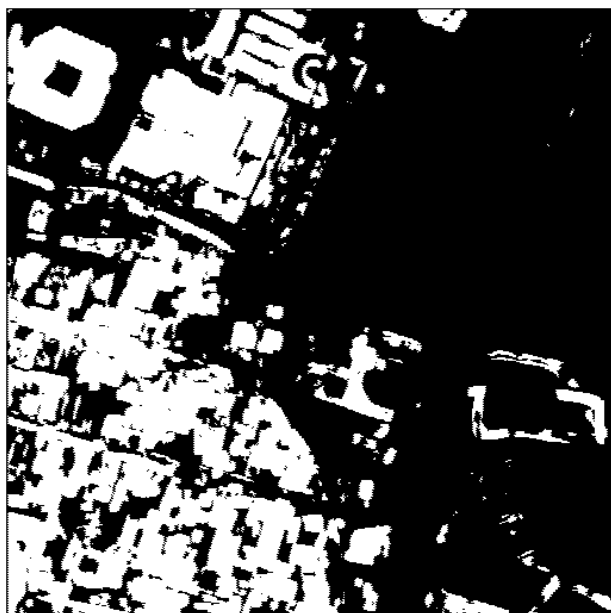
Resultado de la fusión entre las imágenes:
mapa A, Autocorrelacion3 y Cluster shade3



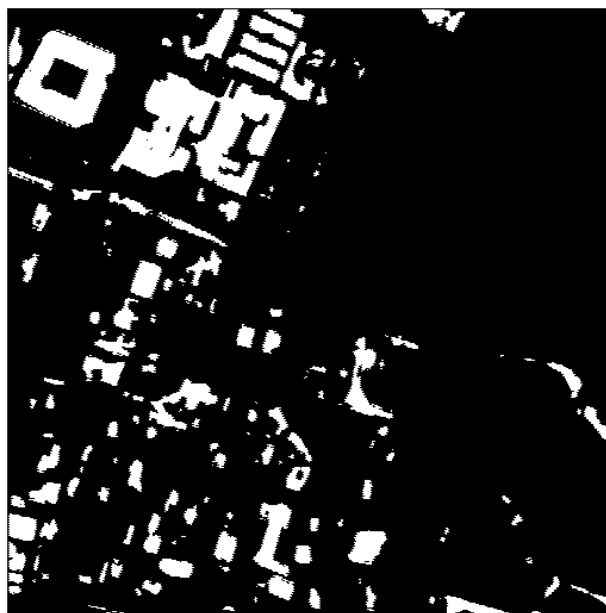
Resultado de la fusión entre las imágenes:
mapa A, Entropía1 y Correlación1



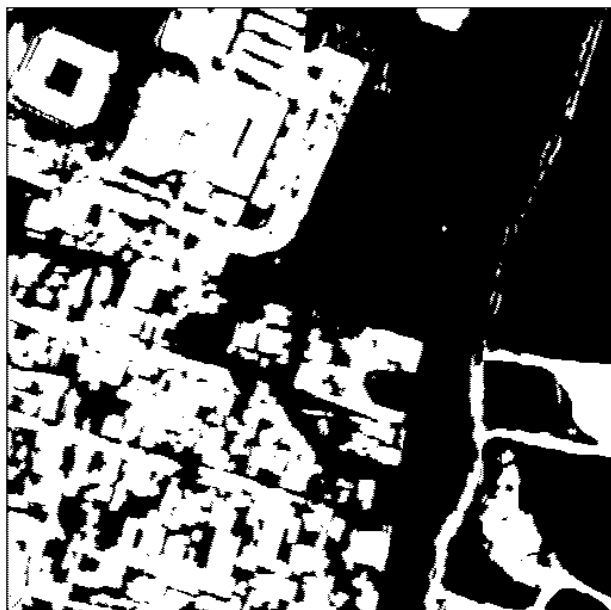
Resultado de la fusión entre las imágenes:
mapa A, Maxprobabilidad3 y Homogeneidad3



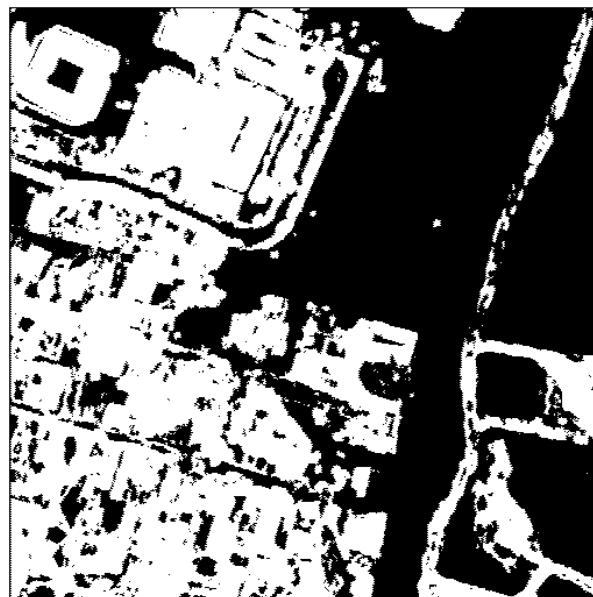
Resultado de la fusión entre las imágenes:
mapa A, Varianza1 y Correlación1



Resultado de la fusión entre las imágenes:
mapa B, Autocorrelacion3 y Cluster shade3



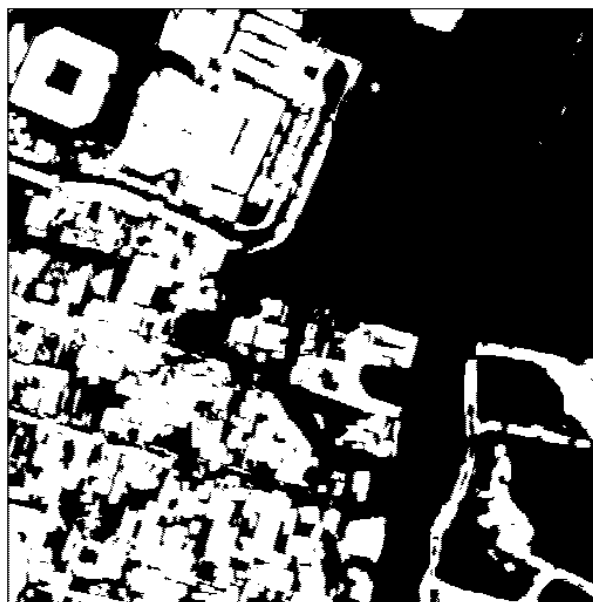
Resultado de la fusión entre las imágenes:
mapa B, Contraste1 y Autocorrelación1



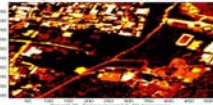
Resultado de la fusión entre las imágenes:
mapa B, Correlación1 y Entropía1



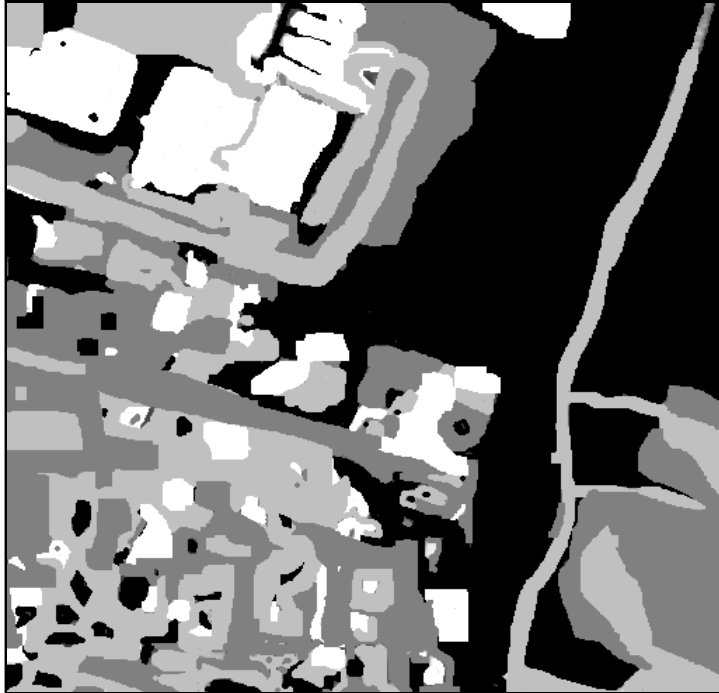
Resultado de la fusión entre las imágenes:
mapa B, Entropía3 y Autocorrelacion3



Resultado de la fusión entre las imágenes:
mapa B, Varianza1 y Correlación1



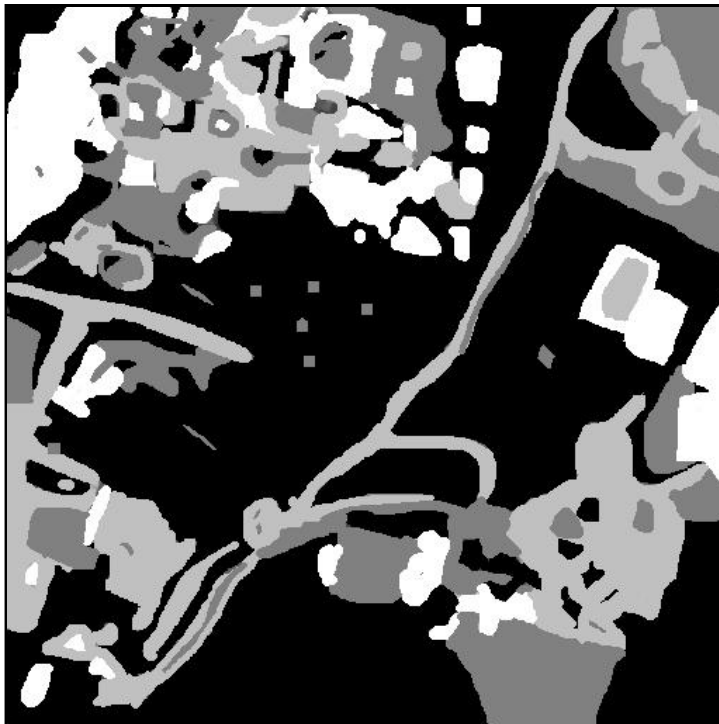
En las siguientes imágenes se muestran los mapas burdos con 4 clases utilizados en la fusión multiclase.



