

**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE INGENIERÍA

**ANÁLISIS CONTEXTUAL PARA LA
SEGMENTACIÓN DE IMÁGENES DE
PERCEPCIÓN REMOTA DEL DISTRITO
FEDERAL**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

INGENIERA EN COMPUTACIÓN

P R E S E N T A

HORTENCIA CABALLERO LÓPEZ

DIRECTOR DE TESIS: DR. MIGUEL MOCTEZUMA FLORES

CIUDAD UNIVERSITARIA

SEPTIEMBRE 2002.



**TESIS CON
FALLA DE ORIGEN**



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.

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recibido.

NOMBRE: Hortencia Caballero López

FECHA: 9 - Septiembre - 2002

FIRMA: Hortencia Caballero L.

“...Cuando yo decía:

 Mi pie resbala,

 tu misericordia, oh Jehová,
 me sustentaba.

 En la multitud de mis pensamientos
 dentro de mí, tus consolaciones
 alegraban mi alma...”

(Salmos 94:18,19)

DEDICADO A:

Dios y Jesucristo

Por haberme traído a este mundo,
Porque me han dado una familia,
Por darme vida a cada segundo que pasa y,
Por haberme permitido y ayudado a finalizar esta etapa de mi vida.

Mis padres

Angela : Gracias mamá, porque me has dado todo y más de lo que yo esperaba
Leobardo: Gracias papá, porque me has ayudado, aún cuando no te lo pedía.

Mis hermanos

Emigdia: Por la gran ayuda que me otorgaste todo el tiempo que duró esta aventura
Efraín y Magdaleno: Por el apoyo económico que sin condiciones me proporcionaron para
la realización de mis estudios.
María de la Cruz: Porque siempre te he considerado y te quiero como mi hermana y me has
proporcionado una invaluable ayuda.

Mi grupo de educación secundaria

Especialmente a Rosa María Ulloa, Celia Guerrero, Juan Gustavo Almanza y Juan Carlos
Roa (q.e.p.d.).

AGRADECIMIENTOS:

Al Dr. Miguel Moctezuma Flores, por haberme dado la oportunidad de dirigir mi tema de tesis, por mostrarme su apoyo y confianza y por haberme tenido tanta paciencia. Gracias Doctor.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por la beca otorgada para la realización de esta tesis.

A la Facultad de Ingeniería, por haberme albergado todo el tiempo que duró mi formación profesional.

A la División de Estudios de Posgrado de la Facultad de Ingeniería (DEPFI) por haber abierto sus puertas para el desarrollo de esta tesis.

Al Ing. Benjamín Ramírez Hernández, porque siempre me animó a terminar mi carrera profesional.

A la Ing. Patricia Salas, por la confianza que siempre me tuvo en la realización de mi Servicio Social.

A mis amigos de la Facultad de Ingeniería:

Agustín Padilla, Nancy Pacheco, Salvador García, Mariela Vargas, José Camaño, Arturo Meneses, Xóchitl, Esther, Manuel, Juan García, Araceli y Janeth. Por la gran amistad que me otorgaron, y porque sé que durará para siempre...

A Miguel Angel, por todos los momentos bonitos que pasamos juntos.

ÍNDICE

INTRODUCCIÓN.....1

CAPÍTULO 1:

FUNDAMENTOS DE PERCEPCIÓN REMOTA6

1.1. – *Definición*7
1.2. – *Elementos básicos*8
 1.2.1. – Fuente de iluminación8
 1.2.2. – Escena8
 1.2.3. – Sistema de adquisición8
 1.2.4. – Sensores9
1.3. – *Análisis de datos*11
1.4. – *Sistemas de Percepción Remota*12
 1.4.1. – Tipos de satélites12
 1.4.2. – Bandas de frecuencia utilizadas por los satélites13
 1.4.3. – Características de algunos satélites14
 1.4.3.1. – SEASAT14
 1.4.3.2. – LANDSAT15
 1.4.3.3. – ERS22
 1.4.3.4. – JERS-123
 1.4.3.5. – ALOS24
 1.4.3.6. – RADARSAT26
 1.4.3.7. – SPOT28
 1.4.3.8. – IKONOS30

CAPÍTULO 2:

ANÁLISIS POR TEXTURAS31

2.1. – *Introducción*32
2.2. – *Definición de textura*32
2.3. – *Clasificación de texturas*33
 2.3.1. – Texturas naturales33
 2.3.2. – Texturas artificiales33
 2.3.3. – Texturas regulares33
 2.3.4. – Texturas irregulares33
2.4. – *Técnicas de análisis de texturas*34
 2.4.1. – Técnica estructural34
 2.4.1.1. – Método34
 2.4.1.2. – Análisis de texturas débiles35
 2.4.1.3. – Análisis de texturas fuertes36

2.4.2. – Técnica estadística	37
2.4.2.1. – Matriz de Co-ocurrencia	37
2.4.2.2. – Campos Aleatorios de Markov	41
2.4.3. – Técnica espectral	44

CAPÍTULO 3:

TÉCNICAS DE SEGMENTACIÓN

3.1. – Definición de segmentación	47
3.2. – Objetivo de la segmentación	47
3.3. – Propiedades de las técnicas de segmentación	48
3.3.1. – Discontinuidad	48
3.3.2. – Similaridad	48
3.4. – Técnicas de segmentación	49
3.4.1. – Segmentación orientada a regiones	49
3.4.1.1. – Segmentación por clasificación	49
3.4.1.2. – Segmentación por crecimiento (unión) de regiones	49
3.4.1.2.1. – Crecimiento de regiones por agregación de píxeles	50
3.4.1.2.2. – Crecimiento de regiones por enlace sencillo	51
3.4.1.2.3. – Crecimiento de regiones por enlace de centroide	52
3.4.1.3. – Segmentación por división y fusión de regiones	52
3.4.1.3.1. – División de regiones	53
3.4.1.3.2. – Fusión de regiones	54
3.4.1.4. – Segmentación basada en el análisis del histograma (umbralización)	56
3.4.1.4.1. – Concepto de umbralización	56
3.4.1.4.2. – Tipos de umbralización	56
3.4.1.4.3. – Método P-cuantil	56
3.4.1.4.4. – Método Isodata	57
3.4.1.4.5. – Método de Simetría de Fondo	58
3.4.1.4.6. – Método del Triángulo	59
3.4.1.4.7. – Limitaciones de los métodos basados en el histograma	60
3.4.2. – Métodos de detección de bordes	60
3.4.2.1. – Definición de borde	60
3.4.2.2. – Tipos básicos de discontinuidades	61
3.4.2.2.1. – Detección de puntos	61
3.4.2.2.2. – Detección de líneas	62
3.4.2.2.3. – Detección de bordes	63
3.4.2.3. – Métodos de segmentación por detección de bordes	64
3.4.2.3.1. – Operador gradiente	64
3.4.2.3.2. – Operador de Roberts	65
3.4.2.3.3. – Operador de Sobel	66

3.4.2.3.4. – Operador de Prewitt	67
3.4.2.3.5. – Operador de Kirsch	67
3.4.2.3.6. – Operador de Robinson	68
3.4.2.3.7. – Operador Laplaciano	69
3.4.2.3.8. – Operador Laplaciano del Gaussiano	71
3.4.2.3.9. – Método de segmentación Watershed	71
3.4.2.3.10. – Método de segmentación basado en la Transformada de Hough	73

CAPÍTULO 4:

TEORÍA DE LOS CAMPOS ALEATORIOS DE MARKOV

4.1. – <i>Probabilidad y Variables Aleatorias</i>	76
4.1.1. – Probabilidad	76
4.1.1.1. – Definición de Probabilidad	76
4.1.1.2. – Medidas numéricas	76
4.1.2. – Variables Aleatorias	78
4.1.2.1. – Concepto de variable aleatoria	78
4.1.2.2. – Variables aleatorias discretas	79
4.1.2.2.1. – Función de probabilidad	79
4.1.2.2.2. – Distribución acumulada	81
4.1.2.2.3. – Valor esperado	82
4.1.2.2.4. – Varianza	82
4.1.2.2.5. – Desviación estándar	82
4.1.2.3. – Variables aleatorias continuas	82
4.1.2.3.1. – Distribución de probabilidad	83
4.1.2.3.2. – Distribución acumulada	84
4.1.2.3.3. – Valor esperado	85
4.1.2.3.4. – Varianza	85
4.1.2.3.5. – Desviación estándar	85
4.1.2.4. – Momentos de una variable aleatoria	86
4.2. – <i>Teoría Bayesiana de Probabilidad</i>	88
4.2.1. – Probabilidad condicional	88
4.2.2. – Eventos independientes y eventos mutuamente excluyentes	89
4.2.3. – Teorema de Bayes	89
4.2.4. – Probabilidad total	90
4.2.5. – Distribución condicional	90
4.2.6. – Valor esperado condicional	90
4.3. – <i>Antecedentes de los MRF</i>	91
4.4. – <i>Estimación bayesiana</i>	92
4.4.1. – Métodos clásicos de estimación	92
4.4.2. – Métodos bayesianos de estimación	93
4.4.2.1. – Probabilidad a priori	93
4.4.2.2. – Probabilidad a posteriori	94

4.5. – Ejemplos de modelos markovianos	97
4.5.1. – Problema de etiquetado	97
4.5.2. – Sistemas de Vecinos y Cliques	98
4.5.2.1. – Sistema de vecinos	98
4.5.2.2. – Sistema de cliques	100
4.5.3. – Distribución de Gibbs	101
4.5.4. – Modelo Ising	101
4.5.4.1. – Aplicación al ferromagnetismo	102
4.5.5. – Modelo Potts	103
4.5.5.1. – Aplicación al ferromagnetismo	104

CAPÍTULO 5:

SEGMENTACIÓN Y FUSIÓN DE DATOS	106
5.1. – Modelo de Fusión	107
5.1.1. – Las medidas de error local y global	109
5.1.2. – Proceso de fusión	109
5.2. – Modelo de Energía	111
5.2.1. – Modelo de energía local	112
5.2.2. – Modelo de energía global	113

CAPÍTULO 6:

RESULTADOS	114
6.1. – Imágenes de prueba	116
6.2. – Resultados con el operador de Entropía	117
6.3. – Resultados de operadores de texturas	118
6.4. – Resultados binarios para la imagen de prueba #1	121
6.5. – Resultado final de la fusión de datos para la imagen de prueba #1	123
6.6. – Resultados binarios para la imagen de prueba #2	124
6.7. – Resultado final de la fusión de datos para la imagen de prueba #2	126

CAPÍTULO 7:

CONCLUSIONES	127
---------------------------	------------

BIBLIOGRAFÍA	130
---------------------------	------------

ANEXO: DISEÑO Y PROGRAMACIÓN	136
---	------------

INTRODUCCIÓN

Desde hace mucho tiempo atrás, uno de los sueños más anhelados por la humanidad ha sido poder observar desde el espacio a nuestro planeta. Pero sólo en fechas recientes se ha dispuesto de los medios técnicos necesarios para hacer realidad este deseo.

Desde ese momento, nació una rama de la ciencia llamada Percepción Remota , lo que nos ha permitido enriquecer notablemente nuestro conocimiento sobre el planeta y sus habitantes.

Uno de los principales objetivos de esta aventura aérea ha sido la búsqueda de una nueva visión de los paisajes terrestres. El afán de remontar la limitada perspectiva de la visión humana es evidente desde los inicios de la aeronáutica llegando a ser, hoy en día, una pieza clave de la investigación sobre la visión computacional.

La Percepción Remota, cuenta con numerosas aplicaciones tales como:

- Cobertura global y periódica de la superficie terrestre.
- Visión panorámica (Deforestación)
- Procesamiento digital de imágenes.
- Agricultura (Predicción de heladas, erosión del suelo, condiciones de clima)
- Meteorología.
- Etc.

Además, el reciente desarrollo de satélites para la observación ha arrojado grandes colecciones de datos en forma de imágenes que tienen que ser procesadas por computadora, dando así un gran campo para el estudio e investigación de la percepción remota

Dentro del procesamiento digital de imágenes es muy interesante el estudio de los diferentes métodos de segmentación de imágenes, los cuáles se han dividido de manera tradicional en dos grupos:

- Segmentación supervisada.
- Segmentación no supervisada.

El método supervisado parte de un conocimiento previo, a partir del cual se seleccionan las muestras para cada una de las categorías.

El método no supervisado realiza una búsqueda automática de las regiones o categorías presentes en la imagen.

Una vez que ya se han definido las clases que se van a considerar para la segmentación, se realiza en sí la segmentación, para esto, se han dividido también en dos grupos:

- Segmentación contextual.
- Segmentación no contextual.

La segmentación contextual requiere un tiempo mayor de procesamiento que los métodos no contextuales, y toma en cuenta la información de el píxel de análisis con sus vecinos. Las imágenes generadas con este método muestran regiones homogéneas y baja incidencia de píxeles aislados, pero a su vez, estructuras finas pueden eliminarse.

La segmentación no contextual no toma en cuenta la información de la vecindad del píxel a clasificar, lo cual puede generar problemas de sobresegmentación.

Dentro de los métodos contextuales se encuentran los Campos Aleatorios de Markov que encuentran enormes aplicaciones en el procesamiento de imágenes.

Para poder realizar la aproximación estadística en análisis de texturas, se hace uso de los descriptores estadísticos siguientes:

- Energía.
- Prominencia.
- Varianza.
- Matíz.
- Tendencia.
- Contraste.
- Correlación.
- Homogeneidad.
- Entropía.
- Autocorrelación.
- Disimilitud.

En este trabajo se presentan varios métodos utilizados para el análisis de fotografía aérea de alta resolución. Usando una combinación de datos texturales, se logra realizar la segmentación de escenas monoculares de áreas urbanas del D.F. Para este efecto, se utilizó una base de datos formada por fotos aéreas digitalizadas de alta resolución del Valle de la Ciudad de México.

A continuación se presenta un resumen del contenido total de la tesis, empezando con la definición del problema, siguiendo con la propuesta para la solución del problema y terminando con el resumen de cada uno de los capítulos.

Definición del problema:

Por medio de un análisis combinado entre Campos Aleatorios de Markov y métodos texturales, realizar la segmentación de escenas monoculares de áreas urbanas del Distrito Federal, haciendo uso de la fotografía aérea de alta resolución.

Propuesta de solución:

A través del empleo de la técnica de análisis de texturas llamada Matriz de Co-ocurrencia, y con la ayuda de los operadores de Haralick, definir un esquema de fusión de datos a efectos de integrar información parcial (texturas y niveles de gris) al proceso de segmentación. La segmentación debe proporcionar datos sobre regiones urbanas y no urbanas (dos clases).

Resumen:

Capítulo 1

Contiene una recopilación de los principales fundamentos y definiciones de lo que es la Percepción Remota, así como datos técnicos de algunos satélites, tratando de enfocar dichos datos a su función recolectora de imágenes.

Capítulo 2

Muestra una perspectiva de los enfoques texturales de segmentación de imágenes, dando un panorama general de definiciones y métodos utilizados para este fin. A fin de presentar el esquema que se utilizó para la segmentación de nuestras imágenes de prueba, se pone especial énfasis en el método de la Matriz de Co-ocurrencia.

Capítulo 3

Presenta un resumen de las técnicas de segmentación en forma general, tratando de explicar de la manera más sencilla cada técnica.

Capítulo 4

Da un panorama general de los Campos Aleatorios de Markov, empezando por un resumen de Probabilidad y Estadística, seguido por la Teoría Bayesiana de Probabilidad y terminando con dos ejemplos de modelos markovianos.

Capítulo 5

Muestra un enfoque de la segmentación y la fusión de datos, tratando de mostrar el proceso utilizado en la solución al problema abordado por esta tesis.

Capítulo 6

En este capítulo se presentan los resultados obtenidos en la solución del problema planteado, se muestran todas las imágenes generadas al aplicar los descriptores de Haralick para la obtención de las texturas, también se muestran los resultados de la fusión de datos, así como el mapa de la imagen original y las tres imágenes utilizadas.

Capítulo 7

Es en este capítulo donde se muestran las conclusiones finales de esta tesis.

Anexo

En esta parte se realiza una esquematización del desarrollo de los programas, así como la justificación del uso de la herramienta de programación.

CAPÍTULO 1

FUNDAMENTOS DE PERCEPCIÓN REMOTA

1.1.- DEFINICIÓ

Percepció Remota es la ciència i arte de obtenir informació relacionada amb un objecte, àrea o fenomen a través del anàlisi de dades adquirides per mitjà d'un dispositiu que no està en contacte físic amb l'objecte, àrea o fenomen de investigació (Lillesand y Kiefer, 1987).

De una manera molt particular, se pot definir com la obtenció i el anàlisi de imàgenes per mitjà de sistemes satel·litaris.

La radiació electromagnètica es rebida per sensors que se troben a bord de los satèl·lits, com se observa en la figura 1.1:

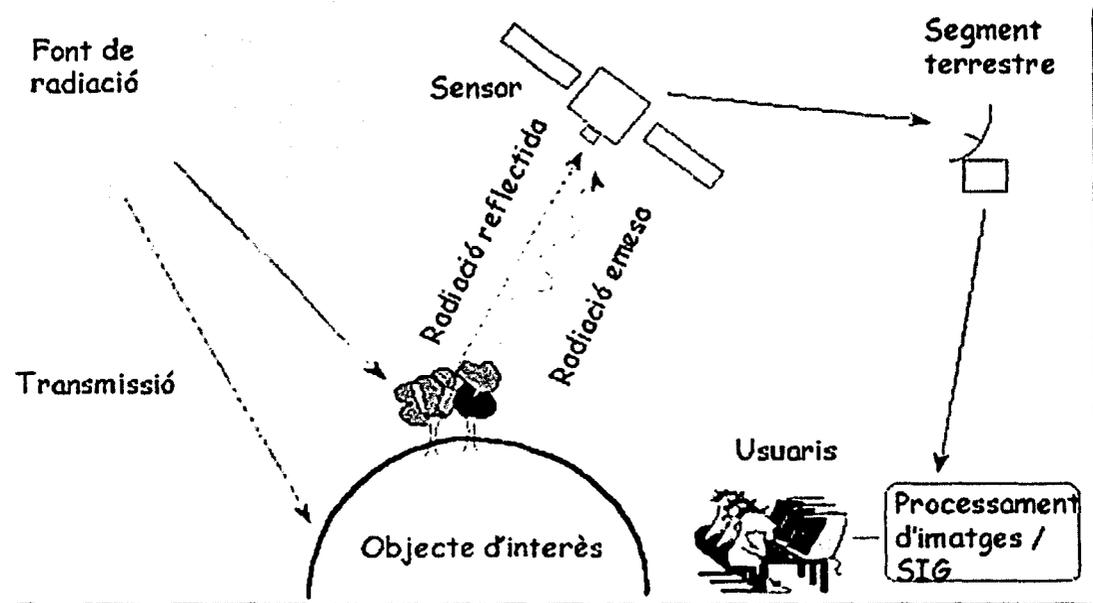


Figura 1.1. – Esquema de la radiació electromagnètica

1.2- ELEMENTOS BÁSICOS:

1.2.1.- Fuente de iluminación

La fuente de iluminación puede ser de dos tipos:

- A. **Externa.**- Es aquella en donde la condición de iluminación puede ser el sol o alguna estrella lejana. Los sistemas que utilizan esta fuente de iluminación se llaman sistemas pasivos.
- B. **Interna.**- Es aquella en donde la condición de iluminación viene integrada en el propio sistema. Los sistemas que utilizan esta fuente de iluminación se llaman sistemas activos, por ejemplo, los radares.

1.2.2. - Escena

La escena está formada por el sistema físico que nos interesa estudiar.

1.2.3. - Sistema de adquisición

El sistema de adquisición es el dispositivo que realiza la adquisición de la imagen o el sistema físico de nuestro interés. Por lo general, es una cámara. Existen dos tipos de cámaras:

- A. **Tipo tubo de rayos catódicos.**- Este tipo de cámara funciona como una televisión, el haz de electrones barre la pantalla de manera vertical de izquierda a derecha. Es de tipo analógico y utiliza un tubo de rayos catódicos (Figura 1.2). A esta tecnología se le llama Tecnología VIDICON:
- B. **Tipo CCD.**- Este tipo de cámara utiliza sensores sensibles a la luz, es de tipo digital. A esta tecnología se le llama Tecnología CCD (Charged Coupled Device). Un dispositivo CCD está compuesto por una matriz de fotodetectores dispuestos en columnas y filas. Cada detector individual se denomina píxel. Cuando la luz incide sobre un dispositivo CCD cada píxel queda impresionado por el fotón que le llega. Si el dispositivo es eficiente al máximo, lo que nunca sucede, cada fotón genera un electrón haciendo que la carga aumente en función del tiempo de integración (equivalente al tiempo de exposición en fotografía convencional) y de la intensidad de la luz incidente. Cuando la

integración ha sido completada, la carga es transferida a un ordenador que lee cada fila y cada columna y seguidamente las cuantifica y digitaliza.

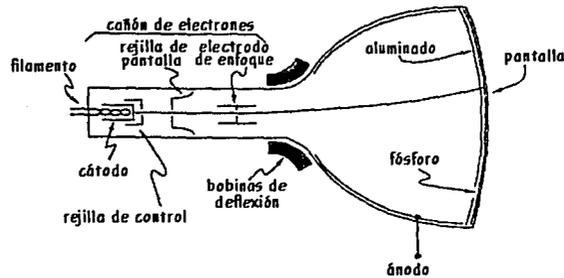


Figura 1.2.- Cámara de rayos catódicos

1.2.4. - Sensores

Son los equipos capaces de coleccionar la energía proveniente del objeto de interés, convirtiéndola en una señal posible de ser registrada y presentándola en forma adecuada para la extracción de información.

- A. **Sensores remotos**: 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.
- B. **Tipos de sensores**: Los sensores pueden ser clasificados en función de la fuente de energía, en función del tipo de producto, o en función del tipo de barrido.

En función de la fuente de energía:

- a) **Pasivos**.- No poseen fuente propia de radiación. Miden la radiación solar reflejada o la radiación emitida por el objeto.
- b) **Activos**.- Poseen su propia fuente de radiación electromagnética.

En función del tipo de producto:

- a) **No imageadores.**- No generan una imagen de la superficie observada, su salida está en forma de gráficos.
- b) **Imageadores.**- Se obtiene como resultado una imagen de la superficie observada.

En función del tipo de barrido:

- a) **Unidimensionales.**- El barrido que realizan los electrones se da en una dimensión, tal como la tecnología VIDICON.
- b) **Bidimensionales.**- El barrido que realizan los electrones se da en dos dimensiones, ya que cuenta con sensores y registros, y al arreglo bidimensional generado se le llama píxel. Una tecnología bidimensional es la tecnología CCD.
- c) **Tridimensionales.**- El barrido que realizan los electrones se da en tres dimensiones. Es decir, a cada píxel se le asocia otro dato de profundidad, que en este caso es un láser o un sonar. Para lo cual el láser se lanza y se recibe, el tiempo que tarda en pegar y rebotar es el valor del píxel de profundidad. A estos sensores se les llama también sensores de imágenes de rango. Dentro de los sensores tridimensionales tenemos de dos tipos:
 - 1. **Sensores tridimensionales activos.**- Aquellos que realizan emisión-recepción.
 - 2. **Sensores tridimensionales pasivos.**- Aquellos que solamente realizan recepción.

1.3.- ANÁLISIS DE DATOS

Un análisis de datos típico es aquel en el que se usan tres tipos de composición de color, los cuales son: Tonos de gris, Falso color y Color verdadero (Figura 1.3, 1.4 y 1.5).

- Tonos de gris:



Para esta composición se utiliza una sola banda. Se asigna un tono negro a un valor 0, y uno claro al 255; a los intermedios, valores de grises.

Figura 1.3

- Falso Color:



Se despliegan tres bandas asignando a cada una de ellas tonos de los colores rojo, verde y azul.

Figura 1.4

- Color Verdadero:



Color artificial, similar al que ven los ojos humanos.

Figura 1.5

1.4.- SISTEMAS DE PERCEPCIÓN REMOTA

En general los sistemas de percepción remota están basados en satélites y las imágenes que estos envían a la Tierra para su análisis. Estos sistemas representan la mayor fuente de datos de Percepción Remota y con mayor potencial de continuidad a lo largo del tiempo.

1.4.1.- Tipos de Satélites.

- **Por su órbita:**

- A. Satélites de órbita geoestacionaria.- Situados a 36 mil kilómetros sobre el nivel del mar. Se sincroniza al movimiento de la Tierra, y observa siempre la misma zona. (GEO)
- B. Satélites de órbita media.- Situados entre 10 mil y 20 mil kilómetros sobre el nivel del mar. (MEO)
- C. Satélites de órbita baja. - Situados entre 300 y 800 kilómetros sobre el nivel del mar. (LEO).(figura 1.6).

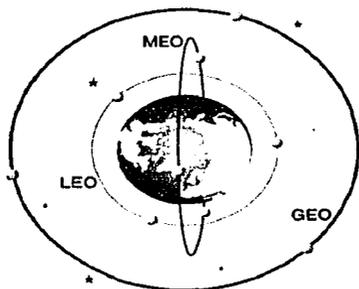


Figura 1.6. - Tipos de órbitas satelitales

- **Por su finalidad:**

- A. Satélites de Telecomunicaciones (Radio y Televisión).
- B. Satélites Meteorológicos.
- C. Satélites de Navegación.
- D. Satélites Militares y espías.
- E. Satélites de Observación de la tierra.
- F. Satélites Científicos y de propósitos experimentales.
- G. Satélites de Radioaficionado.

● **Por su resolución:**

- Espacial.- Mínima separación a la cual los objetos aparecen distintos y separados en la imagen, también llamada tamaño del píxel (elemento geográfico mínimo en una imagen). Esta resolución va desde 1.1X1.1Km, hasta 5 X 5Km.
- Geográfica.- Porción geográfica que cubierta por cada exposición del sensor llega a ser hasta de 60 X 60Km.
- Espectral.- Se refiere al número y ancho de bandas espectrales que puede registrar un sensor.
- Radiométrica.- Depende de la capacidad del sensor para detectar variaciones en la radiación espectral, la mayoría de los sensores en operación poseen 256 niveles de codificación por píxel
- Temporal.- Se refiere a la frecuencia de cobertura que proporciona el sensor. El ciclo de cobertura está en función de las características orbitales

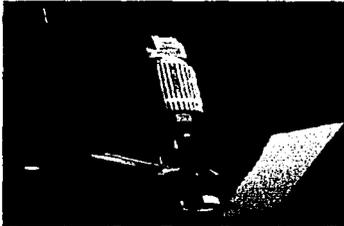
1.4.2.- Bandas de frecuencias utilizadas por los Satélites.

Banda P	200-400 Mhz.
Banda L	1530-2700 Mhz.
Banda S	2700-3500 Mhz.
Banda C	3700-4200 Mhz. 4400-4700 Mhz. 5725-6425 Mhz.
Banda X	7900-8400 Mhz.
Banda Ku1 (Banda PSS)	10.7-11.75 Ghz.
Banda Ku2 (Banda DBS)	11.75-12.5 Ghz.
Banda Ku3 (Banda Telecom)	12.5-12.75 Ghz.
Banda Ka	17.7-21.2 Ghz.
Banda K	27.5-31.0 Ghz.
1 Mhz.= 1000.000 Hz.	
1 Ghz.= 1000.000.000 Hz.	

Tabla 1.1. - Bandas de frecuencia

1.4.3. - Características de algunos satélites:

1.4.3.1. -SEASAT (USA)



El SEASAT es el primer satélite oceanográfico norteamericano, diseñado para estudiar y explorar todos los mares de la tierra. Recogía información sobre los vientos, la temperatura superficial, corrientes, altura de las olas, condiciones de hielo, topografía de los océanos y actividad de las tormentas en las costas.

Los sensores que llevaba a bordo eran:

- ALT (Radar Altimeter) este sensor mide el tamaño de las olas que se producen en la superficie del océano y, mediante una combinación de la altitud medida y la determinación de la órbita, el satélite generaba una imagen topográfica de la superficie del mar.
- SASS(Seasat-A Satellite Scatterometer) medía la superficie del mar, así como la velocidad y dirección del viento.
- SMMR(Scanning Multichannel Microwave Radiometer) medía la velocidad del viento así como la temperatura del mar.
- VIRR(Visible and IR Radiometer) tenía como objetivo proveer una imagen de baja resolución, a partir de una radiación infrarroja en el océano, costa y atmósfera, y así poder medir temperaturas.
- SAR(Synthetic Aperture Radar). Es un radar que genera imágenes de la superficie del océano, para determinar patrones de olas. Es un sensor de microondas.

La transmisión de datos la realiza a 800 kbps.

Sus características son:

Fecha de lanzamiento	Junio de 1978.
Altura sobre la Tierra	790 km
Banda Espectral	1.275 GHz (Banda-L)
Resolución	25 m
Tamaño de la imagen	100 km
Tamaño de registro de datos	110 Mbits/s (5 bits/palabra)

Tabla 1.2. – Características del satélite SEASAT

1.4.3.2. - LANDSAT (Land Satellite)



LANDSAT-1 es el primer observador satelital en el mundo. Le siguieron, LANDSAT-2, 3, 4, 5 y 7.

a) LANDSAT-1:

Tiene dos tipos de sensores, el RBV(Return Beam Vidicon) y el MSS (Escáner Multiespectral). El RBV tiene registro instantáneo de toda la escena, mientras que el MSS tiene registro del terreno por barrido de líneas, es un sensor óptico para observar la radiación solar reflejada en la superficie en 4 diferentes bandas.

En la tabla 1.3 se muestran las características del satélite LANDSAT-1:

<i>RBV Return Beam Vidicon</i>	
Resolución	80 m
Tamaño de la imagen	185 km
Bandas Espectrales	0.475-0.575 0.58-0.68 0.69-0.83
<i>MSS Multispectral Scanner</i>	
Resolución	80 m
Tamaño de la imagen	185 km
Bandas Espectrales	0.5-0.6 0.6-0.7 0.7-0.8 0.8-1.1

Tabla 1.4. – Características del satélite LANDSAT-1

b) LANDSAT-2:

Al igual que el LANDSAT-1, llevaba consigo un sensor tipo MSS y otro de tipo RBV.

c) LANDSAT-3:

Al igual que sus dos antecesores, lleva consigo los mismos dos registros: MSS y RBV, pero ahora las características de los sensores son diferentes (Tabla 1.5).

RBV	
Resolución	40 m
Tamaño de la imagen	185 km
Bandas Espectrales	0.5-0.75
MSS	
Resolución	80 m
Tamaño de la imagen	185 km
Bandas Espectrales	Igual a las del Landsat-1 en adición con la banda térmica
Banda térmica	10.4 - 12.6 micrones
	Resolución 120 m

Tabla 1.5. – Características del satélite LANDSAT-3

d) LANDSAT-4:

Llevaba consigo el sensor MSS, y además uno nuevo, el TM (Thematic Mapper) el cual opera bajo siete bandas espectrales. Este sensor conceptualmente, es semejante al MSS porque realiza un barrido por línea, pero además está mejorado en cuanto a sus componentes ópticos y electrónicos (Tabla 1.6).

MSS	
Resolución	75 m
TM <i>Thematic Mapper</i>	
Resolución	30 m - 120 m
Tamaño de la imagen	185 km
Bandas Espectrales (micrones)	.45 -.52 .52 -.60 .63 -.69 .76 -.90 1.55 -1.75 10.4 -12.5 2.08 -2.35

Tabla 1.6. – Características del satélite LANDSAT-4

e) LANDSAT-5:

Sus características son similares a las del LANDSAT-4.

f) LANDSAT-6:

Este satélite utiliza los sensores ETM, en Banda Pancromática, Banda Multiespectral Pancromática, y la banda térmica. Esto se muestra en la tabla 1.7.

ETM Enhanced Thematic Mapper	
Banda Pancromática	
Resolución	13 x 15 m
Banda Espectral (micrones)	0.50-0.90 (pancromática)
Banda Multiespectral Pancromática	
Resolución	30 m
Banda Espectral (micrones)	0.45-0.52 0.52-0.60 0.63-0.69 0.76-0.90 1.55-1.75 2.08-2.35
Banda Thermal	
Resolución	120 m
Banda Espectral (micrones)	10.4-12.5

Tabla 1.7. – Características del satélite LANDSAT-6

g) LANDSAT-7:

Realiza la observación usando siete bandas. Lleva consigo un instrumento llamado ETM+, este instrumento es un sensor radiométrico multiespectral de ocho bandas, capaz de producir imágenes de alta resolución (3.8 gigabits de datos) tomadas de la superficie de la Tierra.

El sensor ETM+ colecta, filtra y detecta radiación en una cobertura de 183Km en el rango visible, cercano al infrarrojo, infrarrojo y microondas de las bandas de frecuencia; opera sobre 3 ejes, estabilizando la plataforma del satélite.

Además, es un derivado de el sensor TM, el cual era parte del LANDSAT 4 y 5, pero con modificaciones que incluyen la adición de la banda pancromática y dos rangos de ganancia, mejorando la resolución espacial, también se le adicionaron dos calibradores solares. Esto se muestra en la figura 1.7:

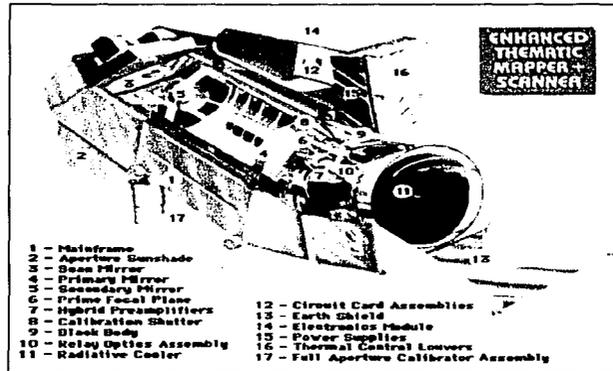


Figura 1.7. - Sensor ETM+

El LANDSAT 7 utiliza las bandas S y X para la comunicación. La banda S es usada para telemetría, mientras que la banda X es usada para la transmisión de datos. Puede recolectar datos a 378 gigabits. El tamaño de la imagen es de 185 X 185 Km. Su resolución es de 30 X 30m.

Su fuente de poder proviene de un array solar y dos baterías de níquel- hidrógeno que proporcionan 50 amperes por hora.

Consta también de 2 canales de giro, un procesador y un torque para controlar los momentos de descarga. El control de la órbita lo realiza con un sistema de ráfaga y un tanque de hidrozina de 270 libras:

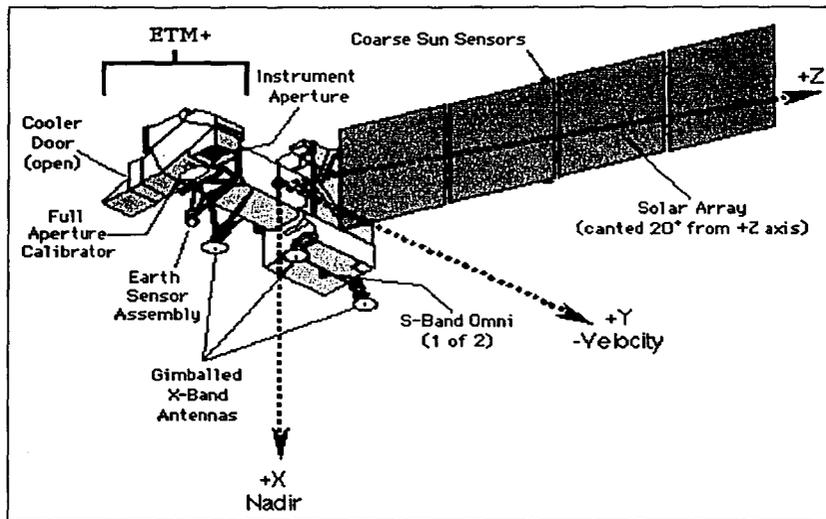


Figura 1.8. - Diagrama de los componentes del satélite Landsat-7.

TESIS CON
 FALLA DE ORIGEN

La Resolución especial del LANDSAT-7 se muestra en la tabla 1.8.

LANDSAT-7 ETM+ CARACTERÍSTICAS DE LAS BANDAS DE FRECUENCIA					
Banda Numero	Rango Espectral (μ)	Resolución (m)	Datos Por Línea (Scan)	Datos Por Línea Longitud (bytes)	Bits Por Muestra
1	.450 a .515	30	16	6,600	8
2	.525 a .605	30	16	6,600	8
3	.630 a .690	30	16	6,600	8
4	.775 a .900	30	16	6,600	8
5	1.550 a 1.750	30	16	6,600	8
6*	10.40 a 12.50	60	8	3,300	8
7	2.090 a 2.35	30	16	6,600	8
8	.520 a .900	15	32	13,200	8

*Valores aplicados a alta y baja ganancia en los datos.

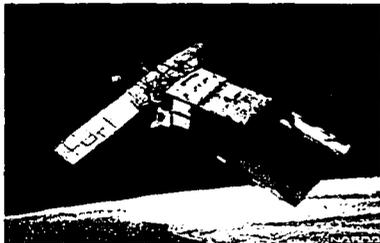
Tabla 1.8. - Características de la resolución espacial y las bandas de frecuencia del satélite LANDSAT-7.

Las características generales de los LANDSAT se muestran en la tabla 1.9

	Fecha de lanzamiento	Fecha de terminación	Altura promedio sobre la Tierra	Sensor	Periodo de cobertura (días)
LANDSAT-1	Julio 1972	Enero 1978	915km	MSS y RBV	18
LANDSAT-2	Enero 1975	Febrero 1982	915km	MSS y RBV	18
LANDSAT-3	Marzo 1978	Marzo 1983	915km	MSS y RBV	18
LANDSAT-4	Julio 1982	Junio 2001	705km	MSS y TM	16
LANDSAT-5	Marzo 1984	Junio 2001	705km	MSS y TM	16
LANDSAT-6	Octubre 1993	Octubre 1993	-	ETM	16
LANDSAT-7	Abril 1999		705km	ETM+	16

Tabla 1.9. – Características generales de los satélites LANDSAT

1.4.3.3. - ERS (European Remote Sensing Satellite)



El ERS-1 fue el primer satélite europeo en tomar imágenes de radar, transportó un gran número de instrumentos. Mide la velocidad y dirección del viento así como parámetros relativos a las olas de los océanos. Posteriormente se lanzó el ERS-2 con la diferencia de que llevaba un nuevo diseño para la medición de la estratosfera y la troposfera.

Sus bandas utilizadas son dos:

- Banda S. Trabaja a 2Kbits/seg, es usada para Sistemas de Telemetría, Telecomandos y Control(TTC).
- Banda X.- Es utilizada para la transmisión de datos a 15Mbits/Segundo.

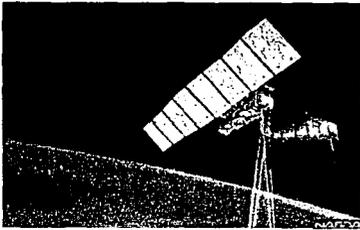
Sus sensores utilizados son:

- AMI (Active Microwave Instrumentation). Este sensor combina las funciones de el SAR y de el sensor WNS (Wind Scatterometer). Realiza medidas del viento y de las olas del océano, y registra todo el tiempo. Genera resoluciones finas en las imágenes del océano, del hielo polar y de las costas.
- RA (Radar Altimeter). Realiza medidas de altitud, viento, hielo y océano.
- ATSR (Along Track Scanning Radiometer). Consta de cuatro canales de infrarrojo para medir temperaturas.
- MWS (Microwave Sounder). Tiene dos canales radiométricos de infrarrojo con los cuales percibe información de la precipitación del vapor de agua en la atmósfera.
- PRARE (Precise Range and Range-rate Equipment). Es un instrumento experimental con alta precisión en datos de órbita.
- LR (Láser Retroreflector). Permite el uso de rangos precisos de calibración.

Sus características generales son:

- | |
|--|
| <ul style="list-style-type: none">➤ Fecha de lanzamiento: ERS-1: Julio 1991. ERS-2: Abril 1995.➤ Altura promedio sobre la Tierra: 800Km.➤ Tamaño de la imagen: 100 X 100 km².➤ Resolución: 30 X 30m². |
|--|

1.4.3.4.- JERS-1 (Japanese Earth Resources Satellite)



El JERS es un satélite de observación para el estudio de la Tierra, se realizan estudios de agricultura, forestación, medios de protección, prevención de desastres, y monitoreo de costas.

Sus sensores son:

- SAR (Synthetic Aperture Radar). Tiene una observación de alto contraste, y determina rasgos topográficos
- OPS (Optical Sensor). Sensor óptico para separar la luz reflejada sobre la superficie de la Tierra en un rango visible, también mide microondas, así como luz infrarroja. Utiliza siete bandas para separar la luz y tiene capacidad de observación estereoscópica. Este sensor a su vez cuenta con dos sub-sensores: VNIR (Visible and Near Infrared Radiometer) y el SWIR (ShortWave Infrared Radiometer).

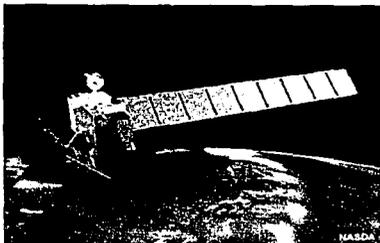
Su sistema de recolección de datos cuenta con el sistema llamado MDR (Mission Data Recorder), el cual trabaja a 30Mbits/segundo en dos canales.

El JERS-1 opera en dos frecuencias, la primera de 8025- 8175MHz y la segunda, de 8215-8400MHz.

Sus características generales son:

- | |
|--|
| <ul style="list-style-type: none">➤ Fecha de lanzamiento: Febrero 1992.➤ Banda de operación: L➤ Altura promedio sobre la Tierra: 568Km.➤ Periodo de cobertura: 44 días. |
|--|

1.4.3.5. - ALOS (Advanced Land Observing Satellite)



ALOS es un observador satelital que consiste en dos satélites. Uno, para la observación atmosférica y marina, y el otro para la observación de la superficie terrestre principalmente. ALOS es usado para estudios de cartografía, observación regional y monitoreo de desastres. Utiliza la banda L.

Tiene dos modos de operación, los cuales se muestran en las tablas 1.10 y 1.11 respectivamente.

Modo Estándar:

Característica	Polarización	Polarización dual
Ancho de banda Mhz.	28	14
Resolución	7 – 44 metros	14 – 88 metros
Tamaño de la imagen	40 – 70 Km	40 – 70 Km
Recolección de datos	240Mbps	240Mbps

Tabla 1.10

Modo SCANSAR:

Característica	Modo SCANSAR
Ancho de banda Mhz.	14 – 28
Resolución	100 metros
Tamaño de la imagen	250 – 350 Km
Recolección de datos	120 -240Mbps

Tabla 1.11

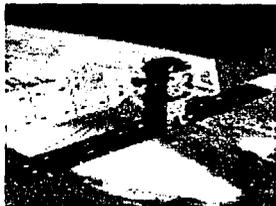
Sus sensores son:

- PALSAR (Phased Array L-band Synthetic Aperture Radar). Este sensor tiene una polarización dual, como la descrita arriba.
- PRISM (Panchromatic Remote Sensing Instruments for Stereo Mapping). Puede generar una imagen de tamaño 3 – 5 metros y tiene un barrido de 35 – 70 Km.
- AVNIR-2 (Advanced Visible and Near Infrared Radiometer-2). Puede generar un barrido de 70Km, genera datos de 160Mbps con compresión de 120Mbps, tiene cuatro bandas espectrales:
 - 0.42-0.50 micrómetros
 - 0.52-0.60 micrómetros
 - 0.61-0.69 micrómetros
 - 0.76-0.89 micrómetros

sus características generales son:

- | |
|---|
| <ul style="list-style-type: none">➤ Fecha de lanzamiento: Agosto 2002.➤ Periodo de cobertura: 46 días.➤ Altura promedio desde la Tierra: 700Km. |
|---|

1.4.3.6.- RADARSAT



RADARSAT es un satélite desarrollado por Canadá conjuntamente con la NASA (National Aeronautics and Space Administration) y la NOAA (National Oceanic and Atmospheric Administration). Fue diseñado para penetrar las nubes y obtener imágenes que se usan para los estudios de exploración de petróleo y gas, agricultura y como soporte de industrias comerciales como son las dedicadas a la pesca y la transportación.

RADARSAT usa el sensor SAR (Synthetic Aperture Radar), el cual, como ya lo habíamos descrito anteriormente es un sensor activo de microondas, este sensor utiliza la banda C y una polarización HH (recepción horizontal y transmisión horizontal), además es selectivo, dependiendo de los diferentes ángulos que muestre el terreno de observación, aplica ciertas condiciones. RADARSAT tiene varios modos de operación, los cuales se muestran en la figura 1.9 y la tabla 1.12.

Las características generales del satélite RADARSAT son:

- Altura promedio sobre la Tierra: 793-821Km.
- Periodo de cobertura: 24 días.
- Bandas espectrales: Banda C.
- Fecha de lanzamiento: Noviembre 1995.
- Tamaño de la imagen: 130 X 130km².
- Resolución: 25 X 25m².

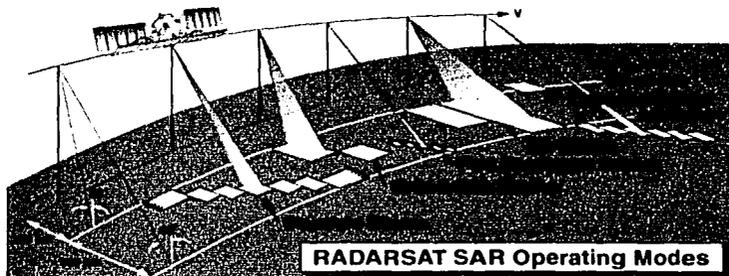
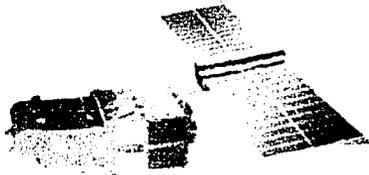


Figura 1.9. – Modos de operación del RADARSAT

Modo	Resolución (m)	Tamaño de la imagen (km)
Fino (Fine)	9.1 x 8.4	50 X 50
	8.7 x 8.4	
	8.4 x 8.4	
	8.1 x 8.4	
	7.8 x 8.4	
Estándar (Standard)	26.0x27.0	110 X 100
	22.0x27.0	
	27.6x27.0	
	25.7x27.0	
	24.2x27.0	
	22.1x27.0	
20.1x27.0		
Ancho (Wide)	35.5x27.0	165 x 165
	26.0x27.0	150x150
	22.8x27.0	130x130
ScanSAR Angosto (Narrow)	50	300 x 300
ScanSAR Ancho (Wide)	100	510 x 500 470 x 500
Extendido Bajo (Low)	36.3x27.0	169 x 170
Extendido Alto (High)	19.8x27.0	85 x 75
	19.4x27.0	
	19.1x27.0	
	18.5x27.0	
	18.2x27.0	
	18.0x27.0	

Tabla 1.12. – Características de los modos de operación del satélite RADARSAT

1.4.3.7.- SPOT (Sisteme Provatoire d'Observation de la Terre)



El SPOT fue desarrollado por Francia, con la colaboración de Bélgica y Suecia, cuenta con 3 satélites en órbita que fueron lanzados en los años 1986, 1990 y 1993.

Este satélite cuenta con dos sensores denominados HRV (*Haute Resolution Visible*), que posibilitan la obtención de imágenes en 2 modalidades: pancromático (PA) y multibanda (XS_n). La modalidad pancromática se encuentra en el rango electromagnético visible y su resolución es de 10m; mientras que la multibanda abarca el rango electromagnético verde, rojo e infrarrojo cercano, con resolución de 20metros.

Sus principales características son:

- Fecha de lanzamiento: Noviembre 1985.
- Altura promedio sobre la Tierra: 822Km.
- Periodo de cobertura: 26 días.
- Tamaño de la imagen: 60 X 60Km².
- Resolución: 20 X 20m² y 10 X 10m².

En el primer trimestre del año 2002, se lanzó el SPOT-5, el cual tiene un ciclo orbital de 26 días, se espera que tenga una vida nominal de 5 años y cuenta con una capacidad de observación dentro del espectro visible, contando con una resolución de 2.5 metros sobre un barrido de hasta 60Km.

Cuenta con capacidades estereoscópicas (también llamadas capacidades HRS). Esta capacidad estereoscópica está compuesta por dos telescopios fijos con una inclinación de 20°, uno orientado hacia adelante y otro hacia atrás de la órbita. Suministrará segmentos de 120Km de alcance y una longitud de 600Km, como se muestra en la figura 1.10.

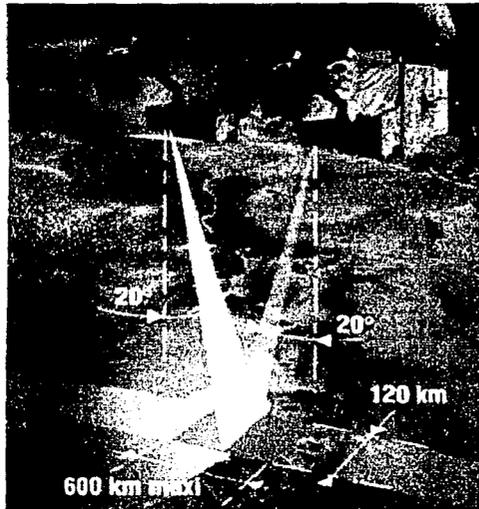


Figura 1.10. – Capacidades estereoscópicas del satélite SPOT-5

Las imágenes HRS están compuestas por líneas de 12.000 píxeles de 10 metros cada una y una separación de 5 metros solamente entre dos líneas consecutivas. El objetivo es obtener una precisión inferior a 10 metros al 90%.

La capacidad máxima que tendrá de cubrimiento diario será de 126.000Km².

Contará también con un nuevo instrumento imageador llamado HRG (Alta Resolución Geométrica), cuyas características son:

- Para imagen pancromática: de 3 a 5 metros de resolución.
- Para imagen multispectral: 10 metros de resolución

Sus aplicaciones son innumerables pero las principales se dan en las siguientes áreas:

- Telecomunicaciones (Modelos de propagación).
- Cartografía.
- Aplicaciones militares.

Algunos de los avances tecnológicos que lleva consigo el SPOT 5 son los siguientes:

- Captores CCD de 12.000 píxeles.
- Sistemas de compresión de datos.
- Memorias de tipo SSM (Solid State Memory) de gran capacidad.
- Capacidad de flujo de datos: 5 canales digitales con 50Mbits /s.

1.4.3.7. - IKONOS



El satélite IKONOS se puso en órbita en Septiembre de 1999, cuenta con capacidad multispectral:

El tamaño del píxel es de 4 metros, y es capaz de adquirir 4 bandas de color: rojo, verde, azul e infrarrojo próximo, siendo la adquisición simultánea en los modos pancromático y multispectral.

La fusión de ambos tipos de datos, permite generar, imágenes de tamaño de píxel de 1 metro y color real o bien, falso color.

Respecto a la resolución radiométrica, IKONOS adquiere imágenes codificadas a 11 bits por píxel, es decir, permite trabajar con 2048 tonos reales de gris.

Las imágenes generadas por este satélite se utilizan principalmente en agricultura de precisión, telecomunicaciones, y aspectos de urbanidad.

Las características de cada modo se presentan en la tabla 1.13.

Característica	Modo Pancromático	Modo Multispectral
Resolución (metros)	1	4
Rango espectral (µmetros)	0,45 - 0,90	# 1: Azul 0,45 - 0,52 # 2: Verde 0,52 - 0,60 # 3: Rojo 0,63 - 0,69 # 4: Cerca de 0,76 - 0,90
Tamaño de la imagen	11km x 100 kilómetros	11km x 100 kilómetros

Tabla 1.13. – Modos de operación del satélite IKONOS

Sus características generales son:

- Altura promedio sobre la Tierra: 681Km.
- Sensor: Pancromático y Multispectral. El pancromático genera imágenes de tamaño de píxel de 1 metro, mientras que el multispectral la genera de 4 metros
- Periodo de cobertura: 98 minutos.

CAPÍTULO 2

ANÁLISIS POR TEXTURAS

2.1. - INTRODUCCIÓN

El análisis de texturas es un campo de investigación que ha tenido un gran auge en los últimos años. Se han utilizado diferentes enfoques, pero aún no se encuentra un modelo que permita generar una gran variedad de texturas. Entre los modelos más interesantes se pueden distinguir los modelos estructurales, los modelos estadísticos y los modelos espectrales.

En este capítulo analizaremos los diferentes métodos (modelos o técnicas) utilizados para el análisis de texturas, así como algunos problemas que han llegado a ser de suma importancia en los campos de investigación. También veremos algunas aplicaciones que estas técnicas han tenido.

2.2. – DEFINICIÓN DE TEXTURA

Aunque no existe una definición formal de textura, este descriptor proporciona medidas de propiedades como suavidad, regularidad y densidad.

Trataremos de explicar a partir de varias definiciones no formales lo que es una textura:

Definición 1. - Manifestación espacial organizada de un arreglo local de tonos de píxeles de la imagen.

Definición 2. - Percepción visual debida a la repetición de patrones en una imagen.

Definición 3. - Región macroscópica, cuya estructura se atribuye a la repetición de patrones de elementos situados según cierta regla.

De estas definiciones notamos que, la textura es un concepto difícil de definir dada la subjetividad de ésta, pero es posible establecer tres características básicas:

- Un cierto orden local se repite sobre una región que es relativamente grande comparada con el tamaño del orden. El orden local implica que se tiene un modelo matemático que aplicado sobre una subregión presenta las mismas características cuantitativas que otra subregión.
- El orden consiste de un cierto arreglo de partes elementales. Es decir, que los píxeles están arreglados de una forma específica, con una cierta estructura geométrica.
- Las partes están compuestas de entidades aproximadamente uniformes, teniendo más o menos las mismas dimensiones dentro de la región de textura, las partes pueden ocurrir en cualquier dirección.

Entonces, para tener una idea general de lo que es una textura, se podría graficar en un espacio de tres dimensiones una banda cualquiera de una imagen, colocando en el plano x - y la posición de cada píxel y en el eje z su valor, y así obtener una superficie. De aquí se podría decir que:

- Una superficie suave es aquella en donde los valores de los píxeles cambian poco, y sería una superficie aplanada.
- Una superficie rugosa es aquella en donde los valores de los píxeles cambian de manera drástica, formando crestas y valles, entonces esta sería una superficie accidentada.

2.3.-CLASIFICACIÓN DE TEXTURAS.

2.3.1. – Texturas naturales

Se clasifican en:

- a) Texturas que se generan en la naturaleza. Como ejemplo tenemos la arena, el musgo, las rocas, etc.
- b) Texturas que no fueron generadas en una computadora. Como ejemplo tenemos los acabados de paredes, ya sean de ladrillos, de madera, de cemento, etc.

2.3.2. – Texturas artificiales

Las texturas artificiales consisten en arreglos de primitivos puestos sobre un fondo neutro.

Estos primitivos pueden ser segmentos de líneas, puntos, estrellas o caracteres alfanuméricos (Pratt, 1981).

Este tipo de texturas se generan por la repetición regular de patrones específicos (Levine, 1985) y por medio de algún algoritmo.

2.3.3. – Texturas regulares

También conocidas como texturas estructuradas o determinísticas, están usualmente hechas por el hombre y consisten en arreglos regulares de líneas, triángulos, cuadrados, círculos, hexágonos. Como ejemplos se pueden mencionar a la pared de ladrillos o los motivos sobre una tela.

2.3.4. – Texturas irregulares

También conocidas como aleatorias o estocásticas. Estas texturas ocurren de manera natural, por ejemplo, la arena de la playa.

2.4. – TÉCNICAS DE ANÁLISIS DE TEXTURAS

Como ya se había mencionado en la parte introductoria de este capítulo, tenemos tres principales técnicas para el análisis de texturas, las cuales analizaremos a continuación.

2.4.1. – *Técnica Estructural*

La aproximación estructural puede describirse por medio de un juego de primitivas y una colocación que rige una relación espacial entre ellos.

El acercamiento estructural proporciona una descripción simbólica buena de la imagen. Como un modelo, es muy útil para texturas regulares con primitivas bien definidas. Los acercamientos estructurales a la textura han sido clasificados en dos grupos, esto es, texturas débiles y texturas fuertes. Estas descripciones se refieren a la interacción espacial entre las primitivas. Una interacción débil puede ser caracterizada por elementos estadísticos, tales como los que se usan en los métodos estadísticos. Por otro lado, la interacción fuerte requiere los detalles de las relaciones de topología reales.

2.4.1.1 – Método:

Un modelo conveniente para explicar la generación de la textura “real” de una imagen es el que se muestra gráficamente en la figura 2.1.