Mapa C

Mapa burdo de 512 x 512 píxeles con 4 clases y niveles de gris:

- 0 negro
- 192 gris claro
- 128 gris fuerte
- 255 blanco



Mapa D

Mapa burdo de 512 x 512 píxeles con 4 clases y niveles de gris:

- 0 negro
- 192 gris claro
- 128 gris fuerte
- 255 blanco



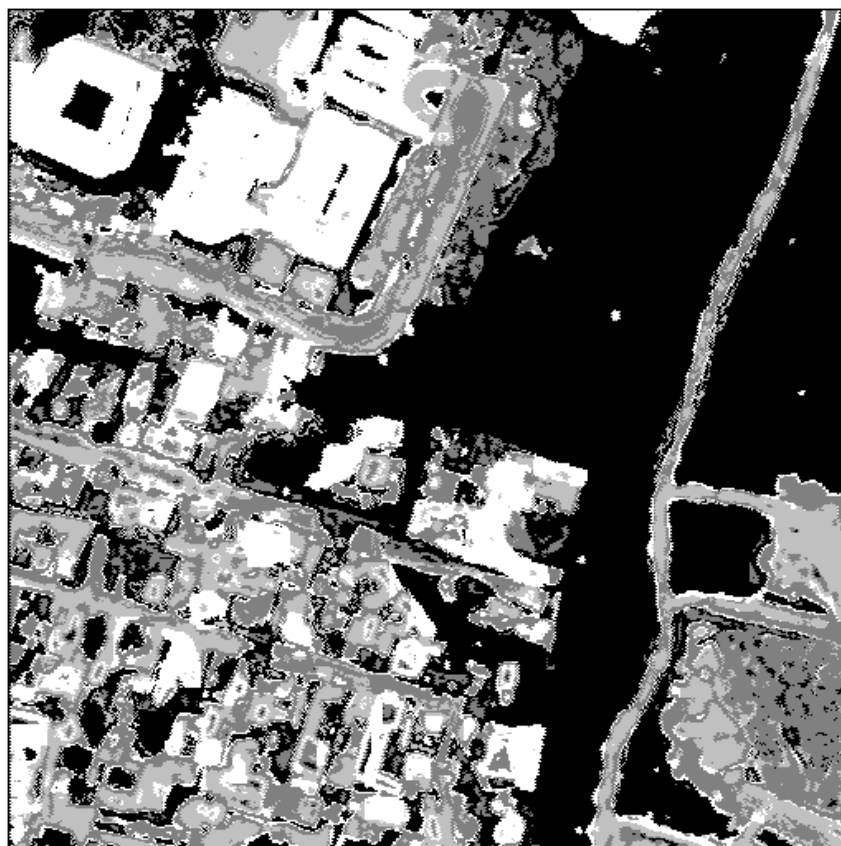
En las siguientes imágenes se muestran los resultados del proceso de fusión multiclases para 2 imágenes diferentes.



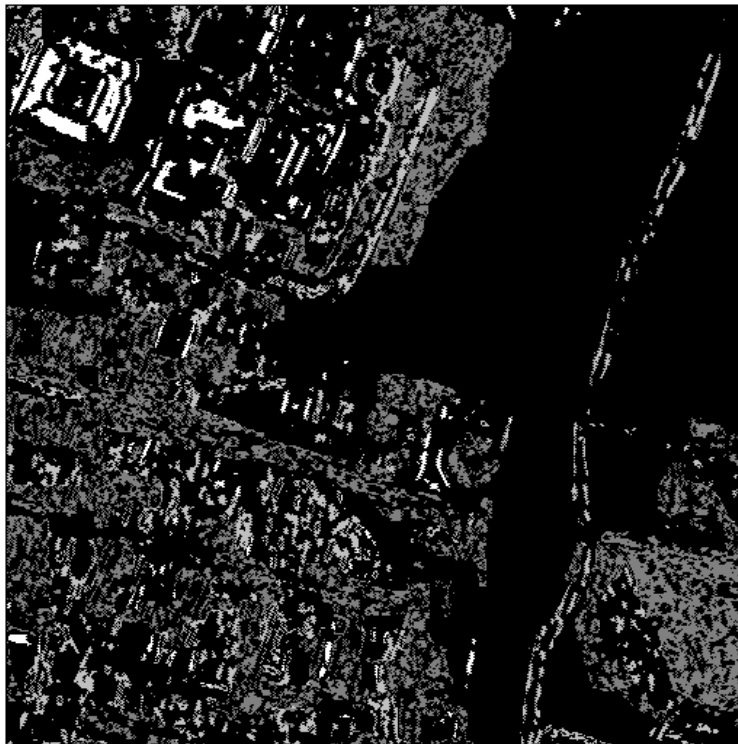
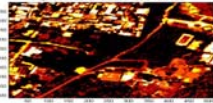
Autocorrelacion4



Entropia4



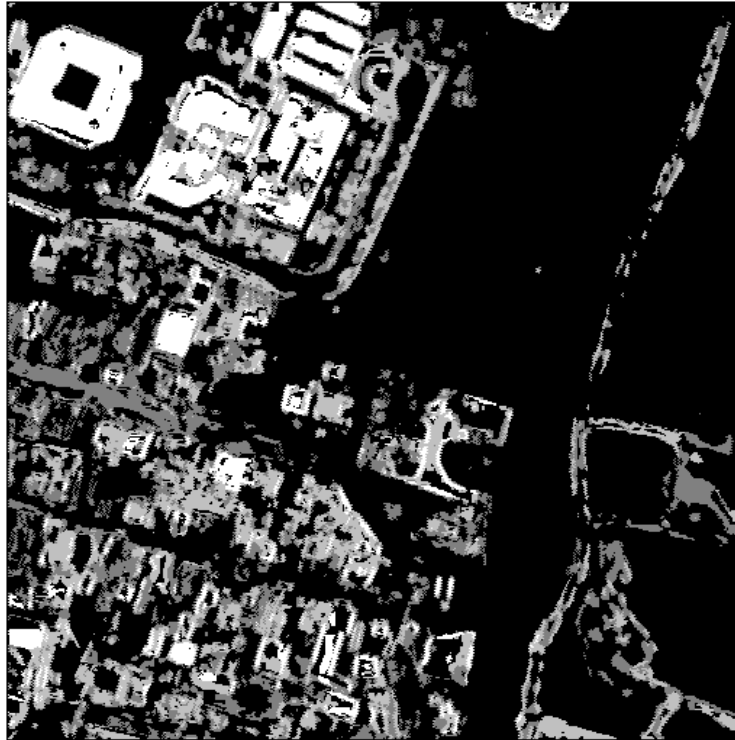
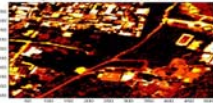
**Resultado de la fusión entre las imágenes:
mapa C , Autocorrelación4 y Entropía4**



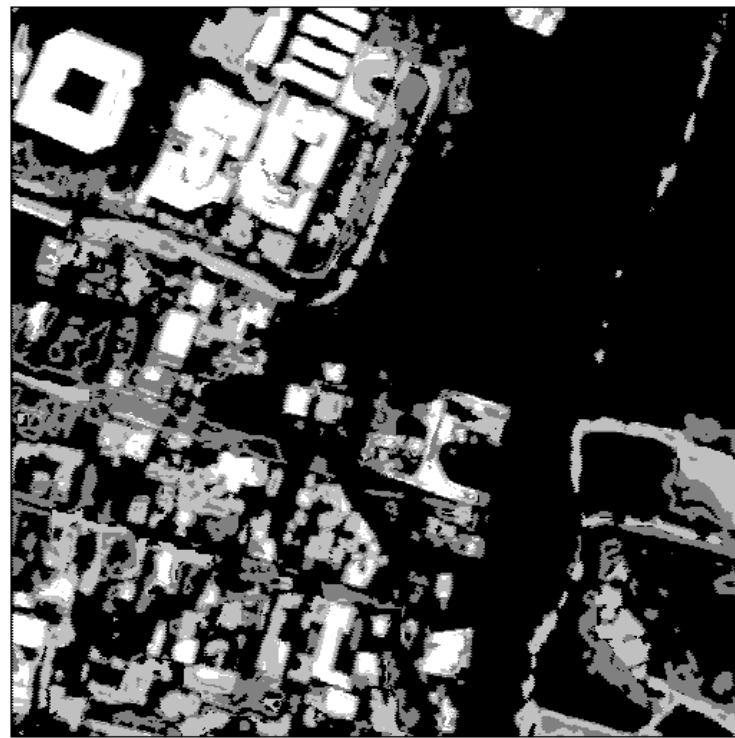
Resultado de la fusión entre las imágenes: mapa C - Dissimilarity4 - Energia4



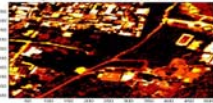
Resultado de la fusión entre las imágenes: mapa C - Dissimilarity4 - Entropia4



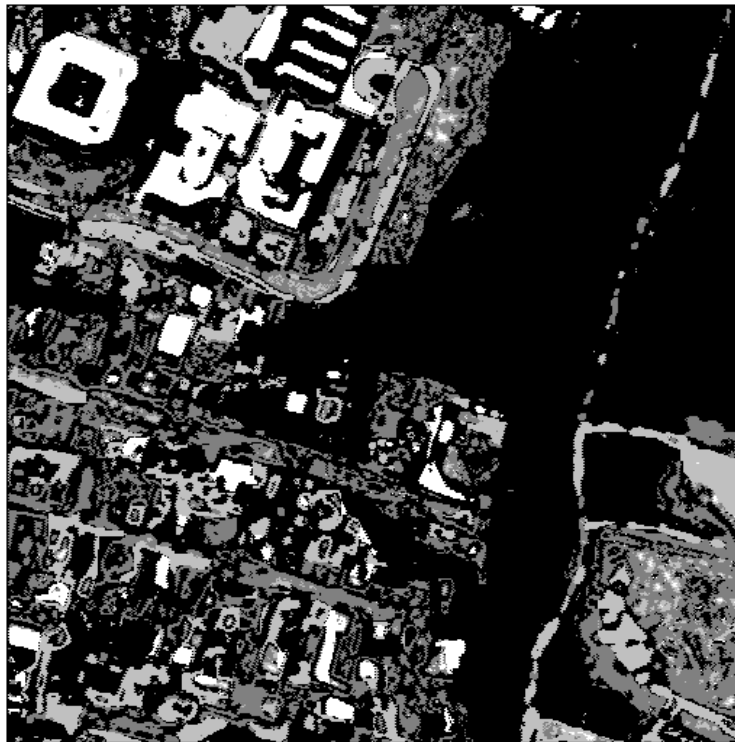
Resultado de la fusión entre las imágenes: mapa C - Dissimilarity4 - Varianza4



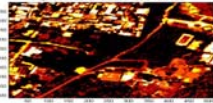
Resultado de la fusión entre las imágenes: mapa C - Autocorrelacion4 – Cluster shade4



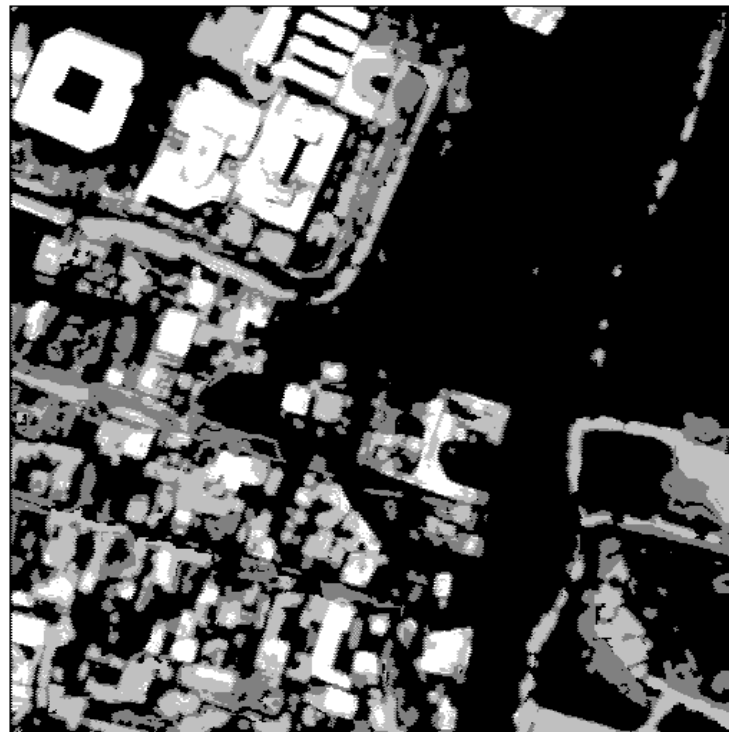
Resultado de la fusión entre las imágenes: mapa C – Autocorrelacion4 – Dissimilarity4



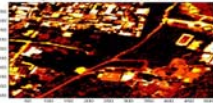
Resultado de la fusión entre las imágenes: mapa C – Autocorrelacion4 – Energia4



Resultado de la fusión entre las imágenes: mapa C - Dissimilarity4 – Cluster shade4



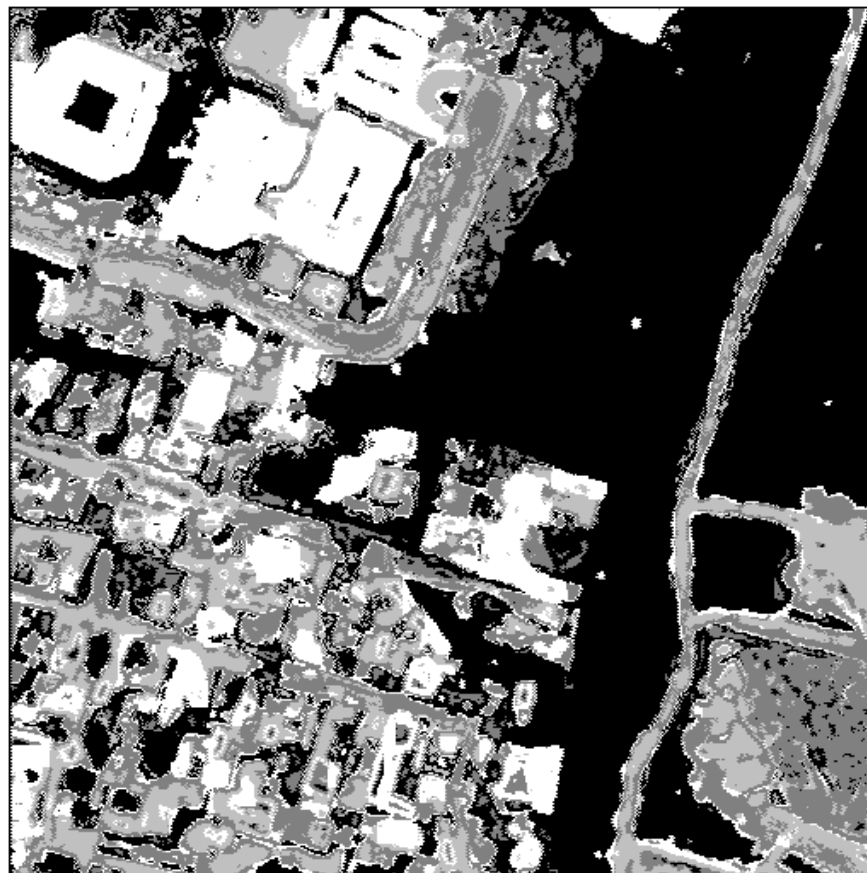
Resultado de la fusión entre las imágenes: mapa C – Autocorrelación4 – Varianza4



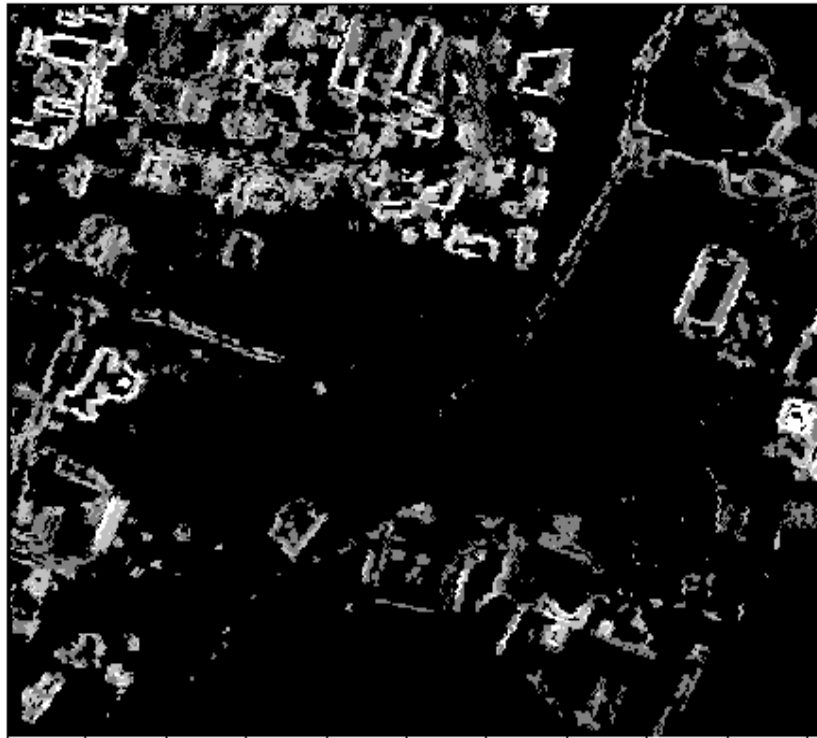
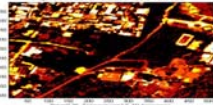
Autocorrelacion4