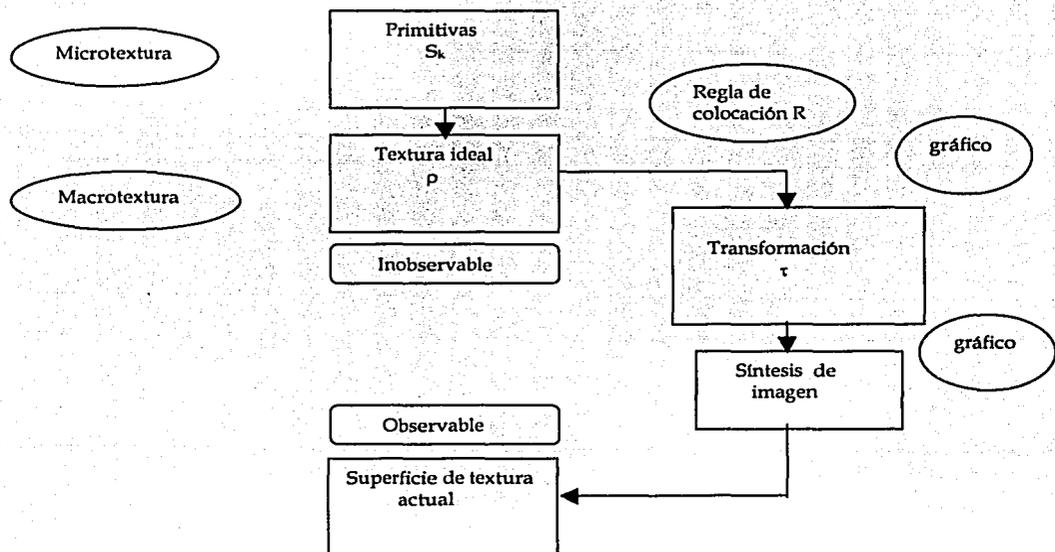


Figura 2.1. – Método de la técnica estructural

Según la figura, sería necesario seleccionar un juego de primitivas de la región S_k y una regla de la colocación R y se obtendría entonces un ρ de la macrotextura. La textura ideal de la macrotextura es inobservable en el sentido que es una versión idealizada de la textura real que se busca. Formalmente, ρ puede ser representado por un gráfico. Así el τ de la transformación es una descripción del gráfico del modelo de la textura ideal ρ en otro gráfico. La imagen $I(i,j)$ es creada entonces por métodos de gráficos de computadora de síntesis de imagen. En el caso de una textura débil, τ es azar; para las texturas fuertes, tiende a ser sistemático.

El análisis de la textura busca generar la textura ideal y el τ de la transformación para una $I(i,j)$ dados. Lo que se debe hacer es buscar una $\rho(i,j)$ de la cual se deben determinar los elementos de la textura, las primitivas de S_k y la regla de colocación R . Las primitivas S_k no son únicas. Pueden utilizarse muchos sub-patrones elementales para generar la misma textura.

2.4.1.2. - Análisis de texturas débiles:

Examinaremos ahora los medios por los cuales las primitivas S_k pueden aislarse. En las texturas débiles, la interacción espacial entre las primitivas es azar. Por lo que primeramente se deben agrupar los elementos S_k de la textura en regiones. Esto se cumple haciendo uso de los principios operativos de similitud y proximidad (similitud es el espaciamiento igual entre los elementos iguales, mientras que proximidad es qué tan próximo está un elemento a otro igual a él).

La segmentación usual de las primitivas de textura involucra la forma anterior. La opción más obvia que encontramos para esto, es la utilización de algoritmos de detección de bordes. Una segunda posible forma es usando las características de campo receptivo, el cual utiliza manchas y bordes convencionales, en este caso la mancha es la que se trata de descubrir. También pueden ser aislados por segmentación de la región. Un modelo artificial a veces se sobrepone intencionalmente en la superficie de un objeto por medio de luz estructurada para facilitar el proceso de segmentación de la imagen.

Habiendo obtenido los elementos de la textura, el próximo paso es agruparlos, para lo cual se utiliza a menudo un método de análisis del histograma, el cual también determina la regla de colocación R . Computando un histograma bidimensional de las posiciones relativas de cada par de elementos de la textura S_k en la imagen. Cada cresta en el histograma define un vector relativo que caracteriza las propiedades de adyacencia del modelo. Se asume que la imagen texturizada es azar si no hay ningún grupo en el histograma. Así el arreglo espacial de los elementos de la textura es gobernado por vectores de regularidad.

También por el método del histograma se pueden encontrar medidas de intensidad, promedio, mínimo, máximo, y/o contraste local, además del arreglo espacial de los elementos.

2.4.1.3. – Análisis de texturas fuertes

En este caso, las propiedades estructurales de R definen los detalles del modelo, dados los S_k a los elementos. La organización espacial representada por R puede describirse estadísticamente o estructuralmente. Las relaciones estructurales entre las primitivas implican una definición de R en forma sintáctica o gramática. Esto se logra usando gramáticas del perfil o dirigiendo el problema en dos dimensiones haciendo uso de descriptores.

Lo importante es encontrar las propiedades de R y S_k y si se sabe que el micropatrón de la textura era casi periódico, se usa un atributo de la textura diseñado especialmente para este propósito. Se usan las propiedades de nivel de gris (GLCM) y se define la inercia de la manera siguiente:

$$A = \sum_{y_r=y_1}^{y_r} \sum_{y_s=y_1}^{y_s} y_r - y_s^2 H_{y_r, y_s, d}$$

En el caso ideal (ρ de la textura inobservable) A será una función periódica en las direcciones particulares de d en el que ρ es también periódica. Los ceros de A corresponden al periodo. La magnitud y direcciones especificadas por esta propiedad definen un paralelogramo del periodo que constituye el primitivo de la textura ideal.

Para las texturas reales no es posible garantizar que todos los mínimos de A corresponden a un múltiplo del periodo de la textura. Así los mínimos locales que normalmente no ocurren no son suficientes para delinear el modelo. Los valores relativamente altos de A sugieren que el proceso no es uniforme.

Usando la ecuación anterior, es posible definir el paralelogramo del periodo. Esta forma geométrica sirve para el modelo de la unidad en el caso de una textura casi periódica. La regla de colocación R está definida por dos vectores asociados a S_k .

Los elementos primitivos de la textura son obtenidos por análisis de la región. Además de brillo, área, tamaño, direccionalidad y forma se usan para describir los S_k . El análisis del histograma de estos atributos proporciona las distintas clases de la imagen bajo estudio. Alternativamente, la textura puede constituirse de dos o más tipos diferentes de elementos primitivos.

Teniendo en cuenta las variaciones estadísticas computadas en S_k y R, es posible sintetizar un modelo de la textura artificial similar al original. Las reconstrucciones son influenciadas fuertemente por la habilidad de segmentar elementos apropiados de S_k .

2.4.2.- Técnica Estadística

Esta técnica se caracteriza por utilizar las propiedades estadísticas de la distribución espacial de los niveles de gris en una imagen. La matriz de co-ocurrencia de niveles de gris (GLCM) es un modelo muy usados en esta técnica. A veces el uso de la matriz de co-ocurrencia produce resultados insatisfactorios, algunas razones son las siguientes:

- La matriz depende, no solo de la relación espacial de los niveles de gris, sino que también de la variación de la intensidad de fondo dentro de la imagen.
- La matriz de co-ocurrencia revela información textural de la imagen a través de un vector de desplazamiento $V=(\Delta_x, \Delta_y)$, entonces la elección de este vector es algo problemático.

Pero a pesar de esto, esta técnica es la más usada para describir las texturas. La preferencia por la utilización de una aproximación por matrices de co-ocurrencia es que éstas modelan las relaciones espaciales de los niveles de gris en un patrón de texturas.

También se utilizan dentro de esta técnica, los Campos Aleatorios de Markov (MRF o CAM), los cuales modelan una textura como una realización de un proceso aleatorio local y estacionario. Pasemos a ver a continuación cada uno de estos modelos.

2.4.2.1. – Matriz de Co-ocurrencia

Una de las aproximaciones más simples para describir las texturas es utilizar los momentos del histograma de niveles de gris de la imagen o región, esto es que la matriz de co-ocurrencia se basa en estadísticas de segundo orden, que es la relación espacial de pares de píxeles en una imagen digital de texturas.

La matriz de co-ocurrencia estima con qué frecuencia ocurren en una imagen digital de texturas los pares de niveles de gris de los píxeles, que están separados por cierta distancia y posicionados a lo largo de una cierta región. A veces, la matriz de co-ocurrencia no se utiliza directamente, pero ciertos descriptores calculados se basan en ella.

Si utilizamos solo el histograma para calcular algunos descriptores de las texturas, no estamos tomando en cuenta las relaciones entre los píxeles, entonces hay que tomar en cuenta también la posición de píxeles con la misma intensidad.

Supongamos que establecemos una dirección (horizontal, vertical, etc.) y una distancia (un píxel, dos píxeles, etc.) en una imagen. Entonces el elemento (i,j) de la matriz de co-ocurrencia P para un objeto, es el número de veces que los niveles i y j ocurren en dos píxeles separados por esa distancia antes definida y en la dirección dada dividido por M , siendo M el número de parejas de píxeles que contribuyen a P . La matriz P tiene tamaño $N \times N$ donde N es el número de niveles de gris.

Claramente podemos calcular diferentes matrices de co-ocurrencia para diferentes distancias y direcciones. Sin embargo, estas matrices suelen tener bastantes ceros y para evitarlo, se suele dividir el rango de los niveles de gris a alrededor de 8.

Para explicar lo anterior, supongamos que f es la imagen a analizar con N niveles de gris, y sea $V = (\Delta_x, \Delta_y)$, un desplazamiento, será entonces P la matriz de dimensión $N \times N$,

cuyo elemento (i,j) indica el número de veces que un píxel, con un nivel de gris n_1 , ocurre en la posición V con respecto a un píxel que tiene un nivel de gris n_2 . Si f es:

```

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

y V es $(1,0)$, entonces la matriz de co-ocurrencia, de los pares de f , P es:

```

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

Y la matriz de co-ocurrencia de niveles de gris es $Q = P / M$.

En otras palabras, la matriz de co-ocurrencia fue calculada de la siguiente manera:

	0	1	2	3
0	4	2	1	0
1	2	4	0	0
2	1	0	6	1
3	0	0	1	2

Empezando de izquierda a derecha y de arriba hacia abajo, y auxiliándose del arreglo anterior, se toma el par $(0,0)$ y se cuenta en la matriz f el número de pares que contengan un cero a la izquierda y un cero a la derecha, y ese número es el que se pone en la matriz P , después se toma el par $(0,1)$ y se hace lo mismo, y así sucesivamente hasta terminar con todos los valores.

Los elementos cerca de la diagonal principal de P corresponden a pares de niveles de gris que son iguales, mientras que los elementos que están alejados de la diagonal principal corresponden a los pares que son diferentes.

Se pueden hacer desplazamientos a la izquierda, a la derecha, arriba, abajo. Estos son los más comunes, ya que los desplazamientos angulares diferentes a los anteriores son más difíciles de hacer.

Una vez que la matriz de co-ocurrencia es calculada, se pueden calcular diferentes descriptores de la región, el objetivo es que estos descriptores capturen algunos rasgos de la textura de una imagen.

Haralick (1979) sugirió catorce descriptores texturales, los cuales son clasificados en cuatro grupos. A continuación se analizan algunos de ellos en su grupo correspondiente:

➤ Grupo 1. - Manifiestan rasgos visuales.

Homogeneidad: Es baja cuando los elementos $Q(i,j)$ son todos iguales y alta cuando son diferentes.

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

Contraste: La imagen tiene poco contraste si la concentración de la diagonal principal de la matriz de co-ocurrencia es alta, y es alto en el caso contrario.

$$C = \sum_{i=1}^N \sum_{j=1}^N (i-j)^2 Q(i, j)$$

Correlación: Es una medida de la dependencia lineal del nivel de gris de los pares de píxeles analizados en la matriz de co-ocurrencia.

$$CO = \frac{\sum_i \sum_j (i * j) Q(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

➤ Grupo 2. – Basados en estadísticas:

Varianza: En particular, cuando una área de la imagen está bajo la investigación textural, la variación estadística de primer orden es igual a la variación de GLCM.

Si el vector de desplazamiento de GLCM es igual a 1 y si su ángulo es igual a 0° o 90°. La variación del nivel de gris del vector del desplazamiento tiende a uno, desde que una región puede tener frecuencias espaciales bajas y el valor del contraste bajo mientras su variación puede tener un valor alto o un valor bajo.

$$VAR = \sum_{i=0}^N \sum_{j=0}^N (i - \mu)^2 \times Q(i, j)$$

Probabilidad máxima: Indica cual es el par con valores más probables en f:

$$PM = \underset{i, j}{\text{MAX}} Q(i, j)$$

- Grupo 3. – Basados en la teoría de la información:

Entropía: Es una medida del desorden, logrando su valor más alto cuando los elementos de GLCM son todos iguales, cuando la concentración de la diagonal principal es alta el valor de la entropía es bajo.

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

- Grupo 4. – Basados en medidas de la información de correlación:

Autocorrelación:

$$AU = \sum_i \sum_j (i * j) Q(i, j)$$

- Grupo 5. – Otros:

Energía:

$$\text{Ener} = \sum_i \sum_j Q(i, j)^2$$

Disimilitud:

$$DIS = \sum_i \sum_j |i - j| * Q(i, j)$$

Tendencia de grupo:

$$T = \sum_i \sum_j (i + j - \mu_x \mu_y)^2 Q(i, j)$$

Matíz de grupo:

$$\text{Matíz} = \sum_i \sum_j (i + j - \mu_x \mu_y)^3 Q(i, j)$$

Prominencia de grupo:

$$\text{Promi} = \sum_i \sum_j (i + j - \mu_x \mu_y)^4 Q(i, j)$$

Donde:

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

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

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

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

2.4.2.2. – Campos Aleatorios de Markov

Como ya se había mencionado anteriormente, los Campos Aleatorios de Markov (CAM) modelan una textura como una realización de un proceso aleatorio local y estacionario. Esto es, cada píxel de una imagen está caracterizado por un pequeño conjunto de píxeles vecinos localmente, y esta caracterización es la misma para todos los píxeles.

Para poder explicar lo que es un proceso aleatorio local y estacionario, imaginemos que se muestra una imagen a un observador, pero solo se le permite observarla a través de una pequeña ventana móvil. A medida que la ventana se mueva, el observador podrá ver distintas partes de la imagen: La imagen es estacionaria si, bajo un tamaño apropiado de la ventana, la porción observable siempre tiene un aspecto similar. La imagen es local si cada píxel es predecible o calculable a partir de un pequeño conjunto de píxeles vecinos y es independiente del resto de la imagen, como se muestra en la figura 2.2.

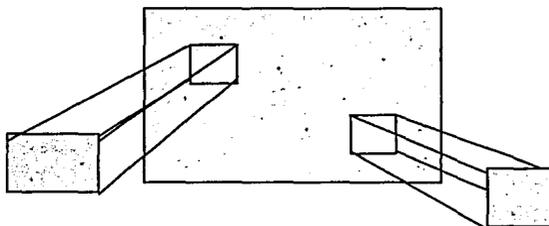


Figura 2.2. - Ejemplo de Localidad y Estacionariedad de una textura

En general, un modelo de textura CAM puede ser especificado por medio de una probabilidad conjunta $p(f)$. La probabilidad determina cómo un patrón de textura ocurre.

A continuación se dan algunas definiciones importantes sobre los CAM

Sea $X(i,j)$ el valor del píxel (i,j) en una imagen P de $N \times N$.

Definición 1. - El punto j se dice vecino del punto i si $p(X(i)|X(1), X(2), \dots, X(i-1), X(i+1), \dots, X(M))$ donde $M = N \times N$ depende de $X(j)$. Esta definición no implica que los vecinos de un punto sean cercanos en términos de distancia, aunque esto es lo usual.

Definición 2. - Un campo aleatorio de Markov (CAM) es una densidad de probabilidad conjunta en el conjunto de todos los posibles X en el arreglo (en este caso imagen) L , sujeto a las siguientes condiciones:

- 1) Positividad: $p(X) > 0$ para todo X .
- 2) Markovianidad: $p(X(i)|\text{todo punto en } M \text{ excepto } i, \text{ es igual a } p(X(i) | \text{vecino de } i)$.
- 3) Homogeneidad: $p(X(i) | \text{vecinos de } i$ depende solo de la configuración de los vecinos y es invariante entre las traslaciones.

En la mayoría de los casos, solo interesan los modelos donde el punto i es vecino del punto j cuando i es cercano a j . Sin embargo, la influencia de la vecindad depende del modelo de función de probabilidad que se haya elegido.

Definición 3. - El orden de un proceso de un campo aleatorio de Markov ($O(N)$) en un arreglo es el valor máximo de i tal que $b(i,1)$ o $b(i,2)$ son distintos de cero.

Definición 4. - Un campo aleatorio de Markov es *isotrópico de orden i* si $b(i,1)=b(i,2)$. En otro caso, se dice que es *anisotrópico de orden i*.

Existen varios modelos de CAM, tales como el modelo autobinomial, el modelo gaussiano y el modelo de pirámides de imágenes (multiniveles). Un modelo CAM en particular, tiende a favorecer a la correspondiente clase de texturas, por medio de una asociación entre ellas con probabilidades más grandes que otras. De aquí que un modelo CAM sea una función de probabilidad en particular, la cual está especificada por su forma y parámetros en esta.

A continuación se da un pequeño resumen de cada uno de los modelos de CAM:

➤ **Modelo CAM Autobinomial**

En este modelo, la probabilidad de que el punto $X(i,j)$ tenga un nivel de gris n es binomial, con parámetro T , determinado a través de sus vecinos.

$$\theta(T) = \frac{\text{Exp}(T)}{1 + \text{Exp}(T)}$$

donde la forma del modelo de primer orden es:

$$T_1 = a + b(1,1)(t+t') + b(1,2)(u+u')$$

y la forma del modelo de segundo orden es:

$$T_2 = T_1 + b(2,1)(v+v') + b(2,2)(z+z')$$

Los términos de orden superior pueden obtenerse sumando las expresiones correspondientes, tal como se muestra en la figura 2.3:

		m		
	v	u	z	
l	t	X	t'	l'
	z'	u'	v'	
		m'		

Figura 2.3. - Vecinos del punto X

➤ **Modelo CAM gaussiano**

En este caso, el nivel de gris de un píxel tiene una distribución conjunta gaussiana y una correlación controlada por el número de parámetros que representan la dependencia estadística de un píxel de los valores de los píxeles en una vecindad simétrica. La expresión de la función de densidad de probabilidad es la siguiente:

$$P(X(i)|X(i+r), r \in N) = \frac{1}{\sqrt{2\pi v}} \exp \left[-\frac{1}{2v} \left(X(i) - \sum_{r \in N} \theta_r X(i+r) \right)^2 \right]$$

➤ **Modelo CAM Pirámides de Imágenes**

La implementación CAM mediante pirámides de imágenes corresponde a un enfoque diferente a los ya expuestos. Las pirámides de imágenes corresponden a un tipo particular de transformación sub-banda.

La característica fundamental de una pirámide de imágenes es que las funciones base de la transformación son copias dilatadas y trasladadas en un factor 2^j , para algún j una de otra. Esta transformación sub-banda se calcula mediante convoluciones y submuestreos sucesivos.

El conjunto de imágenes generado a partir de la aplicación de esta transformación corresponde a una representación multiresolución de la imagen.

2.4.3.- Técnica Espectral

En esta técnica se usa el espectro de Fourier, el cual es ideal para describir la direccionalidad de casi todo el periodo de una imagen en dos dimensiones. Estos modelos de textura globales, aunque fácilmente discernible, generalmente son bastante difíciles de descubrir con los métodos espaciales debido a la naturaleza local de esta técnica.

Se consideran tres características del espectro de Fourier las cuales son útiles para la descripción de la textura:

(1) Las crestas más grandes en el espectro dan la dirección principal de los modelos de la textura.

(2) La localización de las crestas en el plano de la frecuencia da el período espacial fundamental de los modelos.

(3) Eliminando cualquier componente no periódico vía filtración de elementos periódicos de la imagen, ya que estos elementos pueden describirse entonces por las técnicas estadísticas.

Dado que el espectro de una imagen real es simétrico sobre el origen, tan sólo la mitad de la frecuencia necesita ser considerada. Así con el propósito del análisis, cada modelo periódico es asociado con sólo una cresta en el espectro, en lugar de dos.

La detección e interpretación de las características del espectro a menudo son simplificadas expresando el espectro en coordenadas polares con una función $S(r, \theta)$, donde S es la función del espectro y r y θ son las variables en este sistema de coordenadas. Para cada θ de la dirección, $S(r, \theta)$ puede ser considerada una función de una dimensión. De manera similar, para cada frecuencia r , $S(\theta)$ es una función de una dimensión. Se analiza $S_\theta(r)$ para un valor fijo de valores de θ , la conducta del espectro (como la presencia de crestas) a lo largo de una dirección radial del origen, También se analiza $S_r(\theta)$ para un valor fijo de valores de r , la conducta a lo largo de un círculo centrado en el origen.

Una descripción más global se obtiene integrando (sumando para las variables discretas) estas funciones:

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

y

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

donde R es el radio de un círculo centrado en el origen. Para un espectro de $N \times N$, el valor de R típico se escoge como $N/2$.

El resultado de las ecuaciones anteriores constituyen un par de valores $[S(r), S(\theta)]$ para cada par de coordenadas (r, θ) . Variando estas coordenadas, se pueden generar dos funciones de una dimensión, $S(r)$ y $S(\theta)$, las cuales constituyen una descripción de la energía espectral de la textura para una imagen entera o una región bajo consideración. Además, los descriptores para estas funciones pueden computarse para obtener características cuantitativas. Los descriptores que se usan típicamente para este propósito son la localización del valor más alto, la media, variación de la amplitud, variaciones axiales, y la distancia entre la media y el valor más alto de la función.

Una de las aplicaciones de este tipo de técnicas es la de reconocimiento de patrones. Esto es que puede usarse para diferenciar entre dos modelos de textura, para lo cual se deben graficar primero el espectro de la textura que sea periódica, después se debe graficar $S(r)$ y $S(\theta)$. La gráfica de $S(r)$ es una estructura típica, mientras se tiene la energía más alta cerca del origen y los valores más bajos progresivamente para las frecuencias superiores. La gráfica de $S(\theta)$ muestra las crestas prominentes a los intervalos que claramente corresponden a la periodicidad en la textura de la imagen. La diferenciación entre los dos modelos de la textura se da analizando su $S(\theta)$ correspondiente, las formas de onda serían claramente deladoras.

CAPÍTULO 3

TÉCNICAS DE SEGMENTACIÓN

3.1. – DEFINICIÓN DE SEGMENTACIÓN

Definición 1. - Es el proceso por el cual una imagen queda subdividida en las partes u objetos que la forman.

Definición 2. - Es la etapa de un sistema de visión en la que se extraen los objetos de interés.

Antes de aplicar un proceso de segmentación sobre una imagen, es importante el haber realizado una etapa de pre-procesamiento adecuada sobre esta, pues entre otras cosas podría ocurrir que se identificara el ruido como un objeto de interés.

Para poder explicar un poco mejor este concepto veamos la figura 3.1:

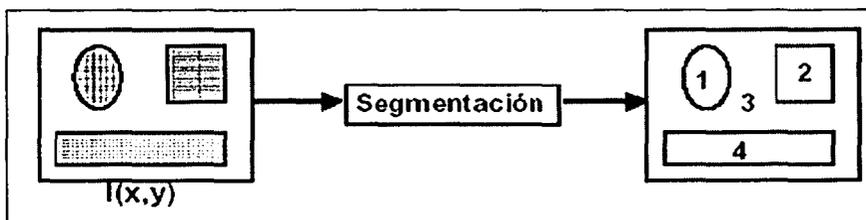


Figura 3.1. - Concepto de segmentación

$I(x,y)$ es la imagen bajo análisis, la cual pasa por un proceso de segmentación y entonces se genera una imagen dividida, extrayendo los objetos que nos interesan.

En general, el proceso de segmentación es particionar las imágenes en algunas zonas de interés, correspondientes a objetos presentes en la escena.

La segmentación subdivide una imagen en sus partes constituyentes u objetos. El nivel al que se lleva a cabo esta subdivisión depende del problema a resolver.

3.2. – OBJETIVO DE LA SEGMENTACIÓN BINARIA

El objetivo de la segmentación es el de evaluar si cada píxel de la imagen pertenece o no al objeto de interés. Esta técnica de procesamiento de imágenes genera una imagen binaria, donde los píxeles que pertenecen al objeto se representan con un uno, mientras que los que no pertenecen al mismo se representan con un cero. Este tipo de particionamiento está basado en el análisis de alguna característica de la imagen, tal como los niveles de gris o la textura.

3.3. – PROPIEDADES DE LAS TÉCNICAS DE SEGMENTACIÓN

Las técnicas de segmentación de imágenes monocromáticas generalmente se basan en una de las dos propiedades básicas de los valores del nivel de gris: discontinuidad y similaridad.

3.3.1. - *Discontinuidad*

En la propiedad de discontinuidad (o medida de contraste con los objetos y sus bordes), el método general consiste en dividir una imagen basándose o teniendo en cuenta los cambios bruscos del nivel de gris

Las principales áreas de interés dentro de esta categoría son:

- Detección de puntos aislados en una imagen.
- Detección de líneas en una imagen.
- Detección de bordes en una imagen.

3.3.2. – *Similaridad*

En la propiedad de similaridad, se realiza la división de una imagen en objetos teniendo en cuenta la similitud que presentan ciertas regiones adyacentes de la imagen.

Las principales áreas de interés dentro de esta categoría son:

- Técnicas de umbralización.
- Crecimiento de regiones.
- División y fusión de regiones.

Estos conceptos se aplican a imágenes estáticas y dinámicas. El movimiento puede alterar el rendimiento de las técnicas de segmentación.

Dentro de estas dos grandes categorías mencionaremos las técnicas que abarcan cada una de ellas

3.4. – TÉCNICAS DE SEGMENTACIÓN

Las técnicas de segmentación pueden dividirse en dos categorías principales:

3.4.1. - Segmentación orientada a regiones

Se trata de encontrar regiones homogéneas hasta que dejan de cumplir al criterio de homogeneidad¹.

Este tipo de técnicas tienen por objetivo, el construir las regiones directamente, solamente tomando en cuenta el valor de gris que el mismo tiene asociado, lo que conlleva a que puntos aislados o pequeñas áreas pueden ser clasificadas como pertenecientes a la región de interés.

Estos métodos de segmentación toman en cuenta un conjunto de puntos de la imagen, a los cuales se les analiza características de posición en el espacio de intensidades, características de conectividad y de frontera entre dos conjuntos. Dependiendo de cómo sea analizada la posición en el espacio y las relaciones espaciales existentes entre los píxeles, se pueden encontrar métodos por clasificación, métodos por crecimiento de regiones, métodos de división y fusión de regiones, y métodos basados en el análisis del histograma. A continuación se analizarán algunos de ellos.

3.4.1.1. – Segmentación por Clasificación .

Se utilizan los niveles de gris presentes en la imagen para obtener una partición del espacio. Se asocia a cada píxel la clase de nivel de gris a la cual pertenece. Las regiones son definidas por los conjuntos de píxeles conexos pertenecientes a una misma clase.

Este método utiliza el cálculo del histograma para realizar la clasificación de las intensidades de la imagen, realizando la búsqueda de los distintos modos del histograma y sus valles correspondientes; las clases son determinadas por los valores entre los valles, de esta forma los puntos de la imagen son etiquetados con la clase correspondiente.

3.4.1.2. – Segmentación por Crecimiento (unión) de Regiones

El crecimiento de regiones se realiza mediante la adición de píxeles a una u otra región si estos píxeles cumplen con determinada propiedad.

En general la propiedad viene a ser nivel de gris, color o textura.

¹ Homogeneidad: Conjunto de características similares

3.4.1.2.1. – Crecimiento de regiones por agregación de píxeles.

Este crecimiento se realiza partiendo de un conjunto de píxeles semilla que denotan a las distintas regiones que se quieren segmentar. A partir de estos píxeles semilla se hace crecer la región añadiendo a dichos píxeles semilla aquellos vecinos que tienen propiedades similares.

Estos puntos semillas servirán como puntos de comienzo del proceso de crecimiento de las regiones, con lo cual, el número final de regiones ha de ser como mucho igual al número de semillas sembradas por el usuario (puede ser menor, pues en algún paso del método se puede decidir unir dos regiones para formar una sola).

Para poder realizar la agregación de píxeles similares será necesario definir el concepto de similaridad, que no tiene porque ser el mismo para todo tipo de aplicaciones.

La propiedad o concepto de similaridad se representa con la forma de un predicado. Cualquier píxel que cumpla con la propiedad dada es asignado a la región.

Posibles criterios, ya utilizados en algoritmos desarrollados, pueden ser que la diferencia entre el valor del píxel a agregar y el valor de la semilla, o el valor medio de la región ya formada, sea menor que un cierto umbral predeterminado.