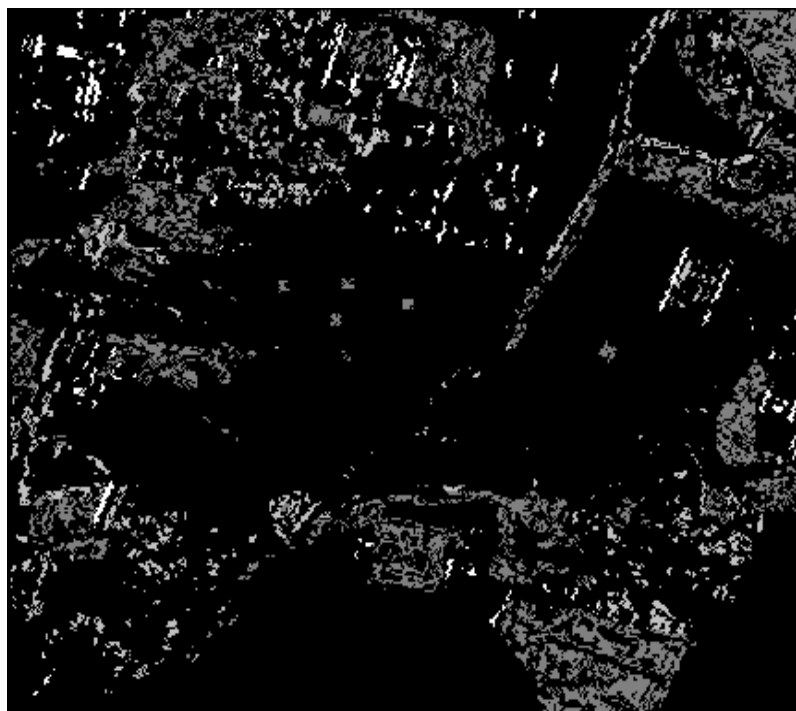
Entropia4



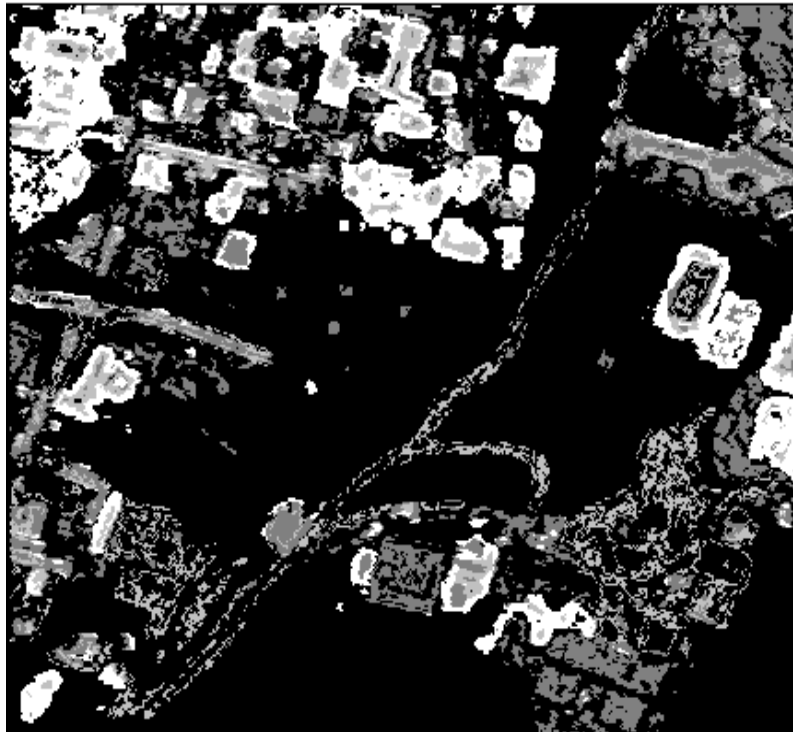
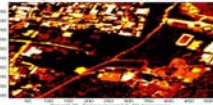
Resultado de la fusión entre las imágenes:
mapa C, Autocorrelación4 y Entropia4



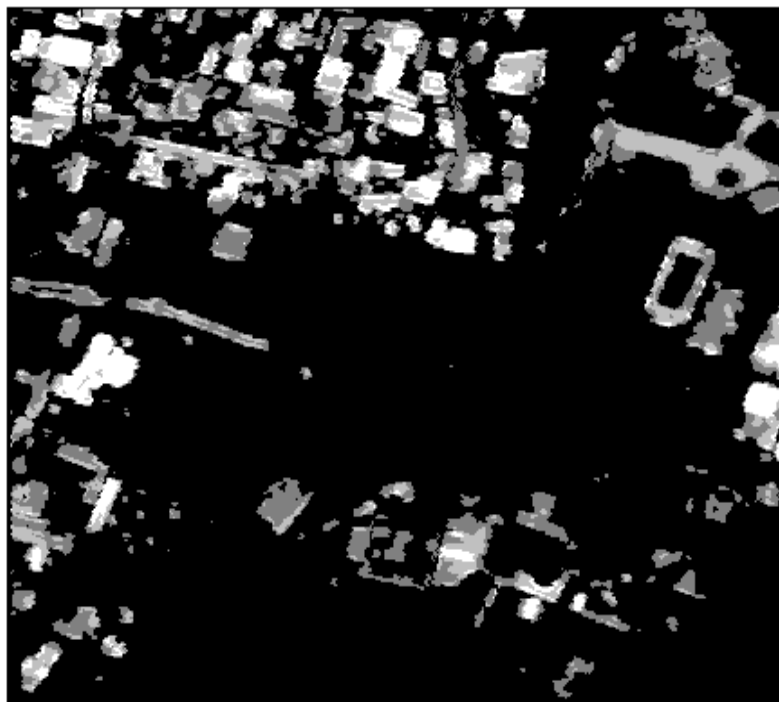
Resultado de la fusión entre las imágenes: mapa D – Dissimilarity5 – Cluster shade5



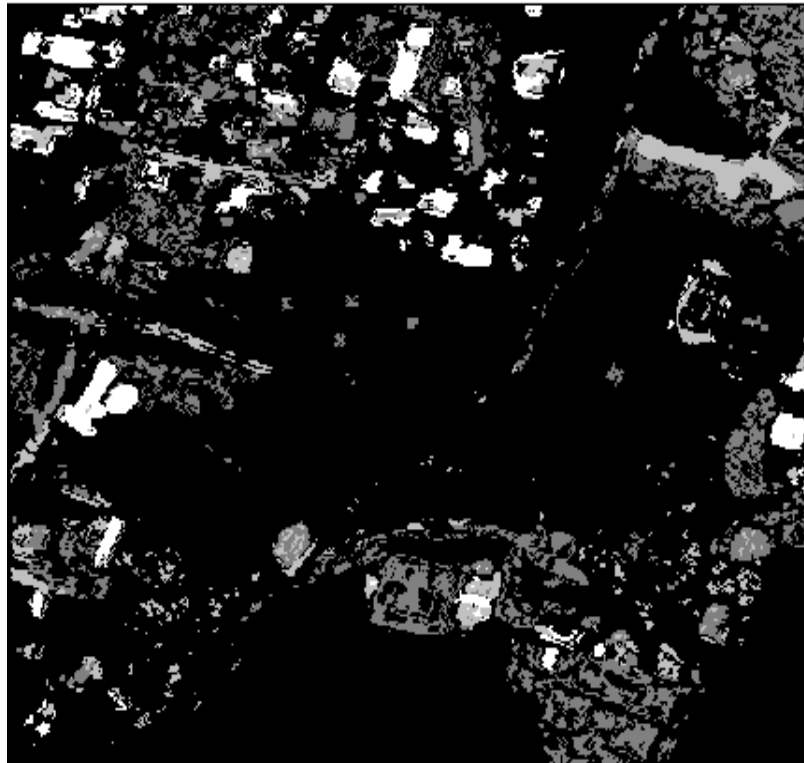
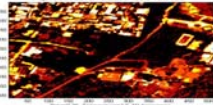
Resultado de la fusión entre las imágenes: mapa D - Energia5 – Dissimilarity5



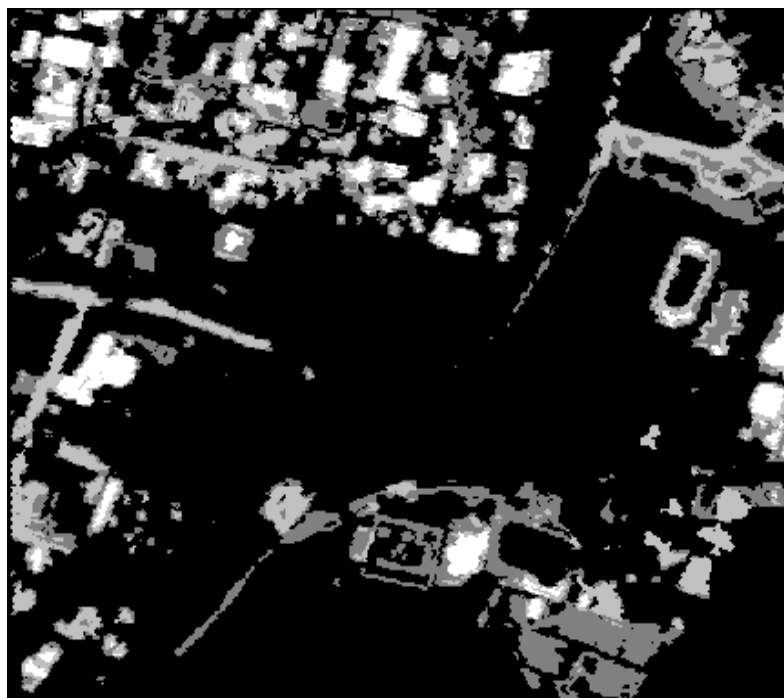
Resultado de la fusión entre las imágenes: mapa D – Entropía5 – Cluster shade5



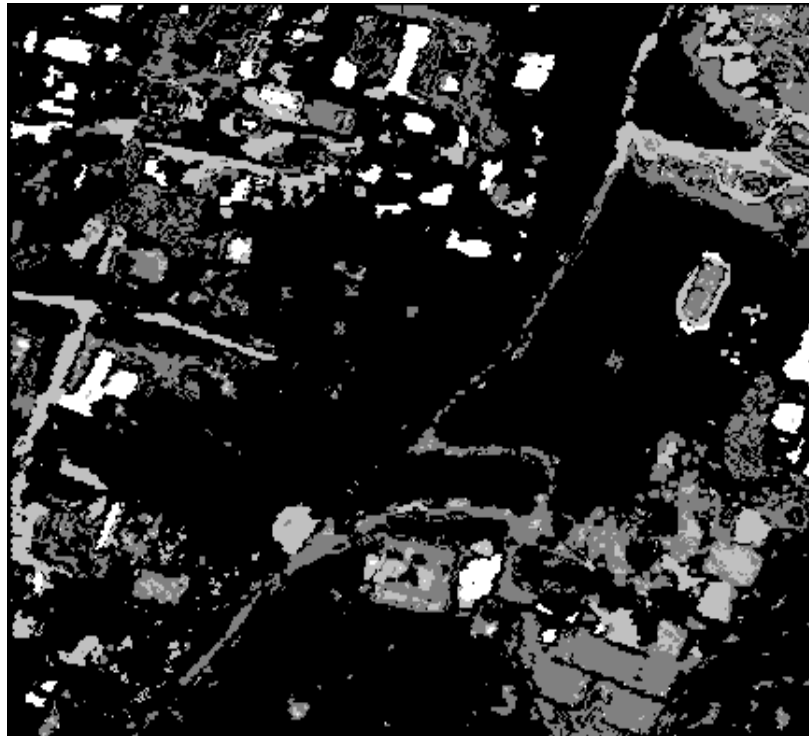
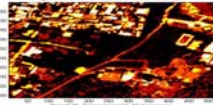
Resultado de la fusión entre las imágenes: mapa D – Varianza5 – Cluster shade5