Para poder comprender un poco más esta técnica veamos el siguiente ejemplo donde las dos semillas introducidas están marcadas por un subrayado:

0	1	5	8	7
1	2	6	<u>6</u>	5
3	<u>2</u>	0	0	7
2	2	8	6	5
7	5	2	2	2

Los números que se encuentran dentro de las celdas son niveles de gris de una imagen dada, y considerar que se ha definido que un punto y una semilla original son similares si y sólo si su diferencia de nivel de intensidad es menor o igual que 2. Si se usa una conectividad 4, el método genera las siguientes iteraciones (a partir de cinco iteraciones ya no es posible realizar más agregaciones de píxeles).

0	1	5		7	0	1				0	1			
1	2				1	2				1	2			
3	2	0	0	7	3	2	0	0		3	2	0	0	
2	2	8	6	5	2	2	8	6	5	2	2	8	6	
7	6	2	2	2	7	6	2	2	2	7	5	2	2	2

0	1				0	1			
1	2				1	2			
3	2	0	0		3	2	0	0	
2	2	8			2	2			
7	6	2	2	2	7	6	2	2	2

En la primera iteración se nota como se unieron los vecinos con nivel igual o menor que 2, en la segunda notamos que se siguen uniendo, hasta llegar al final. Se nota que las celdas en blanco son píxeles que no cumplieron con el criterio de similitud.

El comienzo con un píxel semilla en particular, permitiendo que esta región crezca completamente antes de tratar otras semillas puede tener varios efectos:

- Crecimiento de regiones dominantes: Ambigüedades en torno a los bordes de regiones adyacentes pueden no ser resueltas correctamente.
- La elección de diferentes puntos semilla, puede dar lugar a diferentes segmentaciones.
- Pueden surgir problemas si un píxel semilla (elegido arbitrariamente) pertenece a un borde.

Este método presenta además de los inconvenientes ya mencionados, el problema de un criterio para detener el método.

3.4.1.2.2. – Crecimiento de regiones por enlace sencillo.

Este método de segmentación considera a cada píxel como un nodo de un grafo, luego analizando una vecindad de un píxel cualquiera, se unen por medio de un arco, todos aquellos píxeles con características similares.

La característica de similitud buscada es la intensidad de los niveles de gris, en tal sentido, dos píxeles serán conectados si los mismos difieren en menos de un valor constante preestablecido, y si se encuentran en una determinada vecindad.

Todos los conjuntos de los píxeles así encontrados, forman un sub-grafo de la imagen.

El proceso se repite hasta que las regiones son conjuntos máximos de píxeles que generan la imagen segmentada.

Los pasos del método son:

1. Definir una segmentación inicial que cumpla el criterio de homogeneidad.
2. Definir un criterio para unir regiones adyacentes.
3. Unir las regiones adyacentes si cumplen con el criterio de la unión. Se detiene cuando no puedan unirse dos regiones sin romper el criterio de homogeneidad.

La implementación más simple empieza la unión de regiones con la segmentación de regiones de tamaños 2×2 , 4×4 u 8×8 . Las descripciones de las regiones se basan en los niveles de gris, para lo cual se hace uso del histograma, la descripción de una región se compara con la de otra adyacente, si coinciden, se unen las regiones.

3.4.1.2.3. – Crecimiento de regiones por enlace de centroide

El método de crecimiento de regiones por enlace de centroide, no utiliza la comparación de la similaridad de dos píxeles vecinos, para realizar su unión.

Este método explora la imagen de una manera predeterminada. El valor de un píxel es comparado con la media de las intensidades de una zona o segmento de su vecindad. Si el valor del píxel y el valor medio del segmento están lo suficientemente cercanos, entonces el píxel es adicionado al segmento y su media es actualizada.

Si regiones no vecinas tienen una media suficientemente cercana, entonces un nuevo segmento es establecido, siendo el píxel en estudio su primer miembro. Si dos regiones diferentes están lo suficientemente cercanas, las dos regiones son unidas y el píxel es adicionado a la nueva región.

3.4.1.3. – Segmentación por División (separación) y Fusión (unión) de regiones.

Esta técnica subdivide una imagen en un conjunto de regiones uniformes y luego las une o las separa en un intento de que todas ellas:

- 1) Cubran la imagen completa.
- 2) Cada una de ellas esté conectada.
- 3) Sus intersecciones dos a dos sea el conjunto vacío.
- 4) Algún predicado de homogeneidad sea verdadero en cada región.
- 5) La unión de dos cualesquiera de ellas no cumpla algún predicado de homogeneidad.

El método divide progresivamente la imagen en regiones mientras no se cumpla un predicado P_1 y después se van fusionando las regiones obtenidas mientras se cumpla un predicado P_2 .

Normalmente, el método empieza con la hipótesis de que la imagen completa es una única región, entonces analiza la homogeneidad de la misma (mediante un cierto criterio y propiedades). Si existe homogeneidad, la imagen se encuentra ya segmentada, si no es así, entonces la región es dividida siempre en 4 regiones.

Este proceso se repite para cada una de las regiones generadas hasta que el proceso de división no puede llevarse a cabo. Una vez que se ha llevado a cabo el proceso de división, se comprueba para cada región generada, si es posible, unirla a una región adyacente (lógicamente si los píxeles satisfacen el criterio de homogeneidad establecido). El proceso termina cuando no se pueden fusionar más regiones.

A continuación analizaremos el método de división.

3.4.1.3.1. – División de regiones

Sea el predicado $P_1 =$ desviación estándar de la imagen ≤ 10

Sea la imagen siguiente:

A			B		
129	127	132	148	153	159
129	130	128	142	151	159
C			D		
126	129	133	152	147	150
160	164	150	141	149	160

Media: 143.66
Desviación estándar: 12.40

Entonces, tenemos las siguientes subregiones:

A	Desviación estándar:	1.7						
<table border="1"> <tr><td>129</td><td>127</td><td>132</td></tr> <tr><td>129</td><td>130</td><td>128</td></tr> </table>	129	127	132	129	130	128		
129	127	132						
129	130	128						

B	Desviación estándar:	6.6						
<table border="1"> <tr><td>148</td><td>153</td><td>159</td></tr> <tr><td>142</td><td>151</td><td>159</td></tr> </table>	148	153	159	142	151	159		
148	153	159						
142	151	159						

C	Desviación estándar:	16.5						
<table border="1"> <tr><td>126</td><td>129</td><td>133</td></tr> <tr><td>160</td><td>164</td><td>150</td></tr> </table>	126	129	133	160	164	150		
126	129	133						
160	164	150						

D	Desviación estándar:	6.2						
<table border="1"> <tr><td>152</td><td>147</td><td>150</td></tr> <tr><td>141</td><td>149</td><td>160</td></tr> </table>	152	147	150	141	149	160		
152	147	150						
141	149	160						

De estas cuatro subregiones la única que **no** cumple con el predicado P_1 es la subregión C, entonces subdividimos la subregión C en otras cuatro regiones, pero como ya no sería un arreglo cuadrado, entonces se subdivide en dos regiones:

C			Desviación estándar: 16.5
126	129	133	
160	164	150	

E			Desviación estándar: 3.5
126	129	133	

F			Desviación estándar: 7.2
160	164	150	

Ya que ninguna de las regiones obtenidas tiene una desviación estándar mayor de 10, termina el proceso de división.

Una representación de la imagen vendría dada por:

A	A	A	B	B	B
A	A	A	B	B	B
E	E	E	D	D	D
F	F	F	D	D	D

3.4.1.3.2. – Fusión de regiones

Una vez realizada la división de regiones ahora lo que hay que hacer es la fusión de regiones. Hay que definir un predicado P_2 , y la unión que **sí** cumpla con el predicado P_2 , se fusionan.

Sea $P_2 = \text{“Desviación en dos regiones adyacentes } \leq 10\text{”}$

Las fusiones se realizan comprobando pares de regiones, si dos regiones se fusionan ya no pueden compararse con otras de la misma iteración.

Para realizar la fusión se utiliza un método de unión (crecimiento) de regiones.

$\Rightarrow A \cup B$ Desviación estándar: 12.1

$\Rightarrow A \cup E$ Desviación estándar: 2.2 **sí** cumple, se fusionan A y E y a esta región le llamaremos G.

$\Rightarrow B \cup D$ Desviación estándar: 6.2 **sí** cumple, se fusionan B y D y a esta región le llamaremos H.

$\Rightarrow F$ ya no se comprueba, pues E y D que son adyacentes a F, ya están fusionadas.

Esto da como resultado la siguiente representación de la imagen:

A	A	A	B	B	B
A	A	A	B	B	B
E	E	E	D	D	D
F	F	F	D	D	D

G	G	G	H	H	H
G	G	G	H	H	H
G	G	G	H	H	H
F	F	F	H	H	H

Ahora, hay que fusionar la nueva imagen:

- ⇒ G U H Desviación estándar: 12.1
- ⇒ G U F Desviación estándar: 13.5
- ⇒ H U F Desviación estándar: 6.8 *sí cumple, se fusionan H y F en la región llamada I.*
- ⇒ G U I Desviación estándar: 19.9

Como ya no hay más pares que comprobar, el proceso de fusión finaliza, quedando la representación de la imagen como sigue:

G	G	G	H	H	H
G	G	G	H	H	H
G	G	G	H	H	H
F	F	F	H	H	H

G	G	G	I	I	I
G	G	G	I	I	I
G	G	G	I	I	I
I	I	I	I	I	I

3.4.1.4. – Segmentación basada en el Análisis del Histograma (umbralización).

3.4.1.4.1. – Concepto de umbralización

Establecer un umbral “T” es una forma de segmentar los objetos respecto al fondo, esto es cuando los objetos son oscuros y el fondo es luminoso, aparecen uniformemente iluminados y claramente contrastados, entonces se establece el umbral “T” que separe las crestas de un histograma, de tal forma que los píxeles en los objetos y fondo tienen niveles de gris agrupados en dos modos dominantes.

La forma de extraer los objetos del fondo es seleccionar el umbral “T” que separa estos modos utilizando la siguiente ecuación:

$$g(x,y) = \begin{cases} 0 & \text{si } f(x,y) < T \\ 255 & \text{si } f(x,y) \geq T \end{cases}$$

Pero cuando los objetos no están claramente contrastados, el histograma tiene más de dos crestas, por lo que es necesario utilizar varios niveles, siendo más complicado establecer los umbrales que realmente aislen las regiones de interés.

3.4.1.4.2. – Tipos de umbralización

- Umbral global.- Cuando el umbral T depende sólo de $f(x,y)$.
- Umbral local. – Si T depende tanto de $f(x,y)$ como de alguna propiedad local a dicho punto (por ejemplo, los valores de nivel de gris de sus vecinos.)

Existen diferentes procesos o métodos que se utilizan para obtener el umbral adecuado para una correcta segmentación. Estos son los que veremos a continuación.

3.4.1.4.3. – Método P-cuantil

Este método supone conocido el tamaño del objeto que se desea extraer para segmentar la imagen, por lo tanto se supone que estos ocuparán un porcentaje determinado ($p\%$) de la misma.

Usando este conocimiento se divide el histograma de la imagen buscando el nivel de gris que hace que un $p\%$ de los píxeles pertenezcan a los objetos. Esto se muestra en la figura 3.2:

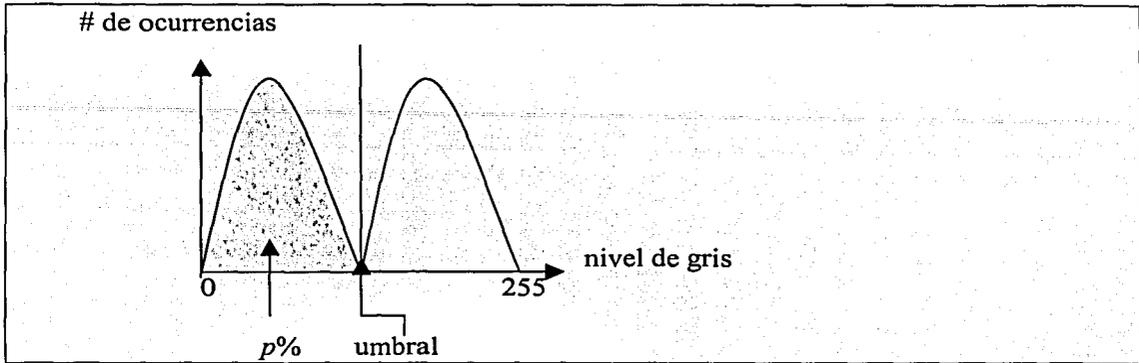


Figura 3.2. – Método P-cuantil

En el histograma anterior se muestra el porcentaje que el método supone que tiene la imagen, y la manera de encontrar el umbral.

Utilizando el conocimiento de la división de la imagen, uno o más umbrales pueden elegirse asignando un porcentaje de píxeles a los objetos.

Este método tiene un uso muy limitado. Solamente unas pocas aplicaciones permiten estimar el área de forma general.

3.4.1.4.4. – Método Isodata

El método Isodata es una técnica iterativa que se utiliza para la obtención del umbral correcto. El histograma es inicialmente segmentado en dos partes utilizando un umbral de comienzo. Este umbral de comienzo se encuentra tomando la mitad del valor máximo del rango de niveles de gris de los que consta la imagen.

A continuación se encuentra la media de los valores asociados con cada una de las partes en que ha quedado segmentado el histograma m_1 y m_2 .

Utilizando esos valores se calcula un nuevo valor de umbral mediante la siguiente ecuación:

$$\alpha = \frac{(m_1 + m_2)}{2}$$

El proceso continúa hasta que en los pasos consecutivos el umbral no cambia.

3.4.1.4.5. – Método de Simetría de Fondo (Background-Simmetry Algorithm)

Este método asume que el fondo (background) de la imagen forma un pico simétrico y dominante en el histograma.

El método busca a la derecha de la parte perteneciente al objeto (en la parte correspondiente al fondo), el valor de la intensidad que corresponde a un cierto porcentaje de nivel de gris ($p\%$) y, como se supone que la cresta del fondo es simétrica, se toma como umbral el máximo menos un desplazamiento igual al de porcentaje p :

$$\text{Entonces } T = T_{\max} - (p\% - T_{\max})$$

Para una mayor comprensión de este método consideremos que $p = 95\%$, entonces veamos la figura 3.3:

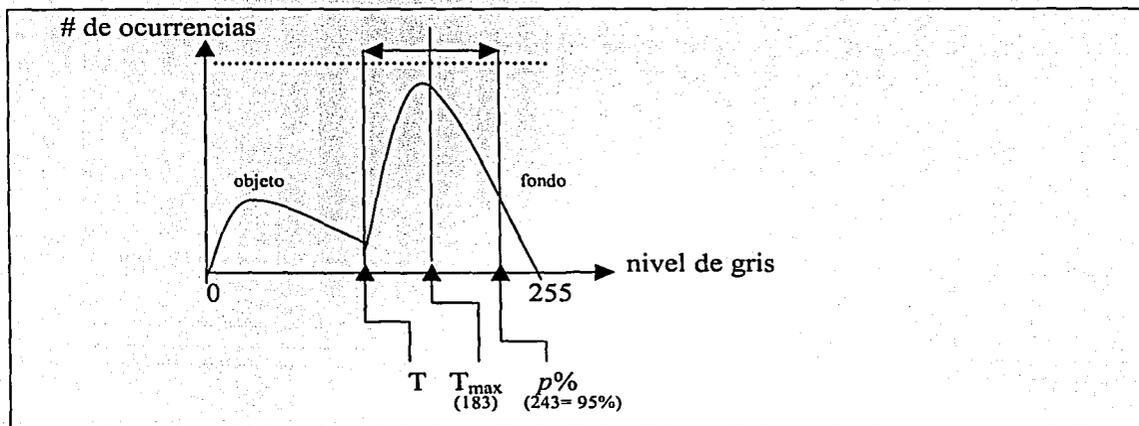


Figura 3.3. – Método Background

Si se toma como porcentaje el 95%, el valor de luminosidad se encuentra a la derecha del pico. Para este valor de brillo (216), el 5% de los puntos se encuentra a la derecha del pico. Como suponemos que existe una simetría en torno a dicho pico, se utiliza como umbral un desplazamiento a la izquierda del pico máximo igual al desplazamiento que existe desde el pico máximo hasta la localización del porcentaje marcado ($p\%$).

Para este caso:

$$T = T_{\max} - (p\% \cdot T_{\max}) \longrightarrow T = 183 - (243 - 183) = 183 - 60 = 123.$$

Entonces el umbral utilizado tiene el valor de 123.

3.4.1.4.6. – Método del Triángulo (Triangle Algorithm)

Esta técnica se basa en la detección del umbral correcto.

El método consiste en construir una línea entre el valor máximo de nivel de gris (v_{\max}) del histograma y el valor más bajo de nivel de gris (v_{\min}). La distancia entre la línea y el histograma ($h(v)$) se calcula para todos los valores de nivel de gris, desde v_{\min} hasta v_{\max} . Entonces se escoge el valor de gris para el cual la distancia entre $h(v_0)$ y la línea es máxima, siendo el umbral $T = v_0$.

Para comprender un poco mejor esta técnica analizaremos la figura 3.4:

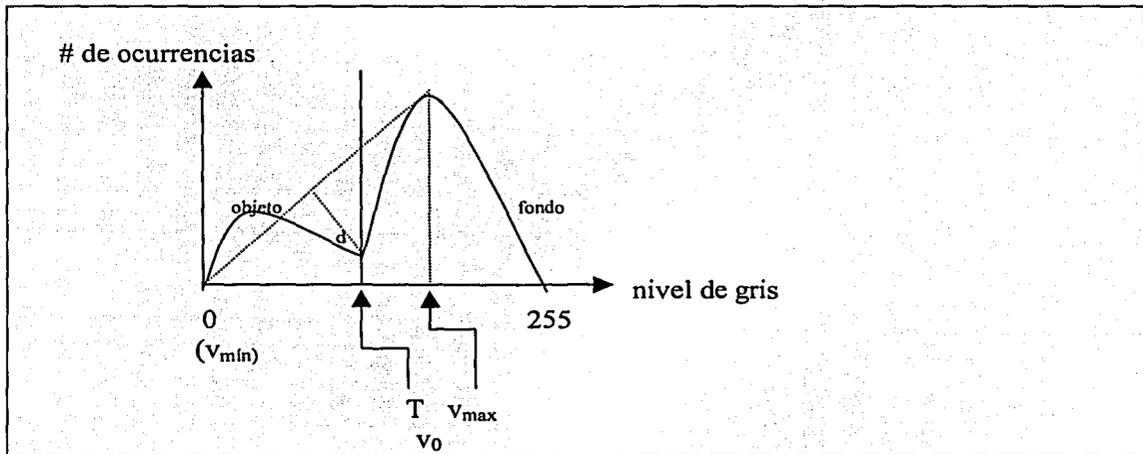


Figura 3.4. – Método del triángulo

Tenemos como v_{\min} el valor mínimo de nivel de gris, que en este caso es cero, tenemos como v_{\max} el valor del pico más alto. Entonces ya identificados estos valores, trazamos una línea desde el pico más alto hasta el nivel más bajo del histograma, posteriormente calculamos las distancias (d) de los puntos del histograma hasta la línea, encontrando el valor más grande, ahí trazamos el umbral.

3.4.1.4.7. – Limitaciones de los métodos basados en el histograma

Los métodos son válidos si los objetos tienen valores de intensidad constantes sobre toda la imagen pero, si la iluminación no es uniforme sobre toda la escena, puede suceder que un único umbral no sea suficiente para poder segmentar la imagen.

Otra de las limitaciones de estos métodos consiste en que el histograma nos da información de la distribución global de la intensidad de una imagen. Imágenes muy diferentes pueden tener diferentes distribuciones espaciales de niveles de gris, pero no tener histogramas muy similares.

3.4.2. – *Métodos de Detección de Bordes*

Se utilizan diferentes operadores que permiten encontrar líneas o bordes y a continuación se aplica algún tipo de post-procesamiento que permita la conexión de puntos para formar los bordes.

3.4.2.1. – Definición de borde

Un píxel de la imagen es de borde, si los píxeles que se encuentran alrededor de él en la imagen (sus vecinos) tienen niveles de gris muy diferentes a él.

Idealmente un borde separa dos objetos distintos de la imagen, pero en la realidad aparecen bordes que son causados por cambios de texturas, cambios de color, o por condiciones de iluminación durante la adquisición.

Una ilustración de los bordes se ilustra en la figura 3.5:

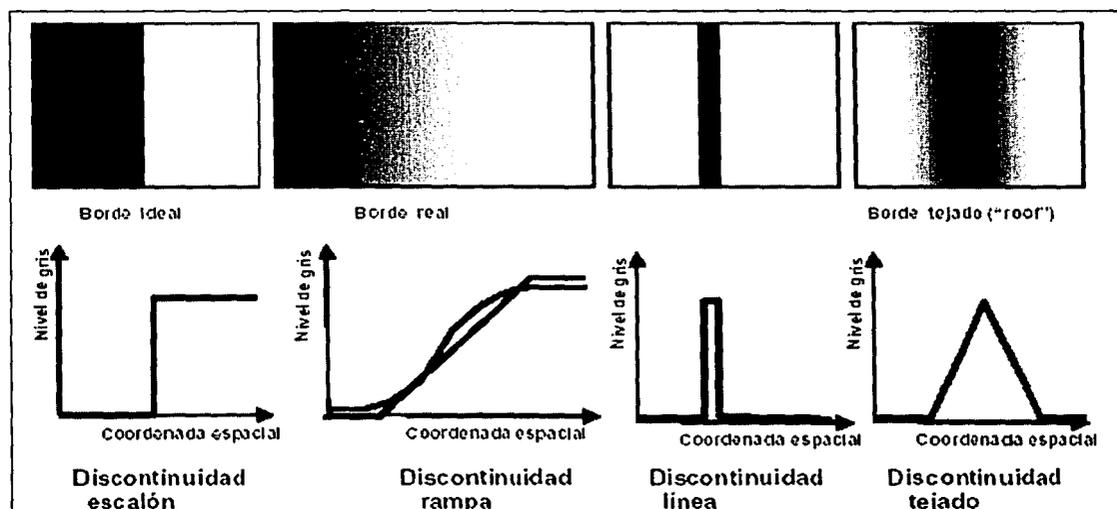


Figura 3.5. – Bordes y discontinuidades

Los cambios significativos de intensidades locales, en la imagen son asociados con discontinuidades.

3.4.2.2. – Tipos básicos de discontinuidades

En una imagen pueden aparecer tres tipos básicos de discontinuidades:

- Puntos.
- Líneas.
- Bordes.

3.4.2.2.1. – Detección de puntos

Se trata de detectar píxeles (puntos) en la imagen, que desentonan su nivel de gris respecto al nivel de gris de sus vecinos.

Para realizar esta técnica, es necesario utilizar una máscara². La máscara utilizada en este tipo de detección es:

-1	-1	-1
-1	<u>8</u>	-1
-1	-1	-1

Donde la respuesta a cada uno de los puntos de la imagen viene dada por:

$$R = I_1 - I_2 - I_3 - I_4 + 8I_5 - I_6 - I_7 - I_8 - I_9$$

donde I es el nivel de gris del píxel asociado con el coeficiente de máscara.

Entonces se tienen los siguientes casos:

- Cuando un punto es detectado y $|R| > T$, siendo T un umbral.
- Si $R=0$, la zona de estudio es una zona homogénea y tiene igual nivel de gris.

² Una máscara es un arreglo matricial, que ya tiene valores dados.

Para ejemplificar estos casos veamos los siguientes ejemplos:

$$\begin{array}{|c|c|c|} \hline 10 & 10 & 10 \\ \hline 10 & \underline{10} & 10 \\ \hline 10 & 10 & 10 \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline -1 & \underline{8} & -1 \\ \hline -1 & -1 & -1 \\ \hline \end{array} = 0$$

$$\begin{array}{|c|c|c|} \hline 13 & 10 & 18 \\ \hline 12 & \underline{20} & 10 \\ \hline 11 & 10 & 16 \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline -1 & \underline{8} & -1 \\ \hline -1 & -1 & -1 \\ \hline \end{array} = 60$$

$$\begin{array}{|c|c|c|} \hline 13 & 10 & 18 \\ \hline 12 & \underline{120} & 10 \\ \hline 11 & 10 & 16 \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline -1 & \underline{8} & -1 \\ \hline -1 & -1 & -1 \\ \hline \end{array} = 960$$

Como se nota en los ejemplos anteriores, un umbral válido para distinguir los puntos que realmente se diferencian de su entorno podría ser 500. La elección del umbral es un factor que depende de la aplicación que queremos realizar.

3.4.2.2.2. – Detección de líneas

El siguiente nivel de complejidad es la detección de líneas en la imagen.

Cada una de las siguientes máscaras tiene una respuesta mayor según el tipo de línea detectada en la imagen:

-1	-1	-1
2	<u>2</u>	2
-1	-1	-1

0°

-1	-1	2
-1	<u>2</u>	-1
2	-1	-1

45°

-1	2	-1
-1	<u>2</u>	-1
-1	2	-1

90°

2	-1	-1
-1	<u>2</u>	-1
-1	-1	2

-45°

Si para un cierto punto de la imagen $|R_i| > |R_j|$, entonces dicho punto está más relacionado con una línea en la dirección de la máscara i .

3.4.2.2.3. – Detección de bordes

La detección de bordes es la técnica más utilizada, ya que en la mayoría de las aplicaciones prácticas no es útil encontrar puntos y líneas. Los pasos para la detección de bordes son los siguientes:

- 1) **Filtrado.** Puesto que el cálculo del gradiente* basado en los valores de intensidad de un número finito de puntos es muy susceptible al ruido, normalmente se filtra la imagen inicialmente para mejorar el rendimiento del detector con respecto al ruido. Este proceso ha de aplicarse con cuidado, puesto que el emborronamiento oculta los bordes.
- 2) **Realce.** Este paso es esencial para determinar cambios en la intensidad de los vecinos de un punto. El proceso es básicamente el cálculo del gradiente.
- 3) **Detección.** Muchos puntos en una imagen tienen un gradiente no nulo. Obviamente no todos ellos son bordes para todas las aplicaciones. Es necesario algún método para determinar qué puntos son bordes. Un criterio ampliamente usado es la umbralización.
- 4) **Estimación.** Esta deberá llevarse a cabo en la resolución apropiada y en algunos casos será necesario también estimar la dirección del gradiente.

* El gradiente se analizará en la sección de operadores.

➤ Algunas definiciones respecto a la detección de bordes son las siguientes:

Punto de borde. - Es un punto de la imagen con coordenadas (x,y) localizado en un cambio significativo de intensidad de la imagen.

Detector de bordes. - Es un algoritmo que produce un conjunto de bordes (o puntos de borde) a partir de una imagen.

Contorno. - Es una cadena de bordes o la curva matemática que permite modelar la cadena de bordes.

Unión de bordes. - Es el proceso de formar una sola cadena ordenada de bordes a partir de una cadena no ordenada.

Seguimiento de bordes. - Es el proceso de buscar en la imagen para determinar contornos.

3.4.2.3. – Métodos de Segmentación por Detección de Bordes

Las técnicas de segmentación basadas en bordes, se basan en los bordes encontrados en la imagen por los detectores de bordes. Otros pasos posteriores se siguen para combinar los bordes en cadenas de bordes que corresponden a fronteras de objetos en la imagen.

Este método se basa en realizar la búsqueda del valor máximo del gradiente, sobre cada línea que forma la imagen. Cuando un máximo es encontrado, el método trata de seguir el máximo del gradiente alrededor del objeto, hasta encontrar de nuevo el punto inicial, para luego buscar el próximo máximo en el gradiente. El objetivo final de estos métodos es alcanzar al menos una segmentación parcial.

Empezaremos por ver los operadores y posteriormente pasaremos a ver los métodos.

3.4.2.3.1. – Operador Gradiente

En el análisis de los operadores hemos de hacer aproximaciones de las derivadas, puesto que lo que se tiene bajo análisis son aproximaciones, no una imagen real.

Los métodos utilizados para generar detectores diferenciales involucran el cálculo del gradiente en dos direcciones sobre la imagen.

El gradiente de una imagen $f(x, y)$ en un punto (x, y) viene dado por el vector:

$$\vec{\nabla} f = \vec{G}[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

La magnitud de dicho vector viene dada por:

$$|G[f(x, y)]| = \sqrt{G_x^2 + G_y^2} \equiv |G_x| + |G_y|$$

Mientras que la dirección del gradiente la da la siguiente ecuación:

$$\alpha(x, y) = \tan^{-1} \left(\frac{G_y}{G_x} \right)$$

donde el ángulo es medido respecto al eje X.

La aproximación más sencilla para calcular el gradiente es realizando el cálculo de las diferencias siguientes:

$$\begin{aligned} G_x &\cong f[i, j+1] - f[i, j] \\ G_y &\cong f[i, j] - f[i+1, j] \end{aligned}$$

Esta aproximación se da utilizando las siguientes máscaras:

$$G_x = \begin{bmatrix} 1 & -1 \end{bmatrix}$$

$$G_y = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Otra alternativa para obtener el gradiente es utilizando máscaras de 3X3, con esto se genera el gradiente en el píxel central.

3.4.2.3.2. – Operador de Roberts

El operador de Roberts nos da información sobre la existencia de un punto de borde, además de darnos información sobre la orientación. El operador es el siguiente:

$$G[f[x, y]] = |f[x, y] - f[x+1, y+1]| + |f[x+1, y] - f[x, y+1]|$$

Donde G_x y G_y son calculados usando las siguientes máscaras:

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$G_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

3.4.2.3.3. – Operador de Sobel

Una forma de evitar que el gradiente se calcule en un punto intermedio, tal y como ocurre en el operador de Roberts, es usar ventanas de tamaño 3x3. Consideremos una distribución de píxeles que se muestren alrededor de (x, y) en una ventana de este tamaño. El operador de Sobel es la magnitud del gradiente calculado mediante

$$M = \sqrt{s_x^2 + s_y^2}$$

Donde las derivadas parciales se calculan mediante las siguientes expresiones:

$$s_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6)$$

$$s_y = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4)$$

Donde el etiquetado de niveles de gris es el siguiente:

a_8	a_1	a_2
a_7	(i,j)	a_3
a_6	a_5	a_4

y con la constante $c = 2$.

Estos operadores de gradiente pueden implementarse utilizando máscaras. Sus máscaras son:

G_x

-1	0	1
-2	0	2
-1	0	1

G_y

1	2	1
0	0	0
-1	-2	-1

3.4.2.3.4. – Operador de Prewitt

Este operador es igual al de Sobel, solamente que la diferencia es en la constante c , que en este caso es igual a 1. Las máscaras se definen de la siguiente manera:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

3.4.2.3.5. – Operador de Kirsch (máscaras de compás)

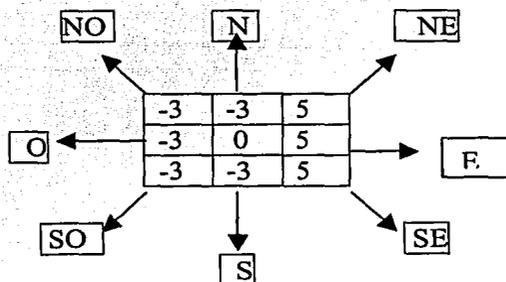
Este operador utiliza máscaras que, se llaman de compás porque se obtienen a partir de una única máscara y girándola en las ocho orientaciones principales: Este, Noreste, Noroeste, Norte, Oeste, Suroeste, Sur y Sureste.

Sea la máscara k_0 , la máscara original:

k_0

-3	-3	5
-3	0	5
-3	-3	5

Entonces tenemos los siguientes giros en ella:



Obteniendo k_1 , giramos la máscara original hacia el Este (el píxel que apunta hacia el SE se mueve ahora hacia el Este) y tenemos:

k_1

-3	5	5
-3	0	5
-3	-3	-3

Obteniendo k_2 , giramos la máscara k_1 hacia el Noreste (el píxel que apunta hacia el Este se mueve ahora hacia el Noreste) y tenemos:

k_2

5	5	5
-3	0	-3
-3	-3	-3

Y sucesivamente se van obteniendo todas las ocho máscaras. Una vez obtenidas, la magnitud del borde se encuentra buscando el máximo valor en la convolución³ de cada una de estas máscaras sobre nuestra ventana de análisis. La dirección del borde es definida por la máscara que produce la máxima magnitud.

3.4.2.3.6. – Operador de Robinson (máscaras de compás)

Este operador se usa de la misma forma que el de Kirsch, pero son más rápidas de calcular porque los coeficientes son 0,1 y 2.

Sólo es necesario calcular el resultado de las cuatro primeras máscaras, ya que las cuatro últimas se obtienen por negación de las primeras.

Sea r_0 la siguiente máscara:

r_0

-1	0	1
-2	0	2
-1	0	1

³ La convolución es en términos prácticos, poner la máscara sobre una ventana de la imagen del mismo tamaño, y realizar una multiplicación punto a punto e ir sumando los valores de la multiplicación, y el valor encontrado es el correspondiente al píxel central de la ventana utilizada.

Realizando la misma metodología de giro vista en el operador de Kirsch, tenemos:

$$r_2 = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}$$

$$r_3 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$r_4 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix}$$

Las siguientes cuatro máscaras se calculan por negación de las ya obtenidas:

$$r_5 = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

$$r_6 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

$$r_7 = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$

$$r_8 = \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Una vez obtenidas las ocho máscaras, la magnitud del borde se define como el máximo valor encontrado por la convolución de cada una de las máscaras con la imagen.

3.4.2.3.7. – Operador Laplaciano

Este operador como su nombre lo indica, utiliza la segunda derivada en dos dimensiones, siendo la segunda derivada de un borde escalón una función que cruza por cero en la posición del borde, tal como se ilustra en la figura 3.6:

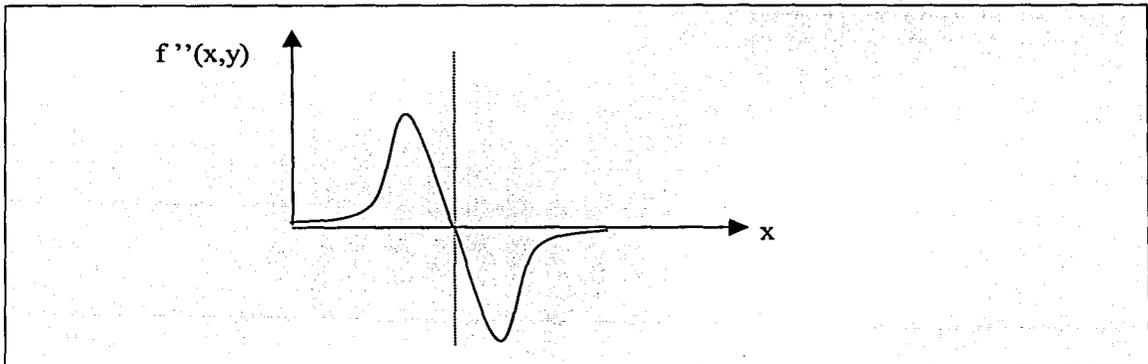


Figura 3.6. – Gráfica de la segunda derivada

El Laplaciano está definido de la siguiente manera:

$$\nabla^2 f = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} = \frac{\partial G_x}{\partial x} + \frac{\partial G_y}{\partial y}$$

La máscara a utilizar para calcular el Laplaciano es:

$$\nabla^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Para poder comprender aún mejor el Laplaciano veamos el siguiente ejemplo:
Sea la siguiente ventana, la porción de una imagen dada

2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8
2	2	2	2	2	8	8	8	8	8

Si se aplica sobre ella el Laplaciano, se obtiene el siguiente resultado:

0	0	0	6	-6	0	0	0
0	0	0	6	-6	0	0	0
0	0	0	6	-6	0	0	0
0	0	0	6	-6	0	0	0

El signo del resultado (positivo o negativo) produce información direccional y nos indica cual es el lado del borde más claro.

3.4.2.3.8. – Operador Laplaciano del Gaussiano

Este operador combina un filtro Gaussiano con un Laplaciano. La máscara de este operador viene dada por:

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	2	-1	0
0	0	-1	0	0

3.4.2.3.9. – Método de Segmentación del Watershed

El método de segmentación del Watershed se basa en la simulación de un proceso de inundación, partiendo de la imagen sometida a un proceso de obtención del gradiente.

Se realiza una inundación progresiva desde los niveles de gris más bajos, partiendo de los mínimos de la imagen del gradiente. Cuando se produzca un choque del agua proveniente de dos mínimos diferentes, obtendremos una línea de watershed, que se mantendrá a manera de dique, hasta que la inundación llegue a los niveles más altos. Esto se ilustra en la figura 3.7:

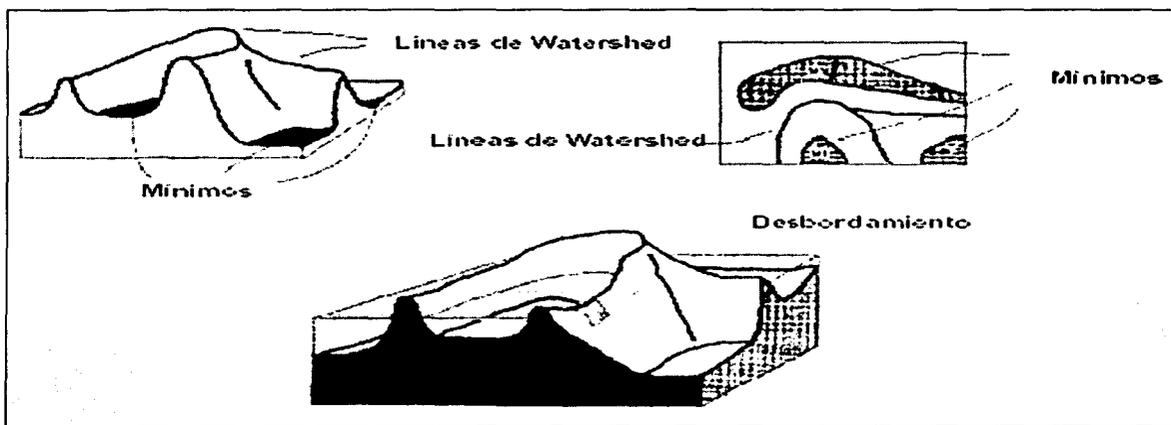


Figura 3.7. – Método de Watershed

Los diques que se hayan obtenido serán las líneas de Watershed, quienes a su vez conformarán los contornos cerrados de los objetos presentes en la imagen. La figura 3.8 muestra una manera gráfica de ver este método.



Figura 3.8. – Líneas de Watershed*

El método de Watershed permite obtener siempre contornos cerrados de grosor de 1 píxel (o como mucho dos píxeles).

En la práctica, el uso de este método produce una sobreesegmentación (Figura 3.9), ya que las zonas homogéneas (con mínimos diferentes), en cuanto a niveles de gris, no suelen coincidir con los objetos que deseamos segmentar:

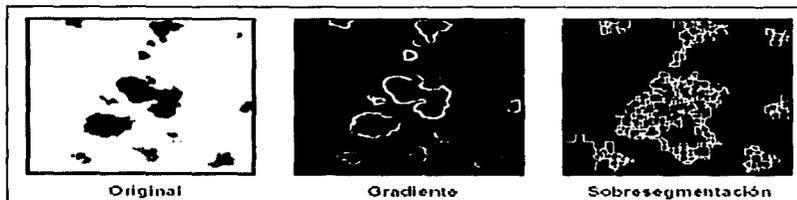


Figura 3.9. – Sobreesegmentación*

Una de las posibles soluciones a este problema es partir de una imagen gradiente, a la que se han añadido una serie de “marcas” que se convertirán en nuevos mínimos esto se muestra en la figura 3.10:

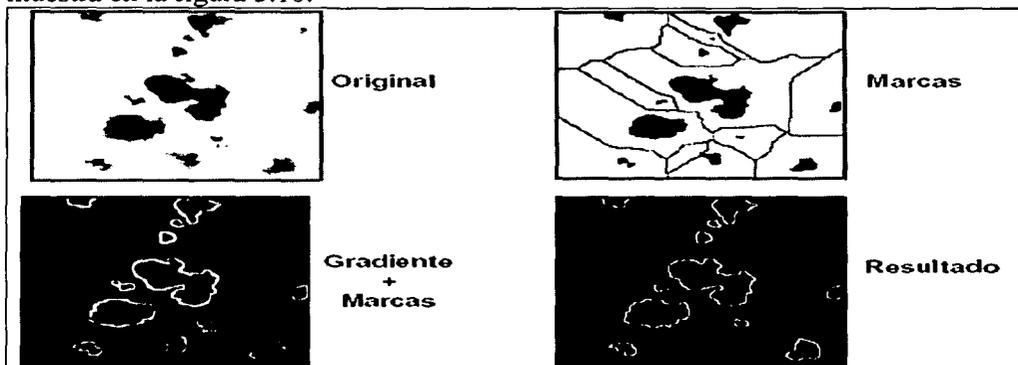


Figura 3.10. – Marcas de Watershed*

* Apuntes de Visión por Computador de la Universidad Politécnica de Valencia

3.4.2.3.10. –Método de Segmentación basado en la Transformada de Hough (Unión de Bordos y Detección de Contornos)

El método realiza una búsqueda de puntos de borde que pueden formar una línea recta, para lo cual se consideran dos soluciones:

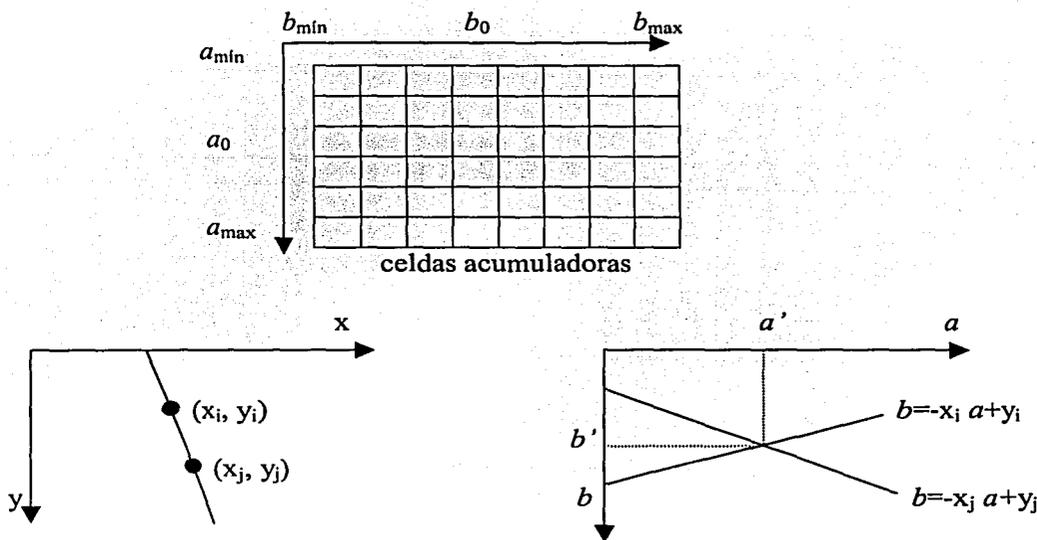
1. - Encontrar todas las posibles líneas rectas determinadas por cada par de puntos de borde. A continuación encontrar todos los subconjuntos de puntos de borde (x,y) que pertenecen a una línea recta en particular.

Computacionalmente esto es muy costoso, ya que implica encontrar aproximadamente n^2 rectas o líneas y realizar n^3 comparaciones de cada punto para todas las líneas.

2. - Transformada de Hough, que es menos costosa que la solución 1.

Esta técnica se basa en considerar relaciones globales entre píxeles y, para llevar a cabo la unión de puntos de borde previamente se analiza si se encuentran en una misma línea (calcular el gradiente de la imagen). Después de esto, se realiza lo siguiente:

Dado un punto (x, y) un número infinito de rectas pueden pasar a través de él, teniendo en cuenta la ecuación general de la recta $y = ax + b$ y variando los valores de a y b (a es la pendiente y b el origen). Entonces Hough transforma el espacio de parámetros y utiliza las llamadas celdas acumuladoras, que representa el espacio de parámetros a b . Cada celda no representa un valor único de a b , sino un intervalo de valores:



Todos los puntos contenidos en la línea que una los puntos (x_i, y_i) e (x_j, y_j) tienen líneas en el espacio de parámetros que intersecta (a', b') .

El método es el siguiente:

Inicialmente las celdas acumuladoras valen cero y, para cada punto (x_k, y_k) en el plano de la imagen (x, y) , teniendo en cuenta los posibles valores de a , se resuelve la ecuación $b = y_k - x_k a$.

El resultado de b es redondeado al valor más cercano en el eje b y se incrementa en uno el contenido de la celda correspondiente.

Al final del procedimiento un valor M en una determinada celda corresponde a que m puntos en el plano $x - y$ caen en la línea $y = a_i x + b_i$.

La división del eje a en K incrementos dados, para cada punto (x_k, y_k) , k valores de b se corresponden con k valores de a .

Pero existe un inconveniente utilizando la ecuación general de la recta, y es que puede ser muy larga e interceptar en el infinito. Para solucionar esto hay que utilizar la siguiente representación de la recta:

$$x \cos \theta + y \sin \theta = \rho$$

donde ρ es la distancia y su rango va de -90° a $+90^\circ$, mientras que θ es la orientación, la cual no puede ser mayor que el tamaño de la diagonal de la imagen, que es la mayor distancia en la que podemos localizar una recta.

La transformada de Hough de una recta es un punto en el plano (ρ, θ) .

Los pasos del método se resumen de la siguiente manera:

- Calcular el gradiente de la imagen.
- Cada celda del espacio de parámetros es un acumulador, que se ha de inicializar a cero.
- Especificar la subdivisión en el plano $\rho\theta$.
- Para cada punto (x, y) de borde de la imagen, incrementar en uno el acumulador que satisface la ecuación.
- Examinar los contadores de las celdas acumuladoras y seleccionar las de valor más alto.
- Examinar la relación, para la continuidad, entre píxel en una celda dada.

CAPÍTULO 4

ANTECEDENTES DE LOS MRF

4.1. – PROBABILIDAD Y VARIABLES ALEATORIAS

4.1.1. – Probabilidad

El estudio formal de la teoría de probabilidad se originó en el siglo XVII y XVIII en Francia, y fue motivado por el estudio de los juegos de azar, sin embargo este punto de vista empezó a cambiar en el siglo XVIII, cuando se desarrolló un modelo probabilístico para el comportamiento de las moléculas de un líquido⁴. Posteriormente el interés sobre la probabilidad fue en aumento.

A pesar que las aplicaciones se fueron extendiendo con rapidez en los inicios del siglo XX, fue hasta los años 1932 a 1934 que apareció una estructura matemática rigurosa para la probabilidad.

4.1.1.1. – Definición de probabilidad

La probabilidad es la posibilidad de que se presente un evento⁵ resultante de un experimento estadístico, evaluado por medio de un conjunto de números reales que caen en el rango de 0 a 1.

A cada punto en el espacio muestral⁶ se le asigna una probabilidad tal que la suma de todas las probabilidades es 1.

Para encontrar la probabilidad de un evento A, se suman todas las probabilidades asignadas a los puntos muestrales en A. Esta suma se llama la **probabilidad de A** y es expresada por P(A.)

4.1.1.2. – Medidas numéricas

➤ Media:

La media o promedio aritmético de un conjunto de números x_1, x_2, \dots, x_n se encuentra sumando los números y dividiendo después la suma entre n, el número de medidas. La fórmula viene dada por:

$$\bar{x} = \sum_{i=1}^n x_i / n$$

⁴ Conocido como movimiento Browniano, ya que Robert Brown, botánico inglés observó este fenómeno en 1827

⁵ Un evento es un subconjunto del espacio muestral del experimento

⁶ El espacio muestral es el conjunto total de resultados posibles

➤ **Varianza:**

La varianza es una medida de la variabilidad debido a que si muchas de las diferencias son grandes (o pequeñas), entonces el valor de la varianza s^2 será grande (o pequeño.) El valor de la varianza puede sufrir un cambio muy desproporcionado, aún más que la media, por la existencia de algunos valores extremos en el conjunto de análisis.

La varianza de las observaciones x_1, x_2, \dots, x_n es el promedio del cuadrado de las distancias entre cada observación y la media del conjunto de observaciones.

La varianza se denota por:

$$S^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)$$

➤ **Desviación estándar:**

La desviación estándar es la raíz cuadrada positiva de la varianza, y se denota por:

$$S = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}$$

➤ **Mediana:**

La mediana de un conjunto de observaciones es el valor medio, cuando todas las observaciones se ordenan de manera creciente. Para determinar la mediana hay que seguir los siguientes pasos:

1. Ordenar los datos de menor a mayor.
2. Si el número de medidas es impar, entonces la mediana será la medida en el centro, pero si el número de medidas es par, la mediana es la media de las dos medidas que ocupan las posiciones centrales.

➤ **Moda:**

La moda, si existe, es el valor más frecuente. La moda muestra hacia qué valor tienden los datos a agruparse. Esta medida tiene dos ventajas:

1. Para ciertas muestras pequeñas, se le determina fácilmente y, en general, no se ve afectada por los valores extremos al final de un conjunto de datos ordenados.
2. La moda puede usarse como una medida de tendencia central para datos numéricos empleados en sentido cualitativo.

➤ **Rango medio:**

El rango medio, es el promedio aritmético de las medidas mayor y menor.

4.1.2. – Variables aleatorias

4.1.2.1. – Concepto de variable aleatoria

Una variable aleatoria es una función que asocia un número real a cada elemento del espacio muestral.

Se utiliza una letra mayúscula, por ejemplo una X, para designar una variable aleatoria y su correspondiente letra minúscula, x para uno de sus valores. Entonces, a cada punto del espacio muestral se le asigna un valor numérico de 0,1,2,...,n. Estos valores son cantidades aleatorias determinadas por el resultado del experimento. Éstos pueden considerarse como los valores que asume la variable aleatoria X. Además, diferentes valores de X pueden conducir a una misma x. (Figura 4.1)

Para comprender un poco más este concepto, analizaremos el siguiente ejemplo:

Se tienen tres eventos S, J y B en un espacio muestral. Suponiendo que el orden correcto de los eventos es SJB, entonces los posibles arreglos en los cuales se podrían organizar los tres eventos y el número de asociaciones correctas (X) es:

Espacio muestral	x (valores que toma la variable aleatoria)
SJB	3
SBJ	1
JSB	1
JBS	0
BSJ	0
BJS	1

En este ejemplo se nota que cuando un evento coincide con el orden correcto, el valor que toma es de 1, en el caso que sean más de un evento los que coincidan con el orden correcto, se van sumando sus valores, por eso en el primer caso en que todos los eventos coinciden con el orden correcto, su valor es de 3.

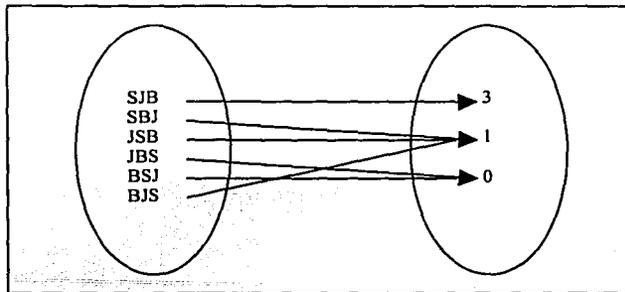


Figura 4.1. – Valores de x

4.1.2.2. - Variables aleatorias discretas.

Cuando un espacio muestral contiene un número finito de posibilidades o una secuencia interminable con tantos elementos como números naturales existen, se le llama **espacio muestral discreto**. A su vez, una **variable aleatoria discreta** es aquella en donde se puede contar su conjunto de resultados posibles (ya sean sus valores finitos o infinitos), y si éstos pueden arreglarse en una secuencia que corresponda a los enteros positivos.

Una variable aleatoria discreta asume cada uno de sus valores con una cierta probabilidad. Tomando los datos obtenidos en el ejemplo anterior, tenemos las siguientes probabilidades:

x	0	1	3
P(X=x)	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

4.1.2.2.1. – Función de probabilidad de una variable aleatoria discreta.

Con mucha frecuencia es conveniente representar con una fórmula todas las probabilidades de una variable aleatoria X. Esta fórmula necesariamente debe ser una función de los valores numéricos x, y recibe por nombre **función de probabilidad** o **distribución de probabilidad** de la variable aleatoria discreta X, y está formada por el conjunto de pares ordenados (x, f(x)); esto es, $f(x) = P(X = x)$.

ESTADÍSTICA
DE LA

La fórmula viene dada de la siguiente manera:

$$f(x) = P(X = x) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}} \quad x = 0, 1, \dots, n$$

donde:

D es el número de eventos que se quiere registrar.

n son los valores que toma la variable aleatoria discreta x.

N es el número de elementos del espacio muestral.

Como un ejemplo, consideremos un embarque de 8 computadoras similares, de las cuales hay 3 defectuosas. Si una escuela realiza una compra aleatoria de 2 de estas computadoras. Queremos encontrar la distribución de probabilidad para el número de computadoras defectuosas entonces tenemos lo siguiente:

$$D = 3$$

$$N = 8$$

$$n = 2$$

$$f(0) = P(X=0) = \frac{\binom{3}{0} \binom{8-3}{2-0}}{\binom{8}{2}} = \frac{10}{28}$$

$$f(1) = P(X=1) = \frac{\binom{3}{1} \binom{5}{1}}{\binom{8}{2}} = \frac{15}{28}$$

$$f(2) = P(X=2) = \frac{\binom{3}{2} \binom{5}{0}}{\binom{8}{2}} = \frac{3}{28}$$

Entonces la distribución de probabilidad de X es:

x	0	1	2
$f(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

4.1.2.2.2. - Distribución acumulada de una variable aleatoria discreta.

La distribución acumulada $F(x)$ de una variable aleatoria discreta X, cuya distribución de probabilidad es $f(x)$, es:

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \quad \text{para } -\infty < x < \infty$$

Muchas veces, existen problemas en los cuales se desea calcular la probabilidad de que el valor observado de una variable aleatoria X sea menor o igual a un número real x. Si se escribe $F(x) = P(X \leq x)$ para cada número real x.

Como ejemplo consideremos el visto anteriormente, el del lote de ocho computadoras. Las distribuciones de probabilidad son las siguientes:

$$f(0) = \frac{10}{28}, \quad f(1) = \frac{15}{28}, \quad f(2) = \frac{3}{28}$$

encontrando la distribución acumulada tenemos:

$$F(0) = f(0) = \frac{10}{28}$$

$$F(1) = f(0) + f(1) = \frac{10}{28} + \frac{15}{28} = \frac{25}{28}$$

$$F(2) = f(0) + f(1) + f(2) = \frac{10}{28} + \frac{15}{28} + \frac{3}{28} = \frac{28}{28} = 1$$

Entonces,

$$F(x) = \begin{cases} \frac{10}{28} & \text{para } x < 0 \\ \frac{15}{28} & \text{para } 0 \leq x < 1 \\ 1 & \text{para } x \geq 1 \end{cases}$$

4.1.2.2.3. – Valor esperado (media) de una variable aleatoria discreta

El valor esperado de una variable aleatoria discreta X es el promedio o valor medio de X y está dado por:

$$E(X) = \sum_x xf(x)$$

Donde $f(x)$ es la función de probabilidad.

4.1.2.2.4. – Varianza de una variable aleatoria discreta

La varianza de una variable aleatoria discreta viene dada por la siguiente fórmula:

$$\sigma^2 = \sum_x [(x - \mu)^2 f(x)]$$

donde μ es el valor esperado o media ($E(X)$) y $f(x)$ es la probabilidad de ocurrencia de x .

4.1.2.2.5. – Desviación estándar de una variable aleatoria discreta.

La desviación estándar de una variable aleatoria discreta se da de la siguiente manera:

$$\sigma = \sqrt{\text{varianza}}$$

4.1.2.3. – Variabes aleatorias continuas

Una variable aleatoria continua nunca asume cualquiera de sus valores exactamente, ya que como su nombre lo indica, toma valores desde $-\infty$ hasta ∞ , y algunos de sus valores no son enteros. Es decir, la variable aleatoria es continua si sus valores consisten en uno o más intervalos de la recta de los reales.

Entonces ahora se utiliza el cálculo de probabilidades para varios intervalos de variables aleatorias continuas, tal como $P(a < X < b)$.

4.1.2.3.1. - Distribución de probabilidad de una variable aleatoria continua

La distribución de probabilidad de una variable aleatoria continua no puede representarse de forma tabular, pero sí puede tener una fórmula. Dicha fórmula, necesariamente debe ser una función de los valores numéricos de la variable continua X y se expresa por la notación funcional $f(x)$ y, por lo general, ya no se llama distribución de probabilidad, ahora se llama **función de densidad de probabilidad** o **función de densidad**.

Definición. – La función $f(x)$ es una función de densidad de probabilidad para la variable aleatoria continua X , definida en el conjunto de los números reales, si cumple con las siguientes especificaciones:

1. $f(x) \geq 0$ para toda $x \in \mathbb{R}$.
2. $\int_{-\infty}^{\infty} f(x) dx = 1$.
3. $P(a < X < b) = \int_a^b f(x) dx$

La mayoría de las funciones de densidad son continuas y sus gráficas pueden tener diversas formas (Ver figura 4.2). Debido a que las áreas se utilizan para representar las probabilidades y éstas tienen valores numéricos positivos, la función de densidad debe estar totalmente arriba del eje x , y se construye de tal manera que el área comprendida bajo la curva es igual a 1, cuando se calcula para el intervalo definido para la función $f(x)$. En la figura 4.3, la probabilidad de que x asuma un valor entre a y b es igual al área delimitada por a y b bajo la función de densidad.