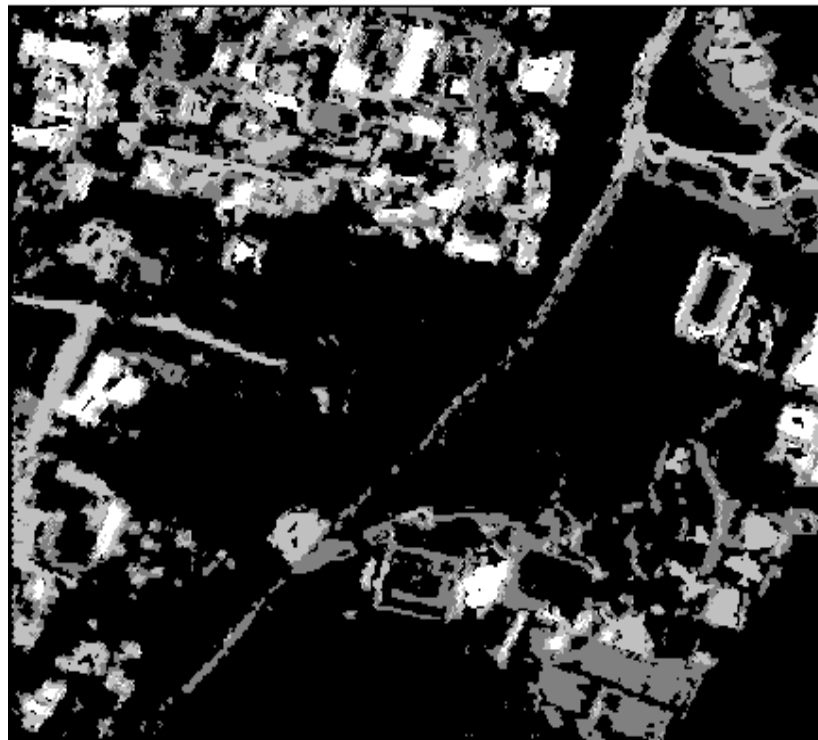
Resultado de la fusión entre las imágenes: mapa D – Varianza5 – Energía5



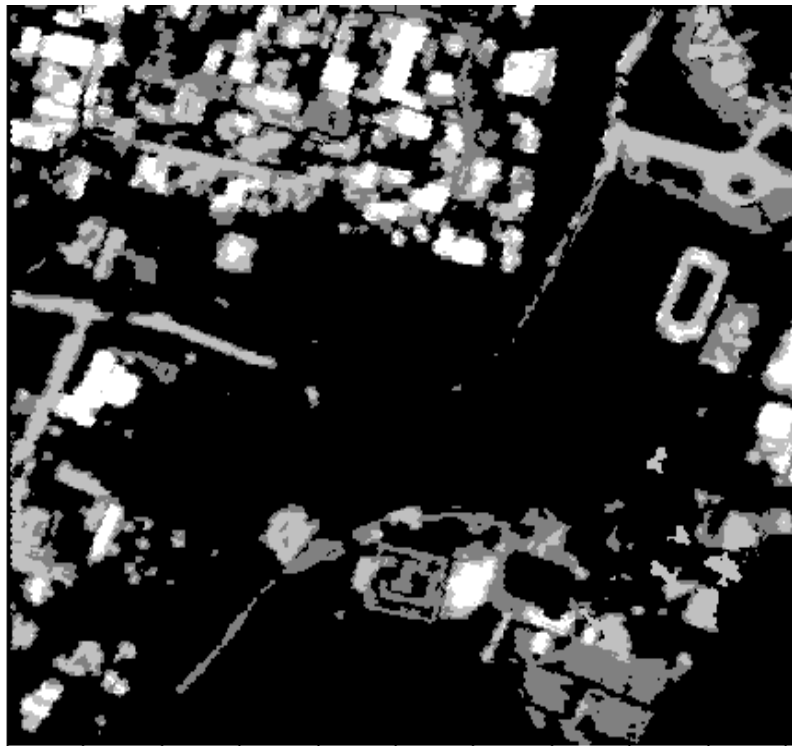
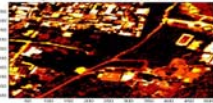
Resultado de la fusión entre las imágenes: mapa D – Autocorrelación5 – Cluster shade5



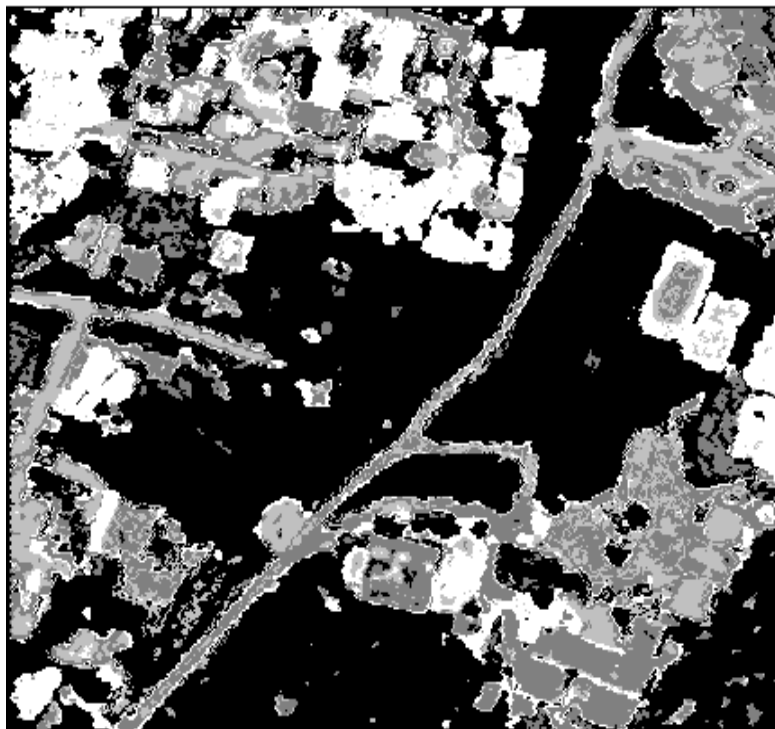
**Resultado de la fusión entre las imágenes: mapa D – Autocorrelación5 –
Energía5**



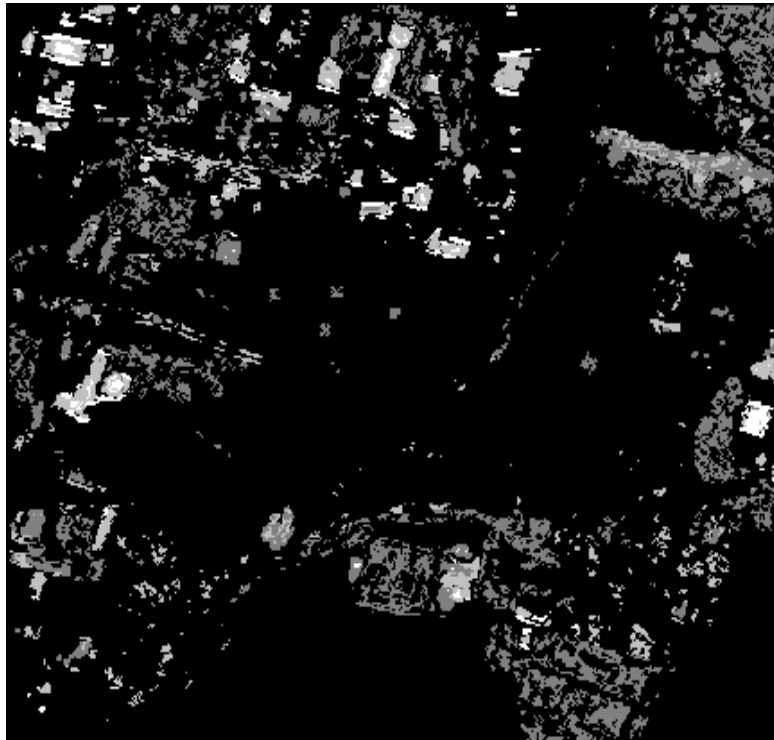
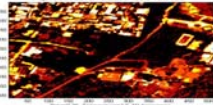
**Resultado de la fusión entre las imágenes: mapa D – Autocorrelación5 –
Dissimilarity5**