El área está dada por:

$$P(a < X < b) = \int_a^b f(x) dx$$

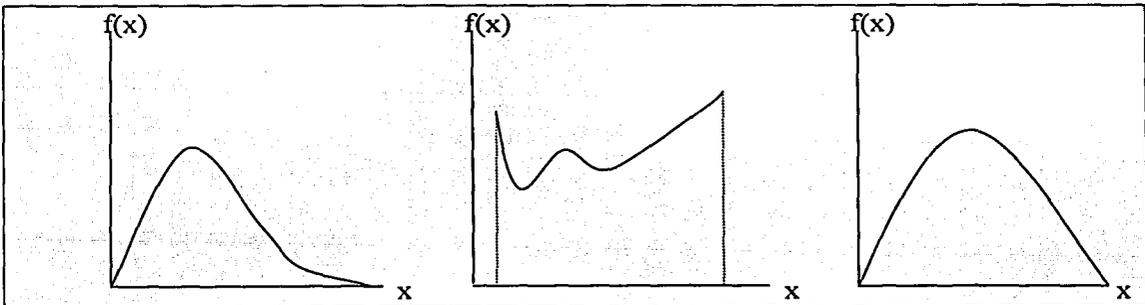


Figura 4.2. – Algunas formas de las funciones de densidad

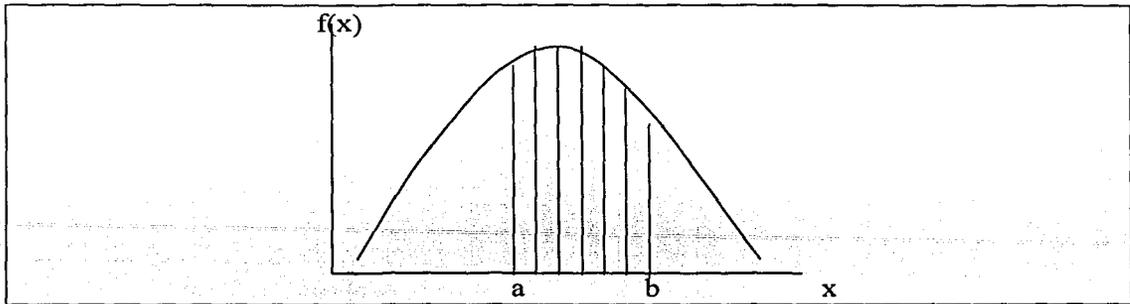


Figura 4.3. - $P(a < X < b)$

4.1.2.3.2. - Distribución acumulada de una variable aleatoria continua.

La distribución acumulada $F(x)$ de una variable aleatoria continua X con una función de densidad $f(x)$ es:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt \quad \text{para } -\infty < x < \infty$$

Y como consecuencia inmediata de esta definición se puede escribir:

$$P(a < X < b) = F(b) - F(a)$$

Veamos el siguiente ejemplo:

$$\text{Sea } f(x) = \begin{cases} \frac{x^2}{3} & -1 < x < 2 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Encontremos la función de densidad $F(x)$ y evaluemos con ella $P(0 < X \leq 1)$

$$\text{Tenemos entonces que } F(x) = \int_{-\infty}^x f(t) dt = \int_{-1}^x \frac{t^2}{3} dt = \frac{t^3}{9} \Big|_{-1}^x = \frac{x^3 + 1}{9}$$

Por lo tanto,

$$F(x) = \begin{cases} 0, & x \leq -1 \\ \frac{x^3 + 1}{9}, & -1 < x < 2 \\ 1, & x \geq 2 \end{cases}$$

$$\text{Entonces } P(0 > X \leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

4.1.2.3.3. – Valor esperado (media) de una variable aleatoria continua.

El valor esperado de una variable aleatoria continua X es el promedio o valor medio de X y está dado por:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Donde $f(x)$ es la función de densidad.

4.1.2.3.4. – Varianza de una variable aleatoria continua

La varianza de una variable aleatoria discreta viene dada por la siguiente fórmula:

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx$$

donde μ es el valor esperado o media ($E(X)$) y $f(x)$ es la probabilidad de ocurrencia de x .

4.1.2.3.5. – Desviación estándar de una variable aleatoria continua.

La desviación estándar de una variable aleatoria discreta se da de la siguiente manera:

$$\sigma = \sqrt{\text{varianza}}$$

4.1.2.4. - Momentos de una variable aleatoria

Los momentos de una variable aleatoria X son los valores esperados de ciertas funciones de X . Éstos forman una colección de medidas descriptivas que pueden emplearse para caracterizar la distribución de probabilidades de X y especificarla si todos los momentos de X son conocidos.

A pesar de que los momentos de X pueden definirse alrededor de cualquier punto de referencia, generalmente se definen alrededor del cero o del valor esperado de X .

Definición. – Sea X una variable aleatoria. El r -ésimo momento de X alrededor del cero se define como:

$$\mu_r' = E(X^r) = \sum_x x^r f(x) = \int_{-\infty}^{\infty} x^r f(x) dx$$

Definición. – Sea X una variable aleatoria. El r -ésimo momento de X alrededor de la media de X se define por:

$$\mu_r = E(X - \mu)^r = \sum_x (x - \mu)^r f(x) = \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx$$

- **Momento central:** El momento central de una variable aleatoria X es 1.
- **Primer momento central:** El primer momento central de una variable aleatoria X es 0.
- **Segundo momento central:** Recibe el nombre de varianza.
- **Tercer momento central:** Este momento nos dan información acerca de la forma de la función de probabilidad de X :

Si	La función es	Gráfica
$\mu_3 < 0$	Asimétrica negativamente	
$\mu_3 > 0$	Asimétrica positivamente	
$\mu_3 = 0$	Simétrica	

Tabla 4.1. –Distribuciones de probabilidad típicas del 3er. momento

➤ **Cuarto momento central:** Este parámetro nos indica qué tan puntiaguda es la función de probabilidad y recibe el nombre de **curtosis**.

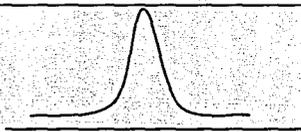
Si	La función es	Gráfica
$\mu_4 > 3$	Con pico alto (Leptocúrtica)	
$\mu_4 < 3$	Plana (Platicúrtica)	
$\mu_4 = 3$	Ni alta ni baja (Mesocúrtica)	

Tabla 4.2. –Distribuciones de probabilidad típicas del 4º. momento

4.2. – TEORÍA BAYESIANA DE PROBABILIDAD

4.2.1. – *Probabilidad condicional*

A la probabilidad de que un evento B se dé cuando se sabe que algún otro evento A se ha presentado se llama **probabilidad condicional** y se denota con la siguiente expresión:

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

O también se denota como:

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

Esto significa que la probabilidad de que tanto A como B ocurran es igual a la probabilidad de que A ocurra multiplicada por la probabilidad condicional de que B ocurra, dado que A ya ocurrió.

Como un ejemplo, supongamos que el espacio muestral S es la población de adultos que han satisfecho los requisitos para graduarse en la escuela. Se deben clasificar de acuerdo con su sexo y a si trabajan o no actualmente:

	Empleado	Desempleado	Total
Hombre	460	40	500
Mujer	140	260	400
Total	600	300	900

Entonces:

M: se escoge a un hombre,
E: tenga empleo.

Se encuentra que:

$$P(M|E) = \frac{P(E \cap M)}{P(E)} = \frac{460/900}{600/900} = 23/30$$

4.2.2. – Eventos independientes y eventos mutuamente excluyentes.

- **Eventos independientes.** - El que un evento B suceda, no tiene impacto en que se presente el evento A. Aquí la ocurrencia de A es independiente de la ocurrencia de B. Entonces:

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

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

y

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

- **Eventos mutuamente excluyentes.** - Si los eventos A y B no tienen elementos en común, es decir no pueden ocurrir simultáneamente. Esto es:

Dos eventos A y B son mutuamente excluyentes o disjuntos, si $A \cap B = \emptyset$ (Figura 4.4)



Figura 4.4. – Eventos mutuamente excluyentes

4.2.3 – Teorema de Bayes

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

$$P(B_r | A) = \frac{P(B_r \cap A)}{\sum_{i=1}^k P(B_i \cap A)} = \frac{P(B_r)P(A | B_r)}{\sum_{i=1}^k P(B_i)P(A | B_i)} \quad \text{para } r = 1, 2, \dots, k$$

El teorema de Bayes se usa comúnmente como un medio de corregir nuestras afirmaciones sobre probabilidades, en base a nuevos datos.

4.2.4. - Probabilidad total

Si A y B forman un sistema completo de sucesos, la probabilidad de cualquier otro suceso M es:

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

Esto quiere decir que A_1, A_2, \dots, A_n son eventos mutuamente excluyentes, y la suma de sus probabilidades son igual a 1.

4.2.5. - Distribución condicional.

Dado un evento C con probabilidad diferente de cero, la distribución condicional de una variable aleatoria X viene dada por:

$$F_x(x|C) = P\{X \leq x|C\} = \frac{P\{X \leq x, C\}}{P(C)}$$

La densidad condicional se obtiene a partir de la distribución en el caso clásico.

4.2.6. - Valor esperado condicional.

El valor esperado condicional de una variable aleatoria X supone que el valor de C está dado por:

$$E(X|C) = \left\{ \begin{array}{l} \int_{-\infty}^{\infty} xf(x|C)dx \\ \sum_n x_n P(X = x_n|C) \end{array} \right\}$$

Para el caso continuo y discreto, respectivamente.

4.3. – ANTECEDENTES DE LOS MRF

Desde su comienzo en los años 60, la investigación en visión artificial ha ido evolucionando en búsqueda de soluciones para esta área. En la búsqueda de soluciones, los investigadores se han dado cuenta de la importancia de la información contextual para la comprensión de las imágenes. Las restricciones de contexto son, en última instancia, necesarias en la interpretación de la información visual.

El uso de restricciones contextuales es muy importante para un sistema de visión. En este proceso se han propuesto una gran cantidad de modelos de visión usando el contexto, y entre ellos está el modelo de Campos Aleatorios de Markov.

La teoría de los Campos Aleatorios de Markov proporciona una manera conveniente y constante de modelar entidades dependientes del contexto tales como píxeles de la imagen y otras características especiales correlacionadas.

Las razones para usar los Campos Aleatorios de Markov son las siguientes:

- Hace más fácil el cálculo de medidas de ejecución cuantitativas para caracterizar cómo trabajan los algoritmos de análisis de imagen.
- Los modelos CAM pueden ser usados para incorporar información contextual a priori, o bien restricciones, de manera cuantitativa.

Los modelos de visión basados en Campos Aleatorios de Markov están formulados dentro del campo de trabajo Bayesiano. La solución óptima de un problema se define como el estimador de probabilidad máxima a posteriori (MAP), el mejor que se puede obtener a partir de las observaciones aleatorias. Los resultados de la teoría CAM nos proporcionan herramientas para codificar restricciones contextuales en la probabilidad a priori. Esta es la razón principal para el uso de CAM en modelos de visión.

Los estudios basados en CAM han tenido éxito al modelar problemas de visión de bajo nivel, como son restauración de imágenes, segmentación, reconstrucción de superficies, análisis de texturas, flujo óptico, integración visual y detección de bordes.

Más recientemente (1989) ha comenzado la investigación y aplicación de los modelos CAM en visión de alto nivel, como para emparejamiento y reconocimiento de objetos. En un desarrollo inicial de un modelo CAM para la interpretación de imágenes, la solución óptima se define como el etiquetado MAP, el cual se calcula minimizando la energía a posteriori, usando la regla de Bayes, a partir de un modelo a priori y un modelo de probabilidad.

Como ejemplo, supongamos un sistema de visión cuya finalidad es clasificar dos tipos de objetos o piezas, tuercas y tornillos. Entonces la clase B_1 serán las tuercas y la clase B_2 serán los tornillos; y eligiendo un vector de características A tenemos que: Entonces puede expresarse el teorema de Bayes de la siguiente manera:

$$P(B_r | A) = \frac{P(X|B_r)P(B_r)}{P(A)} \quad \text{para } r = 1, 2, \dots, k$$

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

- $P(B_r | A)$ Probabilidad de que el vector de características A pertenezca a la clase B_r .
- $P(A | B_r)$ Probabilidad de que dada la clase B_r , el valor de la variable aleatoria sea, precisamente, A . Esto es la función de distribución de la clase B_r , considerada como una variable aleatoria.
- $P(B_r)$ Probabilidad a priori de que se presente un elemento de la clase B_r .
- $P(A)$ Probabilidad a priori de que se presente un objeto a clasificar con un vector de características igual a A .

4.4. – ESTIMACIÓN BAYESIANA

La aproximación bayesiana a los métodos estadísticos de estimación combina la información muestral⁷ con información adicional previa que puede parecer pertinente.

4.4.1. - Métodos clásicos de estimación

➤ Estimación puntual.

Una estimación puntual de algún parámetro poblacional θ es un valor único del estadístico X .

El estadístico que se utiliza para obtener una estimación puntual recibe el nombre de **estimador** o **función de decisión**. La obtención de un estimador no se espera sin error, es decir, no se estimará el valor con exactitud, lo que se busca es que el valor obtenido no se aleje mucho del valor real. Para esto existen varios métodos de estimación.

⁷ Una muestra es un subconjunto de una población

➤ **Estimador insesgado**

Un estimador insesgado es un estadístico X que tiene una media igual al parámetro estimado. Esto es:

$$\mu = E(X) = \theta.$$

Estimador más eficiente. – Si se consideran todos los estimadores insesgados posibles de algún parámetro θ , aquél con la variancia más pequeña recibe el nombre de estimador más eficiente de θ . (Figura 4.5)

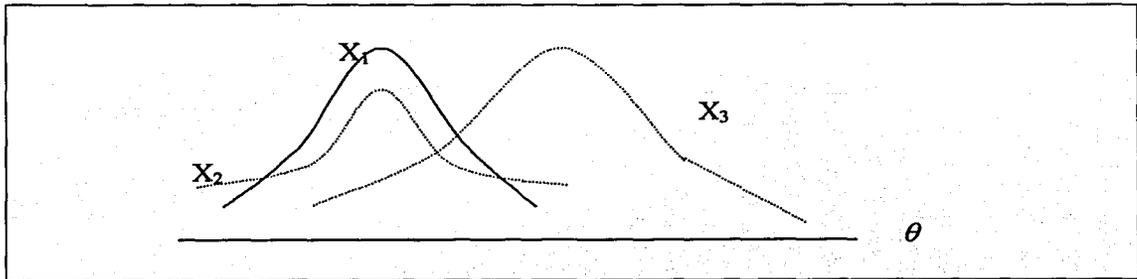


Figura 4.5. – Distribuciones muestrales para diferentes estimadores de θ .

4.4.2. – Métodos Bayesianos de estimación

4.4.2.1. - Probabilidad a priori.

Un concepto fundamental es la probabilidad a priori. Con él se quiere indicar el conocimiento a priori⁸ acerca de un evento antes de realizar un experimento.

Cuando la información adicional que se nos proporciona, nos dice la variación que genera una distribución de probabilidad dada $f(\theta)$, entonces $f(\theta)$ recibe el nombre de **distribución a priori**, con una **media a priori** μ y **variancia a priori** σ^2 . Es decir, se supone que θ es un valor de una variable aleatoria X con la distribución de probabilidad $f(\theta)$ y se desea estimar el valor particular θ para la población de la cual se seleccionó la muestra aleatoria.

⁸ Información que se sabe adicional

- **La probabilidad a priori de una clase.** – Con la probabilidad a priori de una clase se quiere indicar el conocimiento a priori acerca de esa clase antes de realizar un experimento. La probabilidad a priori de una clase j , se expresa como $P(w_j)$.

Las probabilidades a priori de J clases no pueden ser negativas y su suma debe ser la unidad:

1. $-P(w_j) \geq 0, j = 1, 2, \dots, J$
2. $\sum_j P(w_j) = 1$

Si se supone un problema de decisión entre dos clases, se puede plantear un clasificador basado únicamente en las probabilidades a priori:

Decidir por w_1 si $P(w_1) > P(w_2)$
Decidir por w_2 si $P(w_1) < P(w_2)$

Con este planteamiento, la probabilidad de cometer un error en la clasificación viene dado por el menor valor entre $P(w_1)$ y $P(w_2)$.

Para una mayor comprensión de este concepto veamos el siguiente ejemplo:

Supongamos un problema de clasificación en la que se desean discriminar personas sanas (clase 1) frente a personas anémicas (clase 2) y que no se dispone de otra información más que el *conocimiento a priori* sobre la frecuencia relativa de la enfermedad, recogido a partir de estadísticas sobre una población determinada:

- El 90% de la población está sana: $P(w_1) = 0.9$
- El 10% de la población está enferma: $P(w_2) = 0.1$

Si hubiera que clasificar a un paciente (del que desconocemos su analítica), ¿qué clase le asignamos?

A falta de otra información, hay que asignarle la clase con mayor probabilidad a priori: Entonces se puede decir que es una persona sana ya que $P(w_1) = 0.9 > P(w_2) = 0.1$.

4.4.2.2. – Probabilidad a posteriori

La probabilidad a posteriori consiste en la información tanto de la distribución a priori así como de la distribución muestral. La probabilidad a posteriori expresa el grado de credibilidad en la localización del parámetro θ después de que se ha observado la muestra.

- **Probabilidad a posteriori de una clase.** – Es la probabilidad de que la clase cierta sea w_j , dado que el valor observado es x .

Esta probabilidad supone conocidas $P(w_j)$ y $P(x|w_j)$ y supone también que se dispone de una medida x .

La probabilidad a priori y la densidad de probabilidad se combinan en la **Regla de Bayes**, que indica cómo el valor observado, x , modifica la decisión basada en $P(w_j)$, a través de $P(x|w_j)$ introduciendo la probabilidad a posteriori, $P(w_j|x)$.

La probabilidad a posteriori queda expresada de la siguiente manera:

$$P(w_j|x) = \frac{P(x|w_j)P(w_j)}{P(x)} \quad \text{donde } P(x) = \sum_{i=1}^J P(x|w_i)P(w_i)$$

Para un problema de dos clases con iguales probabilidades a priori ($P(w_1) = P(w_2) = P(w)$), tenemos que:

$$P(x) = P(x|w_1)P(w) + P(x|w_2)P(w) = P(w)(P(x|w_1) + P(x|w_2))$$

Por lo que las probabilidades a posteriori tienen la siguiente expresión:

$$P(w_1|x) = \frac{P(x|w_1)P(w)}{P(w)(P(x|w_1) + P(x|w_2))} = \frac{P(x|w_1)}{P(x|w_1) + P(x|w_2)}$$

$$P(w_2|x) = \frac{P(x|w_2)P(w)}{P(w)(P(x|w_1) + P(x|w_2))} = \frac{P(x|w_2)}{P(x|w_1) + P(x|w_2)}$$

Y, en cualquier caso,

1. Si $P(x|w_1) = 0$, $P(w_1|x) = 0$ y $P(w_2|x) = 1$
2. Si $P(x|w_2) = 0$, $P(w_2|x) = 0$ y $P(w_1|x) = 1$
3. Si $P(x|w_1) = P(x|w_2)$, $P(w_1|x) = P(w_2|x) = 0.5$
4. Para todo x , $P(w_1|x) + P(w_2|x) = 1$

El efecto de combinar densidad de probabilidad y probabilidad a priori es que la probabilidad a priori "escala" la densidad de probabilidad. Este efecto se muestra gráficamente en la figura 4.6.

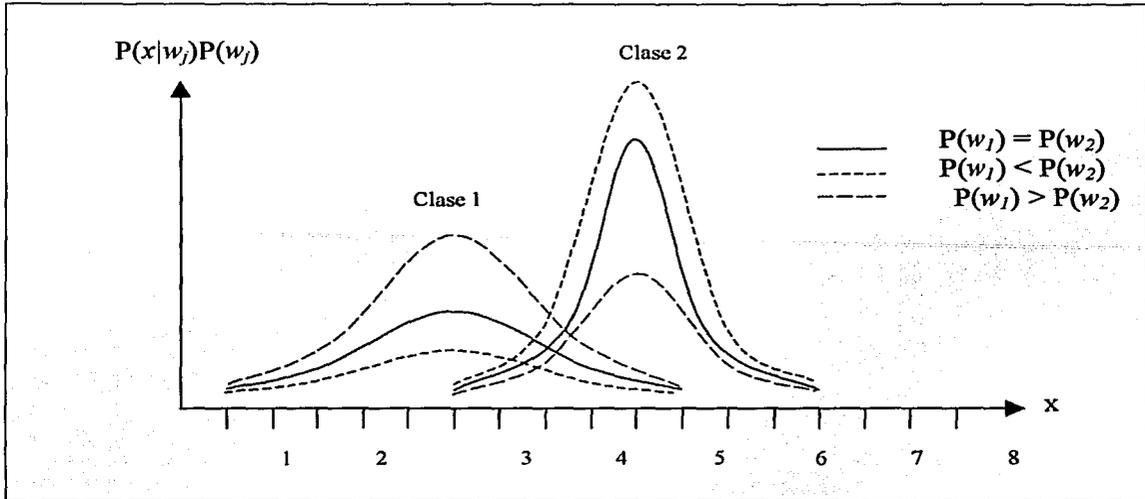


Figura 4.6. – Efecto de la probabilidad a priori sobre la probabilidad a posteriori.

La distribución a posteriori también puede quedar expresada en términos del parámetro θ de la manera siguiente:

$$f(\theta|x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n, \theta)}{g(x_1, x_2, \dots, x_n)}$$

➤ Estimación de Bayes e Intervalo Bayesiano

La media de la distribución a posteriori $f(\theta|x_1, x_2, \dots, x_n)$, representada por θ^* , recibe el nombre de estimación de Bayes de θ .

A su vez:

El intervalo $a < \theta < b$ se llama intervalo de Bayes del $(1-\alpha)100\%$ para θ si:

$$\int_a^b f(\theta|x_1, x_2, \dots, x_n) d\theta = \int_a^{\theta^*} f(\theta|x_1, x_2, \dots, x_n) d\theta = \frac{1-\alpha}{2}$$

4.5. – EJEMPLOS DE MODELOS MARKOVIANOS

4.5.1 - Problema de etiquetado

Un problema de etiquetado puede especificarse en términos de un conjunto de estados y un conjunto de etiquetas, en donde se asigna una etiqueta del conjunto de etiquetas a cada estado.

Un conjunto discreto con m estados es el siguiente:

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

Y a su vez, un estado puede representar un punto o una región del espacio euclidiano.

Un conjunto de etiquetas puede ser un conjunto discreto o continuo. En el caso continuo, el conjunto de etiquetas puede corresponder al conjunto de los números reales o a un intervalo de ellos. En el caso discreto, las etiquetas se toman dentro de un conjunto discreto de M valores.

Caso continuo:

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

Caso discreto:

$$L_d = \{l_1, l_2, \dots, l_M\} = \{1, 2, \dots, M\}$$

El conjunto de etiquetas puede ser caracterizado en términos de su continuidad. Además de la continuidad, otra característica esencial de un conjunto de etiquetas es poder establecer un orden entre las etiquetas. Por ejemplo, los elementos en el conjunto continuo de etiquetas \mathbb{R} pueden ser ordenados en relación a “ser más pequeño que”. En un conjunto discreto como, por ejemplo, la intensidad del tono de gris $\{0, 1, \dots, 255\}$, podemos ordenar los elementos según los valores de la intensidad $0 < 1 < \dots < 255$.

Definición. – Se llama etiquetado a

$$f = \{f_1, f_2, \dots, f_m\}$$

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

Definición. - Se llama “mapping” a toda función

$$f: S \rightarrow L$$

que a cada estado de S le asigna una única etiqueta f_i de L .

En terminología de Campos Aleatorios de Markov, se llama **configuración** al etiquetado. Una configuración de etiquetas puede corresponder a una imagen, a un mapa de bordes de la imagen, a una interpretación de los objetos que aparecen en escena, etc.

4.5.2 – Sistemas de vecinos y cliques

4.5.2.1 – Sistema de vecinos

Es la relación que se da entre los estados de S . Un sistema de vecinos para el conjunto de estados S es el siguiente:

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

donde N_i es el conjunto de los estados vecinos de i para los cuales se cumple:

1. $i \notin N_i$
2. $i \in N_j \Leftrightarrow j \in N_i$

Para un arreglo matricial S , el conjunto de vecinos de i está definido como el conjunto de estados próximos dentro de un radio r :

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

donde $dist(A,B)$ denota la distancia euclidiana entre A y B , y r es un número entero.

➤ Sistema de vecinos de primer orden

Llamado también sistema de 4-vecinos, en este sistema cada estado (interior) tiene cuatro vecinos, según lo mostrado en la figura 4.7, donde x denota el estado considerado y a sus vecinos.

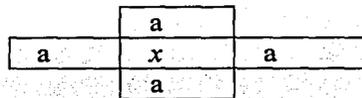


Figura 4.7. - Sistema de 4-vecinos

➤ **Sistema de vecinos de segundo orden**

También llamado sistema de 8-vecinos, hay ocho vecinos para cada estado (interior) según lo mostrado en la figura 4.8

a	a	a
a	x	a
a	a	a

Figura 4.8. – Sistema de 8-vecinos

➤ **Sistema de vecinos de orden n -ésimo**

En este sistema de vecinos, existen n vecinos para cada estado (interior.) Esto se muestra en la figura 4.9

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

Figura 4.9. – Sistema de n - vecinos

Los números $n = 1, 2, \dots, 5$ mostrados en la figura 4.9 indican los estados vecinos exteriores en el sistema de vecinos de orden n .

4.5.2.2. – Sistemas de Cliques

Un clique c para $\{S, N\}$ se define como un subconjunto de S tal que c consiste en un único estado $c = \{i\}$ o un par de estados vecinos $c = \{i, j\}$, o tres estados vecinos $c = \{i, j, k\}$ y así sucesivamente.

Denotamos el conjunto de cliques de un solo estado, el conjunto de cliques de dos estados, ..., por C_1, C_2, \dots respectivamente, donde

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

$$C_2 = \{\{i, j\} / j \in N, i \in S\}$$

$$C_3 = \{\{i, j, k\} / i, j, k \in S \text{ son vecinos dos a dos} \}$$

El conjunto de todos los cliques para $\{S, N\}$ es entonces $C = C_1 \cup C_2 \cup \dots$

El tipo de clique para $\{S, N\}$, de un arreglo matricial, viene determinado por su tamaño, dimensión y orientación. La figura 4.10 muestra distintos tipos de cliques para sistemas de vecinos de primer y segundo orden. A medida que aumenta el orden del sistema de vecinos, el número de cliques crece rápidamente.

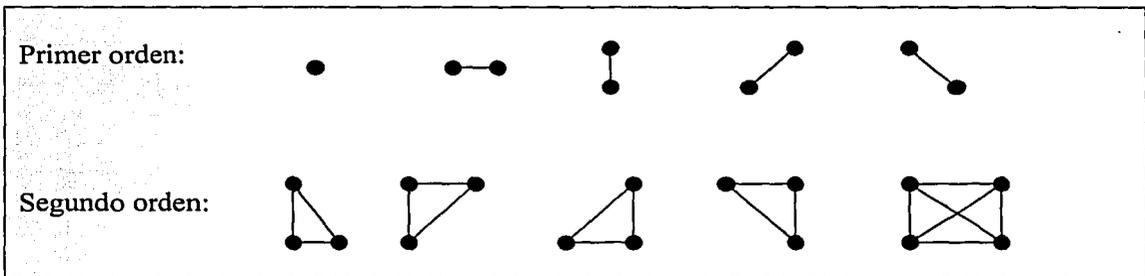


Figura 4.10. – Cliques en un arreglo matricial S

4.5.3.- Distribución de Gibbs

Diremos que un conjunto de variables aleatorias F es un Campo Aleatorio de Gibbs en S con respecto a N si y sólo si sus configuraciones siguen una distribución de Gibbs. Una distribución de Gibbs tiene la siguiente función de densidad:

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

donde

$U(f)$ = función de energía

Z = constante de normalización

T = temperatura.

La función de energía $U(f)$ está definida por:

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

la cual es la suma de las funciones "clique potenciales", $V_c(f)$, para todos los cliques posibles C . El valor de $V_c(f)$ depende de la configuración local del clique c .

$P(f)$ mide la probabilidad de la ocurrencia de una configuración determinada, o patrón f . Las configuraciones más probables son aquellas con energías más bajas.

La temperatura T controla la forma de la distribución. Cuando la temperatura es alta, todas las configuraciones tienden a estar igualmente distribuidas. Con la temperatura cercana a cero, la distribución se concentra alrededor de los mínimos globales de energía

4.5.4. - Modelo Ising

Este modelo, introducido por Lenz en 1920, se resolvió en una dimensión por Ising en 1925, y en dos dimensiones por Onsager en 1944.

El modelo Ising es el modelo más simple usado para transmitir información. Este modelo es muy utilizado debido a su simplicidad y bajo costo computacional. Se codifica en la función de energía de Gibbs como funciones clique potenciales para cada par de estados. Entonces la función clique potencial queda definida por:

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

que llamamos energía de segundo orden porque se calcula para cada clique formado por dos estados. Podemos especificar entonces un Campo Aleatorio de Gibbs o un Campo Aleatorio de Markov mediante la selección apropiada de V_1 y V_2 .

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

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

la ecuación anterior es un modelo que recibe el nombre de **auto-modelo**.

Un auto-modelo es llamado modelo **auto-logístico**, si las f_i toman valores en el conjunto discreto de etiquetas $L = \{0, 1\}$ o $L = \{+1, -1\}$. Entonces la correspondiente función de energía queda de la siguiente manera:

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

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

El modelo auto-logístico se reduce al modelo Ising cuando N es el sistema de vecindad más cercano en un arreglo matricial (en dos dimensiones 4-vecinos, en una dimensión 2-vecinos).

4.5.4.1. – Aplicación al ferromagnetismo

El modelo Ising, considera una secuencia $0, 1, 2, \dots, n$. En donde a cada punto le corresponde un pequeño giro hacia arriba o hacia abajo solamente (figura 4.11)

Se define la medida de probabilidad en un conjunto para todas las posibles configuraciones de los giros $w = (w_0, w_1, \dots, w_n)$. Cada giro es una función de la forma:

$$\delta_i = \begin{cases} 1 & \text{si } w_i \text{ es hacia arriba} \\ -1 & \text{si } w_i \text{ es hacia abajo} \end{cases}$$

La energía $U(w)$ es asignada a cada configuración de la siguiente manera:

$$U(w) = -J \sum_{i,j} \delta_i(w) \delta_j(w) - mH \sum_i \delta_i(w)$$

El término representa la energía causada por la interacción de los giros. La constante J es una propiedad del material que está bajo el análisis del modelo.

- Si $J > 0$, la interacción tiende hacia los puntos vecinos con los giros en la misma dirección (esto es bueno para el análisis en cuestión).
- Si $J < 0$, los giros vecinos se oponen a la orientación favorable (esto es malo para el análisis).

El segundo término de la ecuación representa la influencia de un magnetismo externo H y $m > 0$ es una propiedad del material que estamos analizando.

La medida de probabilidad entonces está dada por:

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

Donde T es la temperatura y k es una constante universal, mientras que Z que es una constante de normalización está definida por:

$$Z = \sum_w \exp\left(-\frac{1}{kT}U(w)\right)$$

Los giros están representados por dos coordenadas y un punto tiene 4 vecinos. En el caso de dos coordenadas, el límite es inestable, esta es la fase de transición.

4.5.5. - *Modelo Potts*

El modelo Potts al igual que el modelo Ising es muy usado para el ferromagnetismo, por lo que analizaremos primero el caso de imágenes y posteriormente como una aplicación al ferromagnetismo.

En el caso de imágenes de varios niveles de gris, es posible modelar las interacciones entre los píxeles a fin de segmentar en varias clases:

Siendo a y b dos estados de una región en 4-vecinos, tenemos lo siguiente:

$$V(a,b) = \begin{cases} +\beta & \text{si } f_a \neq f_b \\ -\beta & \text{si } f_a = f_b \end{cases}$$

también se puede expresar de la siguiente manera:

$$V(a, b) = 2\beta\delta(f_a, f_b)$$

esta ecuación es la expresión más conocida del Modelo Potts.

4.5.5.1. – Aplicación al ferromagnetismo

El modelo Potts es una extensión del modelo Ising, la diferencia es que el modelo Ising es un sistema de interacción de giro entrelazado que puede ser hacia arriba o hacia abajo, además que el modelo Ising se da en dos dimensiones solamente.

El modelo Potts considera un sistema de giros en q dimensiones, en donde cada punto del giro está igualmente espaciado, dependiendo de la dimensión que se esté tratando. Los vectores generados son las combinaciones lineales de los q vectores que apuntan a las direcciones simétricas de un hipertetaedro en $q-1$ dimensiones. (Figura 4.12)

La función de energía del modelo de Potts puede escribirse como:

$$U(w) = \sum_{i,j} J(\Theta_{i,j})$$

donde $J(\Theta)$ es el periodo 2π y $\Theta_{i,j}$ es el ángulo entre los giros de dos vecinos de i y j . El caso cuando $q = 2$ es el equivalente al modelo Ising.

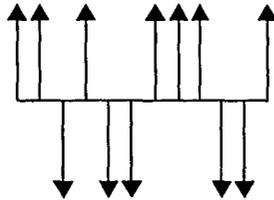


Figura 4.11– Modelo Ising.

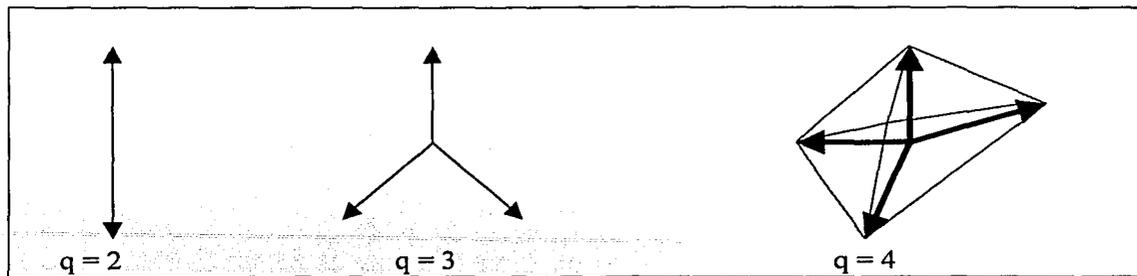


Figura 4.12. – Modelo Potts para diferentes valores de q .

CAPÍTULO 5

SEGMENTACIÓN Y FUSIÓN DE DATOS

5.1. – MODELO DE FUSIÓN

Los métodos utilizados hasta estos momentos para la segmentación de escenas urbanas obtenidas por Percepción Remota, están fundados, la gran mayoría de ellos, en un análisis de texturas, lo que conlleva en esta tesis a seguir un procedimiento basado en texturas.

Para aplicar el modelo de fusión son necesarias cuatro imágenes binarias, tres de las cuales son el resultado de la binarización de tres imágenes generadas del análisis de texturas por el método de Matriz de Co-ocurrencia, la otra imagen es un mapa rudimentario de la imagen original. Las etapas a seguir se enumeran a continuación.

Etapas 1. – Aplicar el método de la Matriz de Co-ocurrencia para un análisis de las texturas presentes en la imagen. De los resultados, seleccionar tres imágenes (las que mejor se adecuen) que nos muestren una segmentación en dos clases, en este caso, urbano y no urbano.

Etapas 2.- Realizar una binarización de las imágenes de texturas, con el objetivo de obtener información de los contornos presentes en la imagen, así como también información de la clase urbano. En este paso también se necesita binarizar el mapa de la imagen original.

Etapas 3.- Realizar el proceso de fusión para realizar una segmentación binaria, con los datos obtenidos en las etapas anteriores.

Dado que existen varios métodos de segmentación, se debe elegir la imagen que más se adecue a las necesidades que en un momento dado se tienen. Lo más utilizado es evaluar la confiabilidad de los algoritmos a través de un error global. El cálculo del error global es generado a partir del cálculo de error local para cada píxel, el cual una vez obtenido, genera la medida de error global para los resultados totales.

La medida de error global determina si un resultado es suficientemente bueno para ser usado en un proceso de fusión. Para las imágenes que son seleccionadas para considerarse en el proceso de fusión, la medida de error local determina cómo cada píxel, en resultados diferentes, debe ser tomado en cuenta en el proceso de fusión.

El píxel con una medida de error local pequeña tiene una gran confiabilidad, por lo tanto, tiene más influencia en el proceso de fusión.

De esta manera, se evalúa el error global para los resultados arrojados por diferentes algoritmos de segmentación, y se elige el que tenga el menor error global.

A continuación veremos el proceso de fusión.

Consideremos que:

s denota un estado de la imagen.

S denota el conjunto de estados de toda la imagen.

$L = \{l_i, 1 \leq i \leq N\}$ denota el conjunto de clases (etiquetas).

$M(s)$ denota la clase del estado s de acuerdo al mapa.

$I(s)$ denota la clase del estado s dado por el resultado del análisis de una imagen.

Se asume que

$$M(s) = I_i \text{ e } I(s) = I_j$$

$$r = 2$$

$$x_{l_1, l_2} = 20$$

$$x_{l_2, l_1} = 1$$

$$x_{l_1, l_1} = 0$$

$$x_{l_2, l_2} = 0.$$

Como los datos analizados son binarios, con valores 1 o 2, las x_{l_1, l_2} son el cambio de que un píxel tenga un valor igual a uno, y cambie a un píxel con valor igual a 2. Esto es para todas las x involucradas en el proceso.

El cambio de un estado de la clase I_i (en el mapa) a la clase I_j (en el resultado del análisis de la imagen) puede incurrir en un error que se puede llamar un "riesgo". Este riesgo viene dado de la siguiente forma:

$$x_{l_i, l_j} \begin{cases} = 0, & \text{si } l_i = l_j \\ > 0, & \text{si } l_i \neq l_j \end{cases}$$

el valor del riesgo depende de la etiqueta específica l_i y l_j . Generalmente el riesgo de cambiar de una etiqueta de un objeto natural a una etiqueta de un objeto construido es más pequeño que el de dirección inversa. El riesgo se estima usando información de mapas de la imagen original, así como de la imagen construida.

Como una regla general, un estado de la imagen no puede analizarse sin relacionar a sus vecinos. Considerando que s' es un sitio de la vecindad V_s del estado s , y $d(s, s')$ es la distancia entre s y s' . Se pueden calcular las siguientes dos medidas:

$$a(s) = \sum_{s' \in V_s} \frac{1\{M(s') = l_j\}}{d(s, s')}$$

$$b(s) = \sum_{s' \in V_s} \frac{1\{I(s') = l_j\}}{d(s, s')}$$

donde $1_{\{ \cdot \}}$ es el indicador de la función:

$$1_{\{ \cdot \}} = \begin{cases} 1 & \text{si } x \text{ es verdadero} \\ 0 & \text{si } x \text{ es otro caso} \end{cases}$$

La medida $a(s)$ es el número de estados en V_s que son clasificados como l_j en el mapa, medidos inversamente por sus distancias al estado s . Esto representa el “soporte” que el estado s obtiene de sus vecinos en el mapa para poder tomar a l_j como su clase (etiqueta).

La medida $b(s)$ representa el “soporte” que el estado s obtiene del resultado del análisis de la imagen para tomar la decisión.

El tamaño de la vecindad depende de la resolución de la imagen. Entre más alta sea la resolución, la vecindad puede ser de mayor tamaño.

5.1.1. - Las medidas de error local y global

La medida de error local $E(s)$ del estado s , tomando l_i como su etiqueta, está definida como:

$$E(s) = \frac{x_{l_i, l_j}}{ra(s) + b(s) + 1}, \quad r \geq 1$$

donde r es un coeficiente positivo. Aún cuando los resultados del análisis de la imagen y el mapa contengan incertidumbre e imprecisión, se usa el coeficiente r para reflejar el grado de confiabilidad que se tiene en el mapa con respecto al resultado de la imagen.

La medida de error global E_{global} es simplemente la suma de las medidas de errores locales en todos los sitios de la imagen, y viene dada por la siguiente fórmula:

$$E_{global} = \sum_{s \in S} E(s)$$

5.1.2. - Proceso de fusión

Una vez obtenida la medida del error global del resultado del análisis de una imagen, podemos decidir si el resultado es suficientemente bueno para ser usado en el proceso de fusión.

Para poder tomar una decisión sobre el resultado obtenido, simplemente se toma K como el mejor resultado entre todos los disponibles. Otra manera de hacer esto es usar un umbral h , y solo aquellos para los cuales E_{global} es menor que h , son retenidos para el proceso de fusión. Determinar el valor de h no es trivial.

Aunque el resultado seleccionado para ser usado en el proceso de fusión pueda tener una confiabilidad relativamente alta, todos los estados en este resultado no pueden tener la misma confiabilidad. Por lo cual, hay que calcular un coeficiente de confianza para cada estado del resultado basado en la medida de error local de este estado. El coeficiente de confianza queda entonces de la siguiente manera:

$$C(s) = 1 - \frac{E(s)}{\sum_{I_i, J_j \in L} x_{I_i, J_j}}$$

donde el denominador es para la normalización de $E(s)$ para que $0 \leq C(s) \leq 1$.

La decisión de la etiqueta (clase) final para un estado depende de la acumulación del coeficiente de confianza del estado s en ese resultado a través de todos los resultados retenidos, quedando de la siguiente manera:

$$I(s) = \arg \max_{I_i \in L} \sum_{k=1}^K C_k(s) 1_{\{I_k(s)=I_i\}}$$

donde $I_k(s)$ es la k -ésima entrada al proceso de fusión, y $C_k(s)$ es el coeficiente de confiabilidad de ese estado s en este resultado.

Algoritmo de fusión de datos:

1.- Para cada ventana de análisis de las imágenes de textura encontrar:

$a(s)$, $b(s)$, $E(s)$ y $C(s)$.

Esto se realiza para cada una de las imágenes de textura.

1.2.- Evaluar para $I_i(s) = 1$ y 2 :

$I_i = 1$ si $\{C_1(s) \delta(\text{ventana-de-textura}_1(s)=1) + C_2(s) \delta(\text{ventana-de-textura}_2(s)=1) + C_3(s) \delta(\text{ventana-de-textura}_3(s)=1)\}$.

$I_i = 2$ si $\{C_1(s) \delta(\text{ventana-de-textura}_1(s)=2) + C_2(s) \delta(\text{ventana-de-textura}_2(s)=2) + C_3(s) \delta(\text{ventana-de-textura}_3(s)=2)\}$.

1.3.- Encontrar el valor máximo de $I(s)$, y poner en la imagen de resultado final el valor I_i correspondiente.

Donde $a(s)$ es el número de veces que ocurre en el mapa el píxel analizado de textura / distancia euclidiana.

$b(s)$ es el número de veces que ocurre en la propia imagen de textura el píxel analizado / distancia euclidiana.

5.2. – MODELO DE ENERGÍA

En el modelo de energía deben reforzarse los rasgos deseables de las imágenes, para lo cual deben considerarse los siguientes factores:

1. La imagen bajo análisis debe ser lisa, esto es que la desviación de valores del píxel alrededor de su vecindario debe ser pequeña.
2. Las proporciones cruzadas permanecen más o menos iguales alrededor de su vecindario.

El modelo de energía es una combinación lineal de los dos factores anteriores. Además de estos factores, existe otro aspecto muy importante para que los valores de la componente espectral de la imagen permanezcan inalterados. Este aspecto del modelo de energía puede ser su dependencia de los valores de su pendiente⁹, o también puede ser información del borde. Entonces lo que se propone hacer es una máscara gaussiana del modelo, con el fin de reforzar los factores a lo largo de las direcciones de pendientes bajas. Esto queda de la siguiente manera:

$$\Phi(i + m + n) = \frac{1}{\sqrt{2\pi\sigma(i, j)}} e^{-\frac{(c(i, j) - \sum_{s=1}^3 (f_s(i, j) I_s(i+m, j+n))^2 / 2\sigma(i, j))}{2\sigma(i, j)}}$$

donde:

- I_s denota al rojo, verde y las componentes espectrales de una imagen de color, y s toma los valores de 1, 2 y 3, respectivamente.
- $f_s(i, j)$ es una función que enmascara a la imagen de la siguiente manera:

$$f_1(i, j) = \begin{cases} 1 & \text{si la máscara en } i, j \text{ es roja} \\ 0 & \text{en otro caso} \end{cases}$$

⁹ Inclinación que tiene un píxel respecto a sus píxeles vecinos

Similarmente, $f_2(i,j)$ y $f_3(i,j)$ están definidas para las máscaras verde y azul.

- $c(i,j)$ es una función que está denotada de la siguiente manera:

$$c(i, j) = \sum_{s=1}^3 (f_s(i, j) I_s(i, j))$$

- $\sigma(i,j)$ es una función del mínimo valor de pendiente (diferencia absoluta entre dos valores de píxeles vecinos), tomándose en consideración la siguiente función para σ :

$$\sigma(i, j) = \underset{\forall (m,n) \in \eta^1(0,0)}{\text{mínimo}} \left\{ \frac{|c(i+m, j+n) - c(i-m, j-n)|}{2\sqrt{m^2 + n^2}} \right\} + \sigma_i$$

el valor de σ_i se toma como un valor constante igual a 10. Puede notarse aquí, que para poder estimar el valor de σ , se considera que las pendientes de las componentes espectrales se ponen en correlación.

Se usa la misma forma para estimar la fuerza del ruido de $(n(i,j))$ a (i,j) para cada componente espectral.

5.2.1. – Modelo de energía local

El modelo de energía local para cualquier píxel se presenta para las siguientes expresiones:

- Para el criterio de suavidad:

$$U^{(sm)}(i, j) = \sum_{m=-p}^p \sum_{n=-p}^p \sum_{s=1}^3 \left\{ (1 - f_s(i, j)) \Phi(i+m, j+n) (I_s(i, j) - I_s(i+m, j+n))^2 \right\}$$

- Para la fiabilidad de las proporciones cruzadas:

$$U^{(\rho)}(i, j) = \sum_{m=-p}^p \sum_{n=-p}^p \sum_{s=1}^3 \left\{ (1 - f_s(i, j)) \Phi(i+m, j+n) \times \left(\frac{I_s(i, j)}{c(i, j)} - \frac{I_s(i+m, j+n)}{\sum_{t=1}^3 f_t(i, j) I_t(i+m, j+n)} \right)^2 \right\}$$

El modelo de energía local es la combinación lineal de los dos términos anteriores y un término adicional para restaurar la imagen del ruido gaussiano que se le adicionó. El modelo de energía local es el siguiente:

$$U(i, j) = \mu U^{(sm)}(i, j) + \lambda U^{(\rho)}(i, j) + \gamma \sum_{s=1}^3 \left\{ (1 - f_s(i, j)) \frac{(I_s(i, j) - I_s^{(0)}(i, j))^2}{2\sigma^2(i, j)} \right\}$$

Es interesante notar que el término $(1 - f_s(i, j))$ ignora cualquier cambio en el valor verdadero de la respectiva componente espectral de (i, j) .

5.2.2. - Modelo de energía global

La función de energía global es la suma de las funciones de energía locales:

$$U = \sum_{\forall i} \sum_{\forall j} U(i, j)$$

CAPÍTULO 6

RESULTADOS

Para la realización de la presente tesis, se eligieron dos imágenes de niveles de gris de tamaño 512 X 512 píxeles, a las cuales se les aplicó el método estadístico de Matriz de Coocurrencia, y se les aplicaron once descriptores de Haralick. Se realizaron diversos experimentos a fin de determinar adecuadamente los parámetros de análisis de texturas. Con ese propósito, se tomó como distancia los valores de 1,2 y 3, además de las orientaciones de 0° y 180°. Los niveles de gris se redujeron a 32, y las ventanas de análisis fueron de 9 X 9.

Posteriormente, de las once imágenes generadas por el método de Matriz de Coocurrencia y los descriptores de Haralick se eligieron tres, las cuales fueron Energía, Prominencia y Varianza. Estas imágenes se seleccionaron porque son las que mejor muestran una diferencia entre las clases urbano y no urbano.

Para poder realizar una fusión de datos, las tres imágenes seleccionadas se binarizaron, para lo cual se utilizó el histograma y, tomando en cuenta el valor de los dos picos más altos, se buscó un umbral de binarización. Esto es que se tomó el valor del pico más alto cercano al nivel de gris 255, así como el que le seguía en altura, y se restó el valor más alto al más pequeño y al resultado se le dividió entre dos, obteniendo así el umbral. Esto se realizó para las tres imágenes seleccionadas.

En nuestras imágenes de resultados, la clase urbano está representada por el nivel de gris 255 (color blanco), mientras que la clase no urbano está representada por el nivel de gris 0 (color negro).

Después de la binarización se realizó la fusión de las tres imágenes, pero para realizarla se necesitó además de un mapa binario rudimentario y sin mucha aproximación de la imagen original, por lo cual también se binarizó la imagen original.

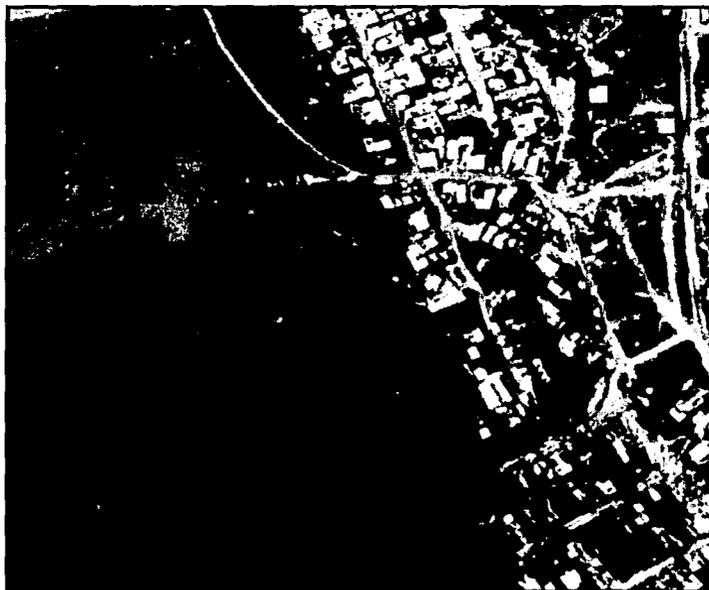
Los resultado se muestran a continuación.

6.1.- IMÁGENES DE PRUEBA

Imagen de prueba #1:



Imagen de prueba # 2:



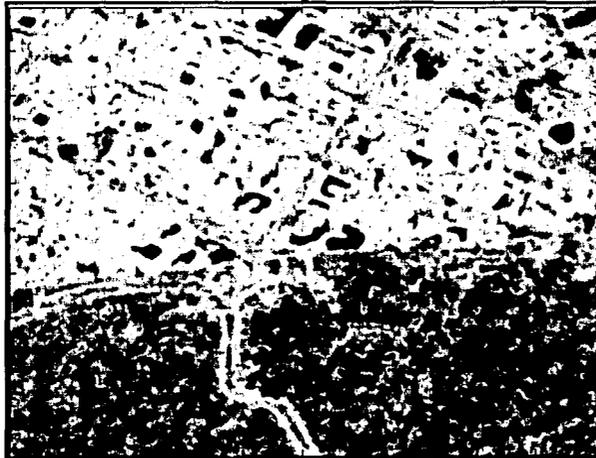
6.2.- RESULTADOS CON OPERADOR DE ENTROPÍA

Para la imagen de prueba #1, tenemos las imágenes generadas con el descriptor de Entropía, tomando como distancias los valores de 1 y 3, así como los ángulos de rotación 0° y 180° :

Entropía $d = 1$:



Entropía $d = 3$:



Como se puede ver, la distancia más aceptable es $d=3$, por lo que esta distancia fue la que se tomó como base para generar las siguientes texturas

6.3.- RESULTADOS DE OPERADORES DE TEXTURAS

Utilizando la imagen de prueba #1, se obtuvo lo siguiente:

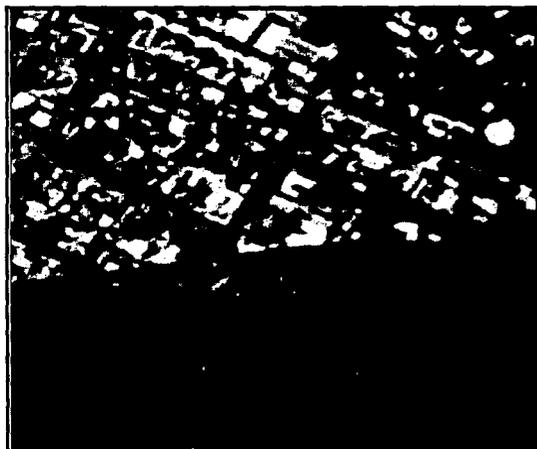
Contraste:



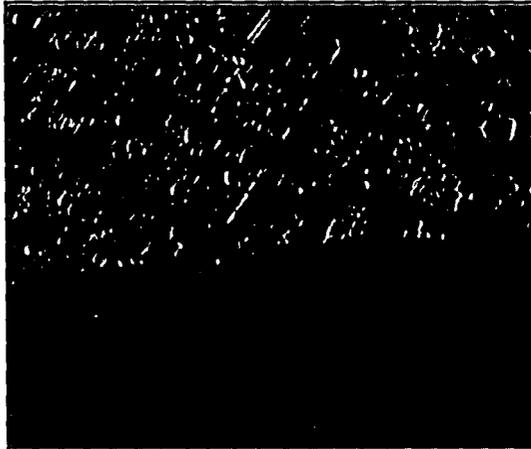
Correlación:



Autocorrelación:

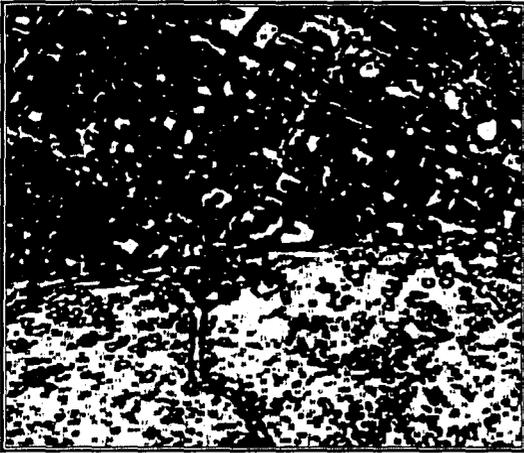


Disimilitud:

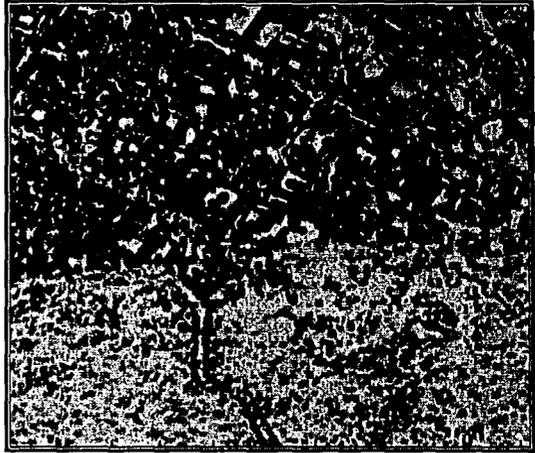


Se puede observar en estas cuatro imágenes, que ninguna de ellas muestra una diferencia entre urbano y no urbano.

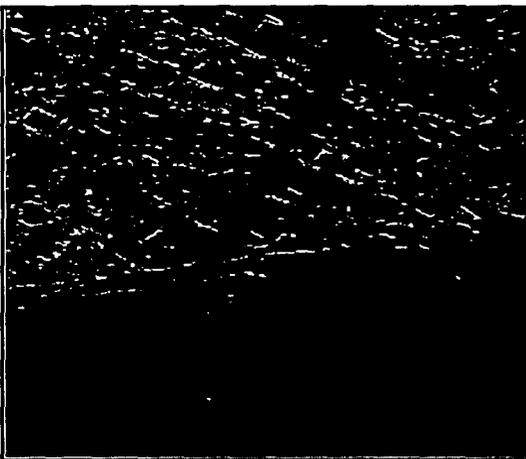
Energía:



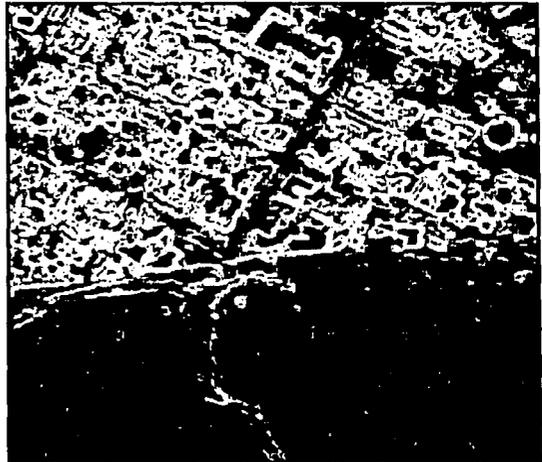
Homogeneidad:



Matiz



Prominencia:



De estas imágenes, se nota que las de energía y prominencia son las que mejor muestran una segmentación de urbano y no urbano.

Tendencia:

Varianza:



En estas dos imágenes se nota que la mejor es la de varianza, aunque la de tendencia también es muy contrastada, en la parte del bosque, se notan algunas regiones que no tiene la de varianza.