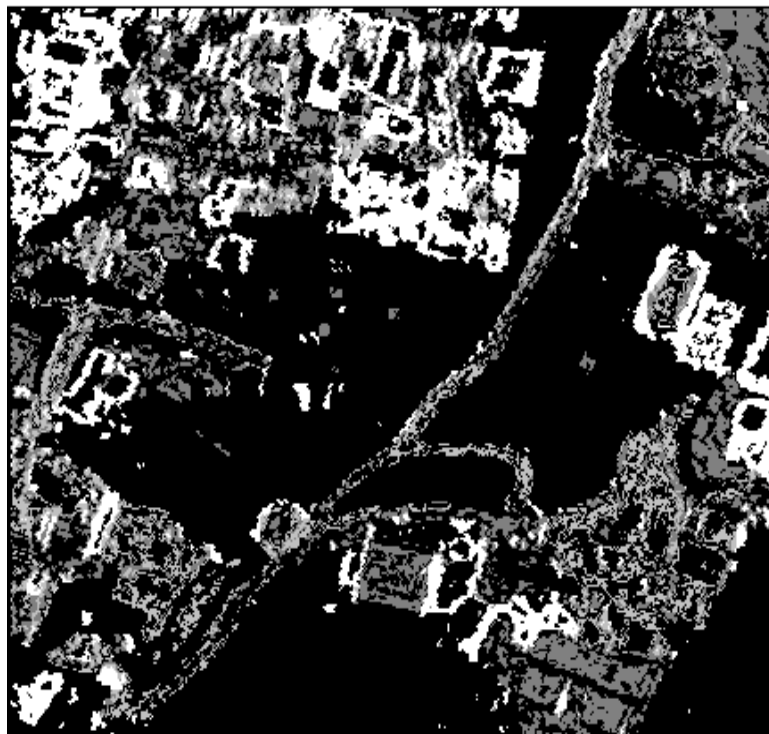
**Resultado de la fusión entre las imágenes: mapa D – Autocorrelación5 –
Varianza5**



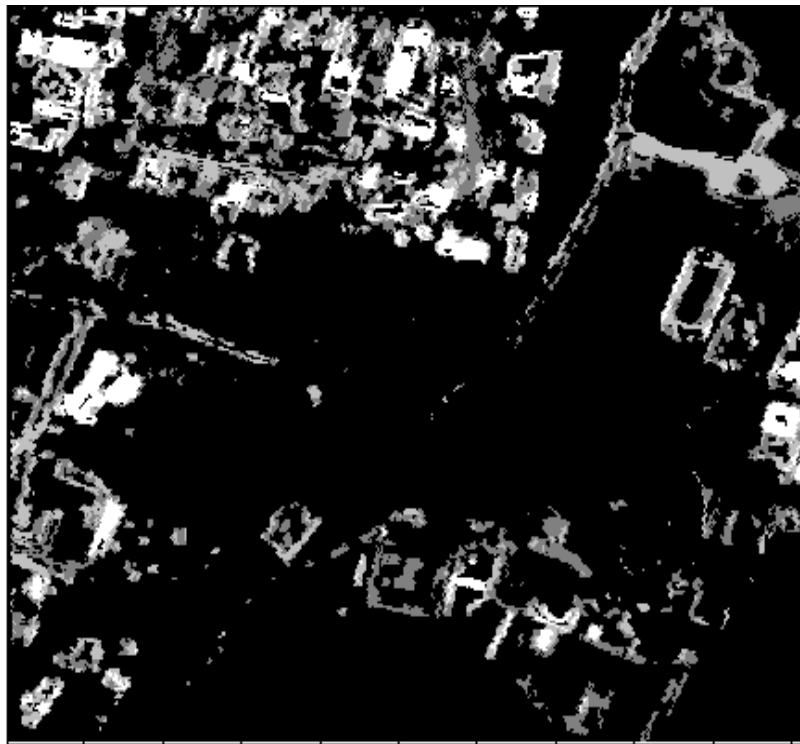
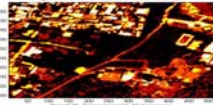
**Resultado de la fusión entre las imágenes: mapa D – Autocorrelación5 –
Entropía5**



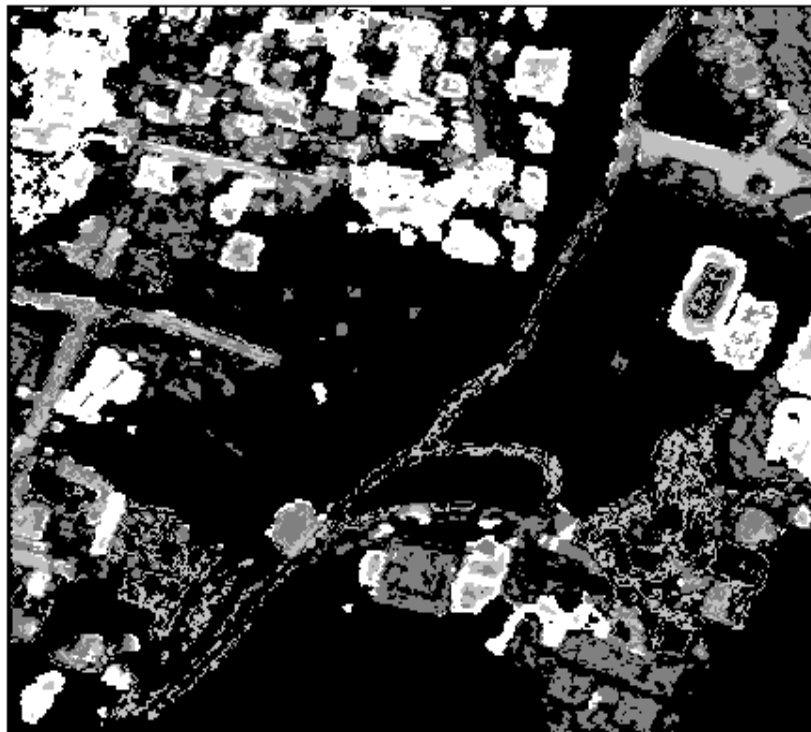
Resultado de la fusión entre las imágenes: mapa D – Energía5 – Cluster shade5



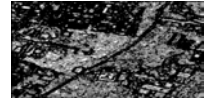
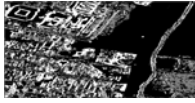
Resultado de la fusión entre las imágenes: mapa D – Entropía5 – Dissimilarity5



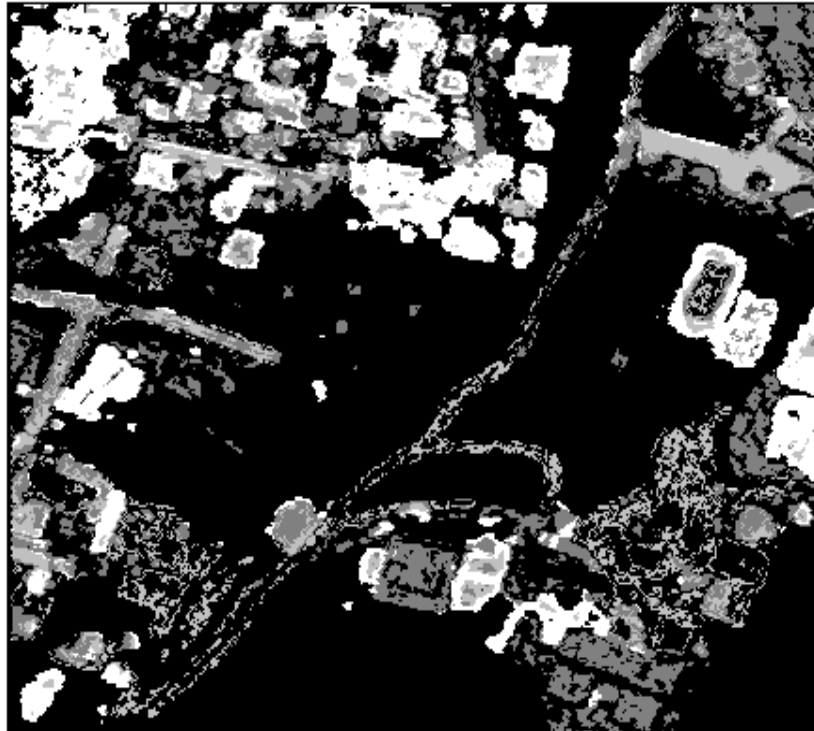
Resultado de la fusión entre las imágenes: mapa D – Varianza5 – Dissimilarity5

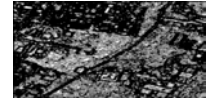
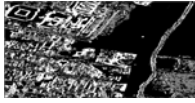


Resultado de la fusión entre las imágenes: mapa D – Varianza5 – Entropia5



8 Conclusiones





8.0 Conclusiones

El tema principal de este trabajo se fundamenta en el desarrollo de un modelo de fusión de imágenes, con el fin de realizar una segmentación y recuperación de información característica de ciertas áreas de la Ciudad de México, con el objeto de proporcionar herramientas para el análisis del desarrollo demográfico del país.

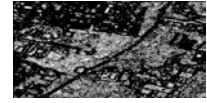
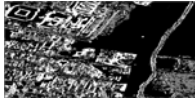
Con ayuda de la percepción remota y el procesamiento digital de imágenes se puede extraer información del mundo exterior, sin tener que estar presentes, por medio de sensores que obtienen fotografías satelitales y en nuestro caso de imágenes aéreas de alta resolución. La fotografía aérea proporciona información que visualmente podemos distinguir, por ejemplo, los patrones texturales que corresponden a los bosques, pero que difícilmente pueden discriminarse a partir de simples análisis a nivel píxel. Tal es la razón de emplear en la presente tesis datos estadísticos texturales que puedan representar más adecuadamente a los diversos elementos de escenas urbanas, tales como bosques, parques y en general zonas urbanas y no urbanas.

Como primer punto podemos decir que gracias a los descriptores de textura, obtenidos por medio de un algoritmo rápido para el cálculo de la matriz de co-ocurrencia, se pudo extraer información importante de algunos rasgos presentes en las imágenes aéreas. Por medio de este algoritmo se redujo en gran medida el tiempo de ejecución del cálculo de la matriz de co-ocurrencia y la obtención de sus descriptores, reduciendo el tiempo aproximadamente en un 50% en relación a métodos tradicionales, tomando en cuenta la reducción a 32 niveles de gris de la imagen original y la capacidad de la maquina empleada, que contaba con un procesador Pentium IV a 1.6 Mhz y 256 Mb en ram.

Al homogeneizar las imágenes de textura de forma supervisada se tuvieron que tomar en cuenta ciertos criterios de umbralaje, los cuales definieron los resultados arrojados por este proceso. El único inconveniente de este procedimiento es el tiempo de ejecución, ya que por ser supervisado emplea un mayor tiempo de usuario, debido a que éste debe de utilizar su criterio de selección, en lo que se refiere a lo que se puede y no despreciar de la información contenida en las escenas. No obstante, este tiempo se ve recompensado al observar la calidad de los resultados.

Como punto final está el proceso de fusión de imágenes, el cual se basa en los esquemas propuestos por Shan Yu en sus artículos de “Urban area detection in satellite using map knowledge by a feedback control technique” [YSMB04] e “Improving satellite image análisis quality by data fusion” [YS04]. Mediante los fundamentos y criterios de la fusión binaria, en esta tesis se propuso un modelo de fusión multiclases, el cual emplea criterios similares al método de fusión binaria de S. Yu. Nuestro modelo fue diseñado para analizar y fusionar 4 clases presentes en las imágenes texturales.

Por medio de la fusión de 2 clases nos podemos dar cuenta de los rasgos globales que encierran las zonas urbanas y no urbanas en general. En la fusión de imágenes de 4 clases distintas se puede observar y distinguir con mayor exactitud las diferentes texturas y rasgos característicos presentes en las escenas, con lo cual se tiene un mejor alcance en la determinación de aspectos urbanos.



También se tiene que tomar en cuenta que en el proceso de homogeneización se requirió la aplicación de diferentes umbrales para cada imagen de textura, ya que éstas presentan rasgos característicos y únicos; para poder distinguirlos fue necesaria la aplicación de estos umbrales.

Como se puede observar en el capítulo de resultados, no fue necesaria la fusión de más de 2 imágenes, ya que con el procedimiento adecuado y una buena selección de imágenes se obtuvieron buenos resultados en el modelo de fusión de 4 clases.

Algunos de los parámetros empleados son mostrados a continuación. Se emplearon tres escenas de fotografía aérea con las siguientes características:

1. Tamaño 512 x 512 píxeles.
2. 256 niveles de gris.
3. Presentan visualmente 4 y 5 clases texturales distintivas.

En el cálculo de los descriptores de textura por medio de la GLCM se utilizaron los siguientes parámetros:

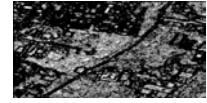
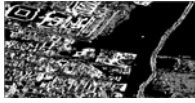
1. Reducción de 256 a 32 niveles de gris, con la menor pérdida de información posible
2. Desplazamiento horizontal para la GLCM de 2
3. Utilización de ventanas de 5x5
4. Número de clases presentes: 5

Se calcularon 11 descriptores de textura, los cuales fueron: autocorrelación, correlación, entropía, contraste, máxima probabilidad, homogeneidad, varianza, dissimilarity, cluster prominente, cluster shade y energía, no obstante no se utilizaron todos en el proceso de fusión, debido a que no arrojaban información relevante.

Para la binarización se emplearon los parámetros mostrados en la tabla 8.1, para cada descriptor de textura:

Autocorrelación (Umbrales en escala de grises)	Cluster prominence (umbrales)	Cluster shade (umbrales)
Zonas urbanas(blancas) = 50-250 no urbanas(oscuras) = 0-50	urbanas(blancas) = 20-221 no urbanas(oscuras) = 0-20	urbanas(blancas) = 50-225 no urbanas(oscuras) = 0-40
Contraste	Correlación	Energía
urbanas(blancas) = 30-225 no urbanas(oscuras) = 0-30	urbanas(blancas) = 0-30 no urbanas(oscuras) = 30-200	urbanas(blancas) = 120-255 no urbanas(oscuras) = 0-120
Entropía	Varianza	Máxima probabilidad
urbanas(blancas) = 0-150 no urbanas(oscuras) = 150-220	urbanas(blancas) = 50-250 no urbanas(oscuras) = 0-50	urbanas(blancas) = 0-50 no urbanas(oscuras) = 50-255

Tabla 8.1 Asignación de umbrales de binarización.



Para la asignación de umbrales en 4 clases se emplearon los parámetros mostrados en la tabla 8.2, para cada descriptor de textura:

Mapas burdos (niveles de grises)	Autocorrelación (T= umbral)	Varianza
0 negro 192 gris claro 128 gris fuerte 255 blanco	T>172 = 255 blanco T>=90 y T<=170 = 192 gris claro T>=50 y T<=89 = 128 gris oscuro T<40 = 0 negro	T>90= 255 blanco T>=50 y T<=90= 192 gris claro T>=21 y T<=52= 128 gris oscuro T<21 = 0 negro

Energía	Entropía
T >172 = 255 blanco T >=90 y T <=170 =192 gris claro T >=41 y T <=89 =128 gris oscuro T<40 = 0 negro	T >190 = 255 blanco T >=149 y T <=190 = 192 gris claro T >=66 y T <=148 = 128 gris oscuro T <65 = 0 negro
Dissimilarity	Cluster shede
T>140 = 255 blanco T>=80 y T<=139 = 192 gris claro T>=50 y T<=79 =128 gris oscuro T<50 = 0 negro	T>120 = 255 blanco T>=80 y T<=120 = 192 gris claro T>=45 y T<=79 =128 gris oscuro T<45= 0 negro

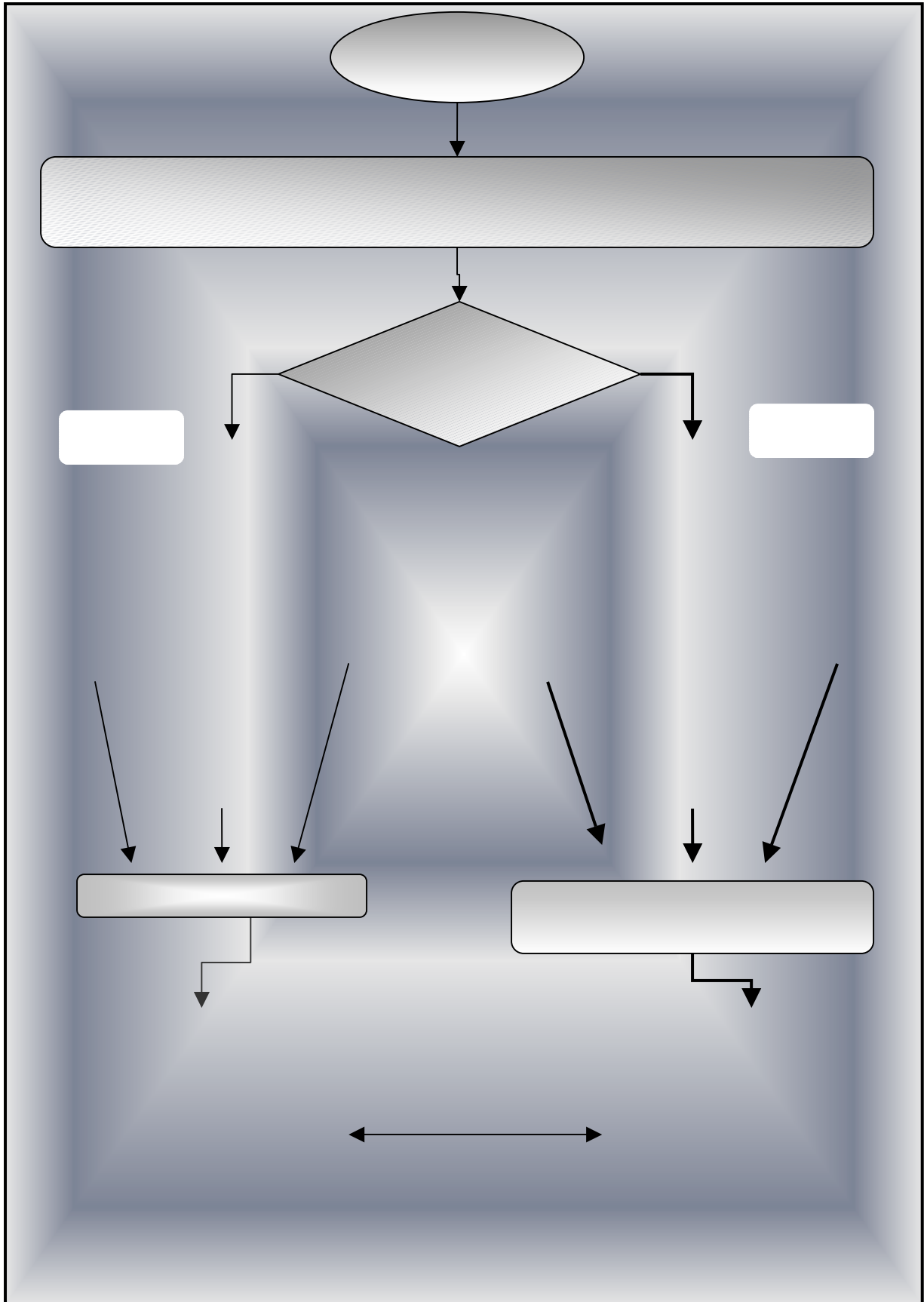
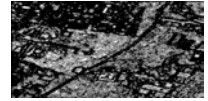
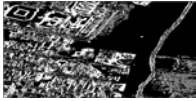
Tabla 8.2 Asignación de umbrales en 4 clases

En la presente tesis también se hace una comparación de resultados, entre el modelo empleado y técnicas de segmentación de imágenes por Markov, teoría de Bayes y el modelo de fusión binaria, concluyendo que el modelo empleado ofrece resultados más detallados y especiales. Los esquemas de procesamiento presentados pueden aportar información específica sobre regiones urbanas, y con ello se podría lograr de una forma más eficiente la planificación urbana.

Como se observa en los resultados, se puede concluir que los descriptores de textura no son lo suficientemente discriminantes en algunos casos, y en consecuencia los resultados no son muy discriminativos y clasificativos para ciertas texturas.

A partir de los resultados obtenidos en la presente tesis, diversos esquemas pueden desarrollarse, principalmente en la determinación de mejores operadores de textura. Consideramos que dada la funcionalidad del esquema de fusión de datos desarrollado, este puede aplicarse en los trabajos por emprender.

A continuación se presenta un diagrama que muestra las etapas de procesamiento entre la fusión binaria y la fusión multiclases desarrolladas en la presente tesis.





Bibliografía

[EC90] Chuvieco, E. 1990. Fundamentos de la Teledetección Espacial. Capítulo 6, Colección Monografías y Tratados GER, Ediciones Rialp, S.A., Madrid. 453 pp.

[RJ86] Richards, J. 1986. Remote Sensing Digital Image Analysis and Introduction. Edited by Springer-Verlag, Berlin Heidelberg, New York.. 281 pp.

[RCG96] Rafael C. González, Richard E. Woods, Digital Image Processing Addison Wesley publishing company Second Edition y ediciones Díaz S.A.1996. cap1,cap7.

[RMHS73] R. M. Haralick, K. Shanmugam and I. Dinstein ,Texture features for image classification , IEEE Trans. on Syst. Man Cybern., vol. 3, pp. 610-621, Nov. 1973.

[JB74] Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society, Series B, 36:192-236.

[JHMC71] Hammersley, J. M. and Clifford, P. (1971). Markov field on finite graphs and lattices unpublished.

[DJS77] Strauss, D. J. (1977). Clustering on colored lattice. Journal of Applied probability, 14:135-143.

[SGDG84] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. IEEE Transactions on PAMI, 6(6):721-741.

[MPRRPA00] M. Pujol , R. Rizo, P. Arques, Revista electrónica de visión por computador Número 4, octubre 2000, Aplicación de los modelos de campos aleatorios de Markov en visión artificial, Departamento de ciencias de la computación e inteligencia artificial Universidad de Alicante, Apartado 99. 03080 Alicante España.

[RC85] Chellappa, R. (1985). Two-dimensional discrete gaussian Markov random field models for image processing. In Kanal, L. N. and Rosenfeld, A., editors, Progress in pattern recognition 2, pages 79-112.

[EFA00] Eliud Fernández Aguirre Tesis: Análisis y Clasificación Semi-Automática de Escenas Urbans del Distrito Federal. Facultad de Ingeniería. México, D.F. 2002. pp 55-60, 103-123.

[BTZ] Sing-Tze Bow. Pattern Recognition and Image Preprocessing, Department of Electrical Engineering Northern Illinois University, Dekalb, Illinois. pp 61- 79.

[MMF] Dr. Miguel Moctezuma Flores, apuntes de señales aleatorias.



[FJAMM00] J. A. Franco, M. Moctezuma and F. Parmiggiani, Remote sensing images segmentation: A mixed approach adapted for fine structure detection, V Ibero-American symposium on pattern recognition, Siarp2000, Lisbon, pp. 393-404, September 11-13, 2000.

[SIRCJIMM00] S.I. Zavala, R. Cordero, J.I. Ferrera, M. Moctezuma and F. Parmiggiani, "Adaptive Segmentation on Synthetic Aperture Radar Images, International conference on telecommunications, ICT-2000, Acapulco, pp. 315-319, May 2000.

[FJMM99] J. A. Franco Bello, M. Moctezuma y F. Parmiggiani, Aproximación Contextual a la segmentación de fotografía aérea, IV simposio Iberoamericano de reconocimiento de patrones, SIARP99, La Habana Cuba, pp. 583-590, Marzo 21-26, 1999.

[DAC] David A. Clausi, Yongping Zhao, Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, Email : dclausi@uwaterloo.ca, Tel: (519) 888 – 4567 x2604, <http://www.elsevier.com/gej-ng/10/13/38/73/57/30/abstract.html>

[AJJ02] Ing. Alvaro Jiménez Jaramillo, Tesis: Análisis del comportamiento de descriptores de la matriz de co-ocurrencia en la extracción de características de textura en imágenes digitales. Universidad nacional de Colombia sede Manizales, Facultad de ingeniería y arquitectura, Ingeniería electrónica Manizales 2002, cap1, cap2.

[YS04] Shan YU, Improving satellite image analysis quality by data fusion. IEEE, INRIA Sophia-Antipolis.2004 Route de Lucioles, B.P. 93.6902 Sophia Antipolis, France. pp. 2164-2166.

[YSMB04] Shan Yu, Marc Berthod (2004), Urban detection in satellite images using map knowledge by a feedback control technique, IEEE, INRIA Sophia-Antipolis.2004 Route de Lucioles, B.P, Sophia Antipolis, France. pp. 98-102.

Referencia Electrónica

[AFG05] Dr. Ángel franco García Dpto. física aplicada universidad del país vasco 1998-2005
<http://www.sc.ehu.es/sbweb/fisica/cuantica/negro/espectro/espectro.htm>

[TCK05] Timothy Loh, César Salazar, Karen Wu-The optics
http://library.thinkquest.org/C003776/espanol/book/espectro_electromagnetico.htm

[AJB05] Dr. Antonio J. Bravo V. Universidad de los Andes Mérida Venezuela
<http://www.ing.ula.ve/> y <http://www.fing.ula.ve/>

[DPI] Sistema de procesamiento de infotmacoes georeferenciadas
http://www.dpi.inpe.br/spring/usuario_spa/entrada.htm#inform
http://www.dpi.inpe.br/spring/usuario_spa/sensorr.htm#espectro
http://www.dpi.inpe.br/spring/usuario_spa/sensorr2.htm#recepcao_ima



[FCAGR] <http://www.fcagr.unr.edu.ar/catedras/mdt/GTS/Zonaedu/perremot11.htm>

[PTD05] Patricio T. Díaz Pazos, astrocosmos
http://www.astrocosmo.cl/astrofis/astrofis-03_15.htm

[GBTEL] German Bense Teledet S.R.L. Juan Benito Blanco 822 Apto. 804 Montevideo CP
11300 Uruguay Teléfono - Fax (598 2) 712.1740 <http://www.teledet.com.uy/quees.htm>

[LSS05] L. Solé Sugrañes, Digital Image Systems S.L propiedades de las imágenes digitales
http://www.dimages.es/TutorialAI/introduccion/marcos_imdig.htm

[OSB] Dr. Oleg Starostenko Basarab, depto de ISC UDLA-P (Universidad de las Américas - Puebla). <http://mail.udlap.mx/~oldwall/docencia/cur2000.html>
<http://mail.udlap.mx/~oldwall/docencia/IMAGENES/curso.html>

Referencias Recomendadas

1. *Digital Image Processing a Systems Approach* [Green, William B. Ed. Van Nostrand Reinhold, New York, 1989]
2. *Digital Image Processing* [Pratt, William K. Ed. John Wiley & Sons, Inc. USA, 1991]
3. *Handbook of Imagen and Video Processing* [George M. and B.S Manjunath, Academic Press]
4. *Modelo matemático del proceso de desertificación empleando variables canónicas y técnicas de percepción remota* [Hevia Montiel Nidiyare, Fac.Ing.UNAM]
5. *Análisis del comportamiento de descriptores de la matriz de co-ocurrencia en la extracción de texturas en imágenes digitales* [Jiménez Jaramillo Álvaro, Fac.Ing y Arq. UNC.]
6. *Clasificación de imágenes utilizando descriptores estadísticos de textura* [Rivero de Nacimiento Paulo, Universidad Federal de Panamá] [Lagavulin.ltid.inpe.br: 1905/col/ltid.inpe.br/sbsr/2003/03.27.11.57/doc/15_185.pdf](http://Lagavulin.ltid.inpe.br:1905/col/ltid.inpe.br/sbsr/2003/03.27.11.57/doc/15_185.pdf)
7. *Classification of binary random patterns* [Abend, K., Harley, T. J., and Kanal, L. N. (1965).. *IEEE Transactions on Information Theory*, 11(4):538-544]
8. *Segmentación de Imágenes. Incorporación de las Características de Color a la Función de Energía* [Arques, P., Compañ, P., Molina, R., Pujol, M. and Rizo, R. (1999)]



9. *Multiscale Minimization of Global Energy Function for Image Segmentation [Arques, P., Compañ, P., Molina, R., Pujol, M. and Rizo, R. (2000) (RecPad2000), Porto, 11th Portuguese Conference on Pattern Recognition Portugal, May 2000, pp 153-158.]*
10. *Multiscale Image Segmentation: Adding Colour Features to the Global Energy Function [Arques, P., Compañ, P., Molina, R., Pujol, M. and Rizo, R. (2000). The 4th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2000), Orlando, July 2000]*
11. *Decision making in context [Haralick R. M. (1983). IEEE Transactions on Pattern Analysis and Machine Intelligence, 5(4):417-428.]*
12. *Markov Random Fields and Their Applications [Kindermann R. and Snell J. L.(1980) American Mathematical Society, Providence, R.I].*
13. *Simultaneous parameter estimation and segmentation of gibbs random fields using simulated annealing [Lakshmanan, S. and Derin, H. (1989) IEEE Transactions on Pattern Analysis and Machine Intelligence, 11:799-813]*
14. *A Markov random field model for object matching under contextual constraints [Li, S. Z. (1994).. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 866-869, Seattle, Washington]*
15. *Scene segmentation from visual motion using global optimization [Murray, D. and Buxton, B. (1987) IEEE Transactions on Pattern Analysis and Machine Intelligence, 8:220-228]*

Referencias electrónicas recomendadas

1. <http://www.sc.ehu.es/sbweb/fisica/cuantica/negro/espectro/espectro.htm#EI%20espectro%20electromagnético>
2. <http://www.ciat.cgiar.org/dtmradar/teledeteccion.htm#1>
3. http://www.altamira-information.com/450-techniques_ES.html
4. <http://www.itch.edu.mx/academic/industrial/sabaticorita/>
5. <http://www.es-geo.com/merge.htm>