Como se observa en las once imágenes de textura generadas por cada uno de los descriptores utilizados, las que más nos acercan a una segmentación, son las de Energía, Prominencia y Varianza.

A estas tres imágenes se les aplicó el algoritmo de binarización explicado al principio de este capítulo, dando como resultado, las imágenes binarias necesarias para el proceso de fusión de datos.

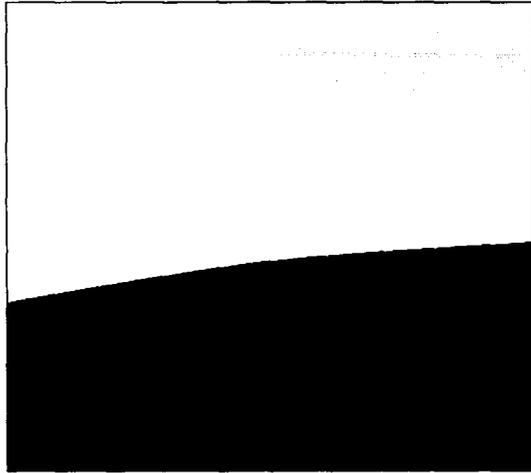
También, como se explicó en el capítulo 5, se necesita un mapa de la imagen original. Los resultados para la imagen de prueba #1, se muestran a continuación.

6.4.- RESULTADOS BINARIOS PARA LA IMAGEN DE PRUEBA #1

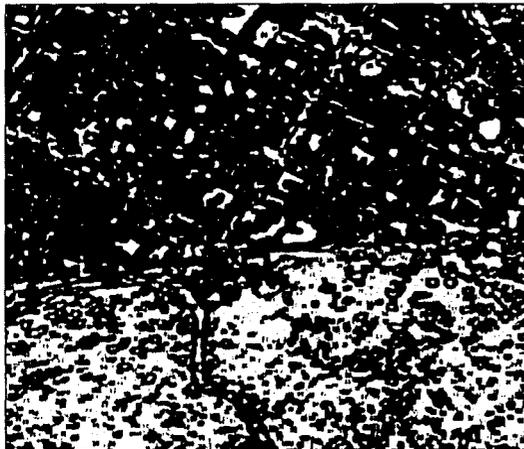
Imagen original de prueba #1 :



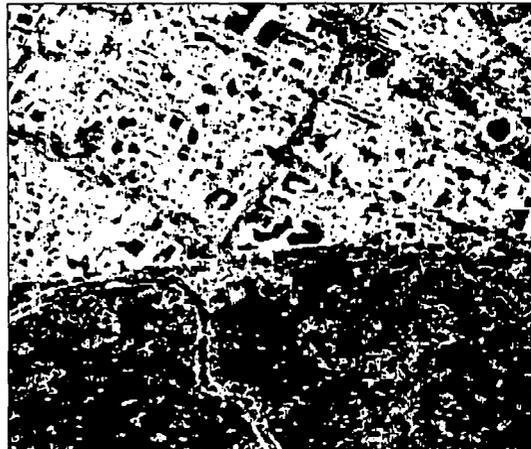
Mapa:



Energía:



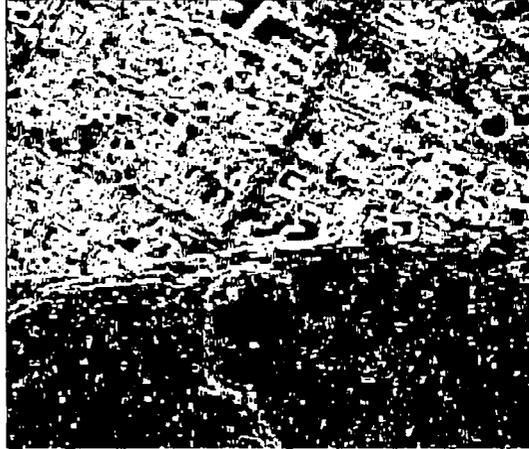
Energía binaria:



Prominencia:



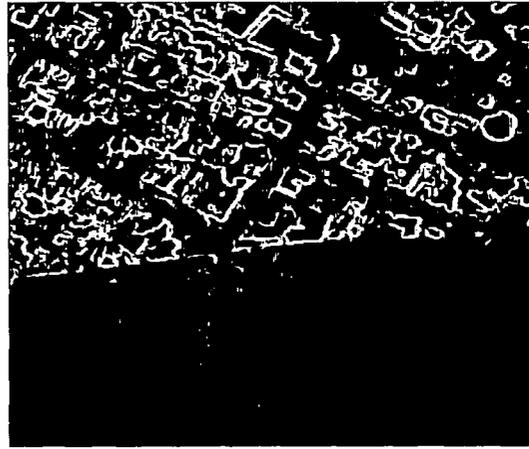
Prominencia binaria:



Varianza:



Varianza binaria:



6.5. – RESULTADO FINAL DE LA FUSIÓN DE DATOS PARA LA IMAGEN DE PRUEBA #1

Imagen original:

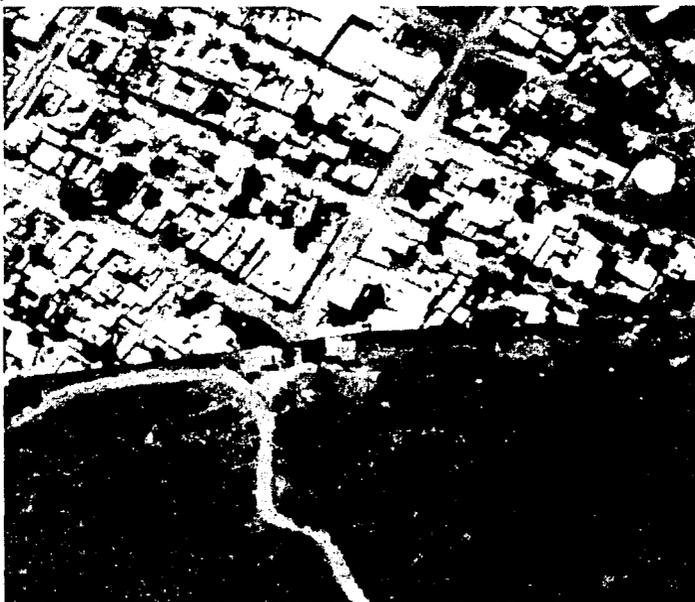


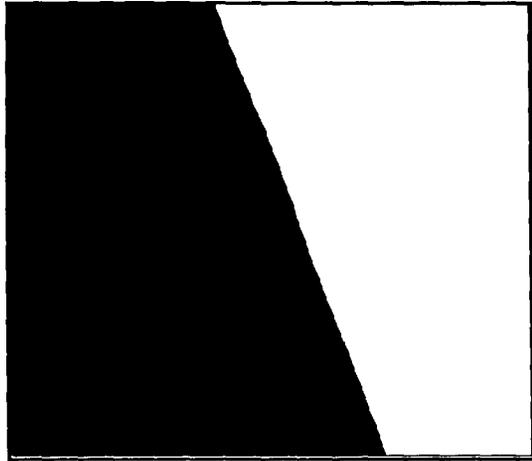
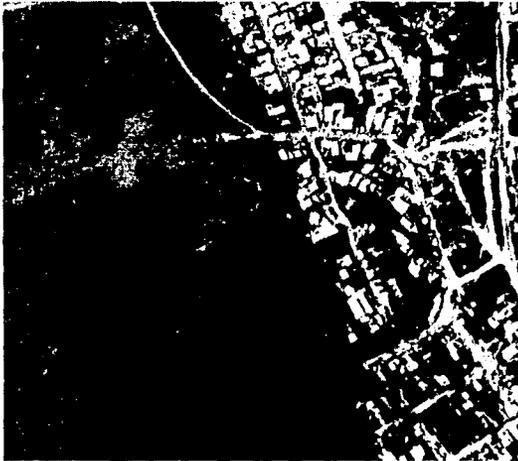
Imagen de resultado final:



6.6.- RESULTADOS BINARIOS PARA LA IMAGEN DE PRUEBA #2

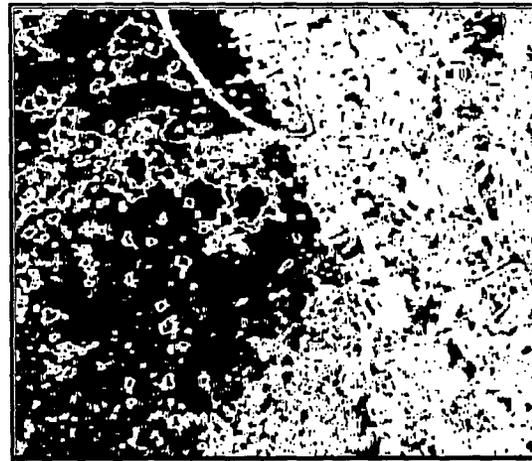
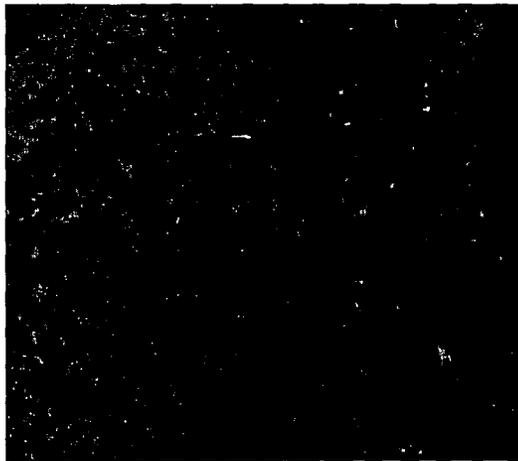
Imagen original de prueba #2 :

Mapa:

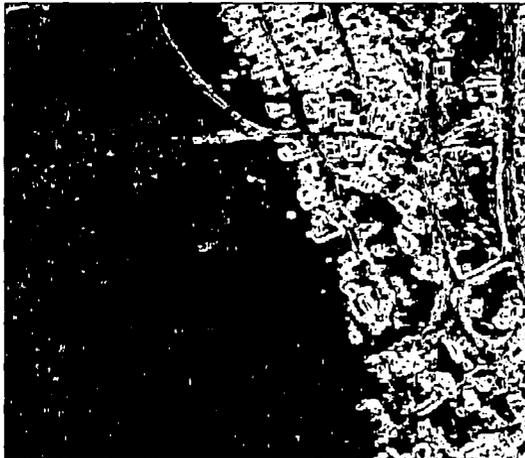


Energía:

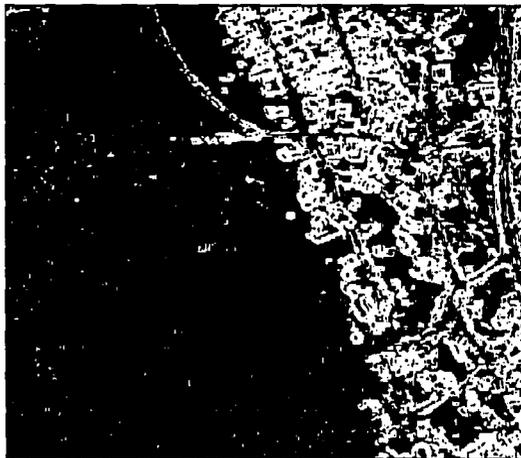
Energía binaria:



Prominencia:



Prominencia binaria:



Varianza:



Varianza binaria:



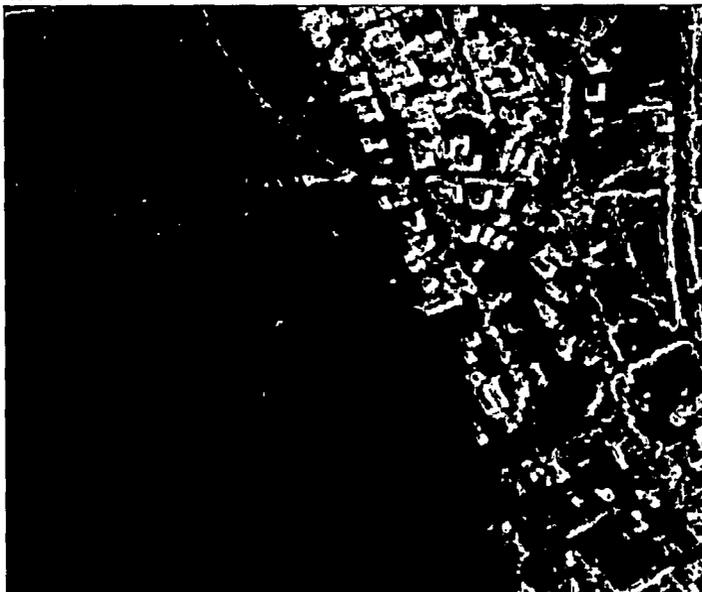
TESIS CON
FALLA DE ORIGEN

6.7. – RESULTADO FINAL DE LA FUSIÓN DE DATOS PARA LA IMAGEN DE PRUEBA #2

Imagen original:



Imagen de resultado:



CAPÍTULO 7

CONCLUSIONES

Una de las mayores aplicaciones de la Percepción Remota ha sido el análisis de imágenes, dentro de este campo, destacan los campos de investigación enfocados al procesamiento digital de imágenes, esto con la finalidad de realizar aplicaciones en diferentes tipos de regiones, ya sean urbanas o no urbanas.

Es por lo tanto, necesario el uso de diversos algoritmos que generen diferentes tipos de segmentación de imágenes para poder distinguir las diferentes zonas de interés, ya sean urbanas o no urbanas. Las imágenes de regiones urbanas suelen utilizarse para realizar planeación urbana, mientras que las imágenes de zonas no urbanas se utilizan para cartografía, valoración de recursos forestales y la agricultura.

Diversos experimentos se emprendieron empleando Campos Aleatorios de Markov. Al emplear solo información de niveles de gris en el proceso de segmentación markoviano, se observaron notorios errores en los resultados, que sirvieron para fundamentar nuestro propósito de incorporar la información textural a nuestro esquema de segmentación.

El esquema planteado en esta tesis fue la realización de una fusión de datos, esto con el objetivo de integrar información parcial al proceso de segmentación. Se logró obtener la segmentación deseada, pero como se nota en los resultados finales, las zonas homogéneas urbanas presentan cierta confusión con las zonas no urbanas (en este caso, el bosque).

El algoritmo planteado fue el siguiente:

- 1.- Segmentación de la imagen original.
- 2.- Binarización.
- 3.- Clasificación (fusión de datos).

Dentro de la segmentación, se utilizó el método de la Matriz de Co-ocurrencia y los descriptores de Haralick, estos experimentos presentaron una desventaja, la cual fue el tiempo de procesamiento del programa (para la imagen de 512 X 512, una ventana de análisis de 9 X 9 y una cuantización de niveles de gris de 32, el tiempo estimado real fue de 72 horas), y como se había mencionado antes, el método confunde las zonas homogéneas urbanas con zonas no urbanas. En este paso se generaron once imágenes de la imagen de prueba #1 que, como se indicó en el capítulo correspondiente, solamente se eligieron tres de ellas para el proceso de fusión de datos, las cuales fueron las que mejor mostraban una segmentación de dos clases (urbano y no urbano).

En el paso de la binarización se hizo uso del histograma de cada una de las seis imágenes seleccionadas (tres de la imagen de prueba #1, y tres de la imagen de prueba #2), aplicándoles el algoritmo de binarización a cada una de ellas y, como se nota en el capítulo de resultados, se obtuvo una muy buena binarización, distinguiéndose perfectamente dos clases, a excepción de las zonas homogéneas urbanas, las cuales pasaron a ser zonas no urbanas.

Para realizar la clasificación de la imagen de prueba #1, se propusieron como datos las imágenes de textura llamadas energía, prominencia y varianza todas ellas binarizadas, obteniéndose como resultado una imagen en dos clases (urbano y no urbano), el cual había sido nuestro objetivo principal, y como se ve en la imagen de resultado final, la cual solamente fue una superposición de la imagen de resultado binaria, sobre la imagen original, esto es, que a las regiones no urbanas se les asignó un valor único de nivel de gris (en este caso el color negro), mientras que a las regiones urbanas se les asignó todo el rango de niveles de gris presentes en la imagen original.

Para la clasificación de la imagen de prueba #2, ya no se realizaron todos los experimentos de segmentación por texturas, puesto que ya se había comprobado con la imagen de prueba #1 cada uno de los descriptores de Haralick mencionados en esta tesis, solamente se realizaron los experimentos necesarios para obtener las imágenes de texturas con distancia $d=3$ para energía, prominencia y varianza, obteniéndose una fusión de datos en donde se notan claramente dos clases. Observando la imagen de resultado se nota que el algoritmo, al igual que para la imagen de prueba #1, confunde las zonas homogéneas urbanas con zonas no urbanas.

De manera general, se puede concluir que nuestros resultado finales proporciona datos sobre regiones urbanas y no urbanas, y una desventaja del esquema planteado es la confusión que tiene de zonas homogéneas urbanas con zonas no urbanas. Puesto que el objetivo era la generación de datos de las regiones, se deduce que se logró el objetivo y con ello la solución del problema planteado.

BIBLIOGRAFÍA

1. LIBROS DE CONSULTA

- **Digital Image Processing**
González, Rafael C. & Woods Richard E.
Editorial Addison-Wesley Publishing Company
Estados Unidos de América, 1993.
- **Vision in Man and Machine**
Martin D. Levine
Editorial McGraw-Hill
Estados Unidos de América, 1985.
- **Digital Image Processing**
Pratt William K.
Editorial John Wiley & Sons, 2ª. ed.
Estados Unidos de América, 1991.
- **Computer and Robot Vision**
Robert M. Haralick & Linda G. Shapiro.
Volumen I
Editorial Addison Wesley.
Estados Unidos de América, 1992.
- **Markov Random Field. Theory and Application.**
Dr. Rama Chellapa & Anil Jain.
Editorial Academic Press.
Estados Unidos de América, 1993.
- **Digital Image Processing.**
Kenneth R. Castleman.
Editorial Prentice-Hall.
Estados Unidos de América, 1996.
- **Markov Random Field Modeling in Computer Vision.**
S.Z. Li.
Editorial Springer-Verlag Tokio.
Hong Kong, 1995.

- **Probabilidad y Estadística**

Ronald E. Walpole & Raymond H. Myers
Editorial McGraw-Hill, 4ª. Ed.
México, 1993.

- **Probabilidad y Estadística para Ingeniería y Administración**

William W. Hines & Douglas C. Montgomery
Editorial: Compañía Editorial Continental, 3ª. Ed.

- **Probabilidad y Estadística. Aplicaciones y Métodos**

George C. Canavos
Editorial McGraw-Hill, 1a. Ed.
México, 1994.

- **Probability and Stochastic Processes for Engineers**

Carl W. Helstrom
Editorial Prentice-Hall, 2ª. Ed.
Estados Unidos de América.

- **Estadística**

Richard C. Weimer
Editorial CECSA, 2ª. Ed.
México, 1999.

2. DIRECCIONES ELECTRÓNICAS:

- **Introducción al Procesamiento y Análisis de Imágenes Digitales**

Javier Abad & R. Molina.

Departamento de Ciencias de la Computación e I.A.

Universidad de Granada, España.

http://www-etsi2.ugr.es/depar/ccia/mia/complementario/Procesamiento-Imagenes/Procesamiento_Imagenes.html

- **Visión por Computador.**

Gabriela Andreu y José Miguel Valiente.
Departamento de Informática de Sistemas y Computadores.
Universidad Politécnica de Valencia.
http://miron.disca.upv.es/vxc/Apuntes_VxC.htm
<http://carpanta.dc.fi.udc.es/~cipenedo/cursos/lp/Tema6/nodo6/1.html>
<http://www.elprisma.com/apuntes>

- **Teoría de los MRF**

<http://www-etsi2.urg.es/depar/ccia/rf/www>
<http://www.terraes/personal2>
<http://physics.categu.ac.kr/~wphong/parallel/BOOK.html>
<http://www.netlib.org/utk/lis/pcwLSI/text/node45.html>

- **Satélites**

http://eosmis.cr.usgs.gov:5725/DATASET_DOCS/landsat7_dataset.html
<http://www.ccrs.nrcan.gc.ca/ccrs/tekrd/satsens/sats/landsate.html>
http://poda.ac.jpl.nasa.gov:2031/SOURCE_DOCS/seasat.html
http://www.asf.alaska.edu/source_documents/jers1_source.html
<http://www.nasda.go.jp/Home/Satellites/eljers-e.html>
http://www.space.gc.ca/publications/pdf/csa_sec/ear_env/rsat1/pdfs/davidbenneralos.pdf
<http://www.crisp.nus.edu.sg/rsat/rsat.html>

- **Texturas**

<http://www.cvc.uab.es/castella/research/arees/taca/taca.html>
<http://plutarco.disca.upv.es/~jcperez/pupas/>
<http://www2.ing.puc.cl/viing/ed430/busqueda-y-sintesis-de-texturas.html>
<http://www.depeca.alcala.es/docencia/ITC-El/va/-11k>
<http://www.icaen.uiowa.edu/~LECTURE/Texture1.html>

- **Reconocimiento de Patrones y Análisis de la Imagen**

<http://codazi4.igac.gov.co/gsdi5/documentos/y-bechack-paper.pdf>
<http://bolfor.chemonics.net/Publicaciones/Simposio/martinez.pdf>
<http://www.cfnavarra.es/territorial2000/COMUNICA/TELEDETECCION.PDF>

3. PUBLICACIONES

- **Toward Robust Analysis of Satellite Images Using Map Information-Application to Urban Area Detection.**

Shan Yu, Marc Berthod, and Gérard Giraudon.
IEEE Transactions on Geoscience and Remote Sensing.
Volumen 37, No.4, Julio 1999.

- **Improving Satellite Image Analysis Quality by Data Fusion.**

Shan Yu.
IEEE, 1995.

- **Aplicación de los Modelos de Campos Aleatorios de Markov en Visión Artificial.**

M. Pujol, R. Rizo, P. Arques, P. Compañ, F. Escolano y R. Molina.
Revista Electrónica de Visión por Computador.
Número 4, Octubre 2000.

- **Image Classification Based on Color and Texture Analysis.**

Begoña Acha y Carmen Serrano.
Escuela Superior de Ingenieros, Universidad de Sevilla, España .
Junio 2000.

- **An Investigation of the Textural Characteristics Associated With Gray Level Co-occurrence Matrix Statistical Parameters.**

Andrea Baraldi y Flavio Parmiggiani.
IEEE Transactions on Geoscience and Remote Sensing.
Volumen 33, No.2, Marzo 1995.

- **Texture Análisis of SAR Sea Ice Imagery Using Gray Level Co-occurrence Matrices.**

Leen-Kiat Soh, and Costas Tsatsoulis.
IEEE Transactions on Geoscience and Remote Sensing.
Volumen 37, No.2, Marzo 1999.

4. TESIS

- **Segmentación de texturas en imágenes radar utilizando la matriz de co-ocurrencia conjunta.**

Juan Manuel Valdes Casillas.
Facultad de Ingeniería.
Ciudad Universitaria, México D.F. 2001.

- **Fusión de datos y segmentación de imágenes de Percepción Remota.**

María Enriqueta Barilla Pérez y José Antonio Franco Bello.
Facultad de Ingeniería.
Ciudad Universitaria, México D.F. 2001.

- **Modelo matemático del proceso de desertificación empleando variables canónicas y técnicas de Percepción Remota.**

Nidiyare Hevia Montiel.
Facultad de Ingeniería. División de Estudios de Posgrado.
Ciudad Universitaria, México D.F. 2001.

5. NOTAS DE CURSO

- Probabilidad y Estadística

Ing. Mario Moreno Jiménez.
Facultad de Ingeniería.
UNAM, México D.F. 1995.

- Procesamiento Digital de Imágenes

Dr. Víctor García Garduño.
Facultad de Ingeniería.
UNAM, México D.F. 2001.

- Comunicaciones Digitales

Ing. Juan Fernando Solórzano Palomares.
Facultad de Ingeniería.
UNAM, México D.F. 1999.

- Reconocimiento de Patrones

Dr. Miguel Moctezuma Flores.
Facultad de Ingeniería.
UNAM, México D.F. 2001.

ANEXO

DISEÑO Y PROGRAMACIÓN

1. – DISEÑO

Puesto que nuestro planteamiento de solución es el empleo de la técnica de análisis de texturas llamada Matriz de Co-ocurrencia, la cual, con la ayuda de los operadores de Haralick, nos permitirán definir un esquema de fusión de datos, esto con el fin de integrar información parcial al proceso de segmentación, los pasos del diseño se enumeran y explican a continuación.

1.- Análisis del problema:

En este paso se analizaron las imágenes de prueba con el objetivo de obtener un panorama visual para poder encontrar el tamaño mínimo de la ventana necesaria que permita integrar la información parcial al proceso de segmentación.

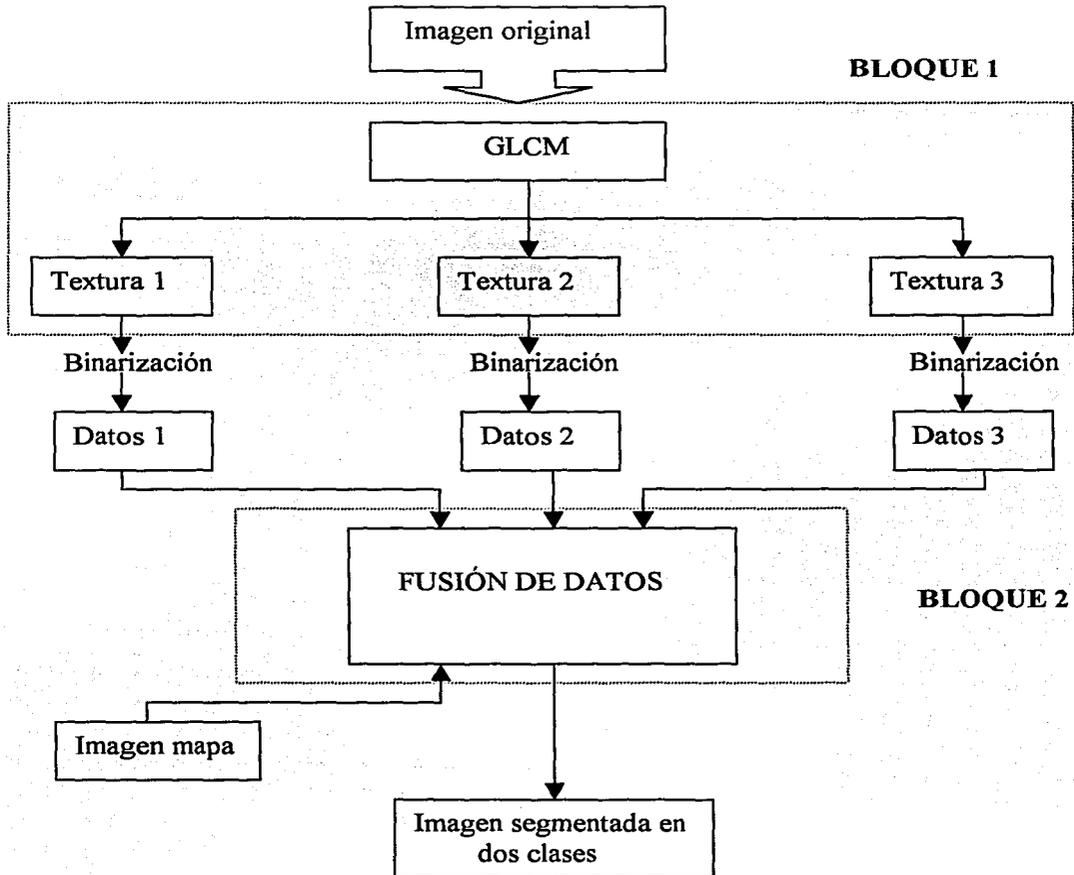
2.- Desarrollo:

Dentro de este paso, existen otros que son los siguientes:

- **Elección del lenguaje de programación.-** Es en este paso en donde se empieza a elegir el lenguaje de programación, para este caso se eligió el lenguaje de programación MATLAB puesto que este lenguaje cuenta con el Toolbox de Imágenes, el cual es un lenguaje que ya tiene incorporadas funciones especiales para el tratamiento específico de imágenes. A diferencia del lenguaje de programación C, el cual es un lenguaje estructurado y que no cuenta con las funciones de tratamiento de imágenes. Aunque una desventaja de MATLAB frente a C, es el tiempo de ejecución del programa, ya que MATLAB utiliza demasiados procesos para poder generar su salida, C es muchísimo más rápido.
- **Programación.-** Este punto lo trataremos un poco más adelante, pero se podría decir que es aquí donde se generaron los diagramas de bloque y de flujo para los programas, así como la ejecución de cada uno de ellos.
- **Análisis de resultados.-** Aquí se analizan cada uno de los resultados obtenidos, verificando que sean los correctos.

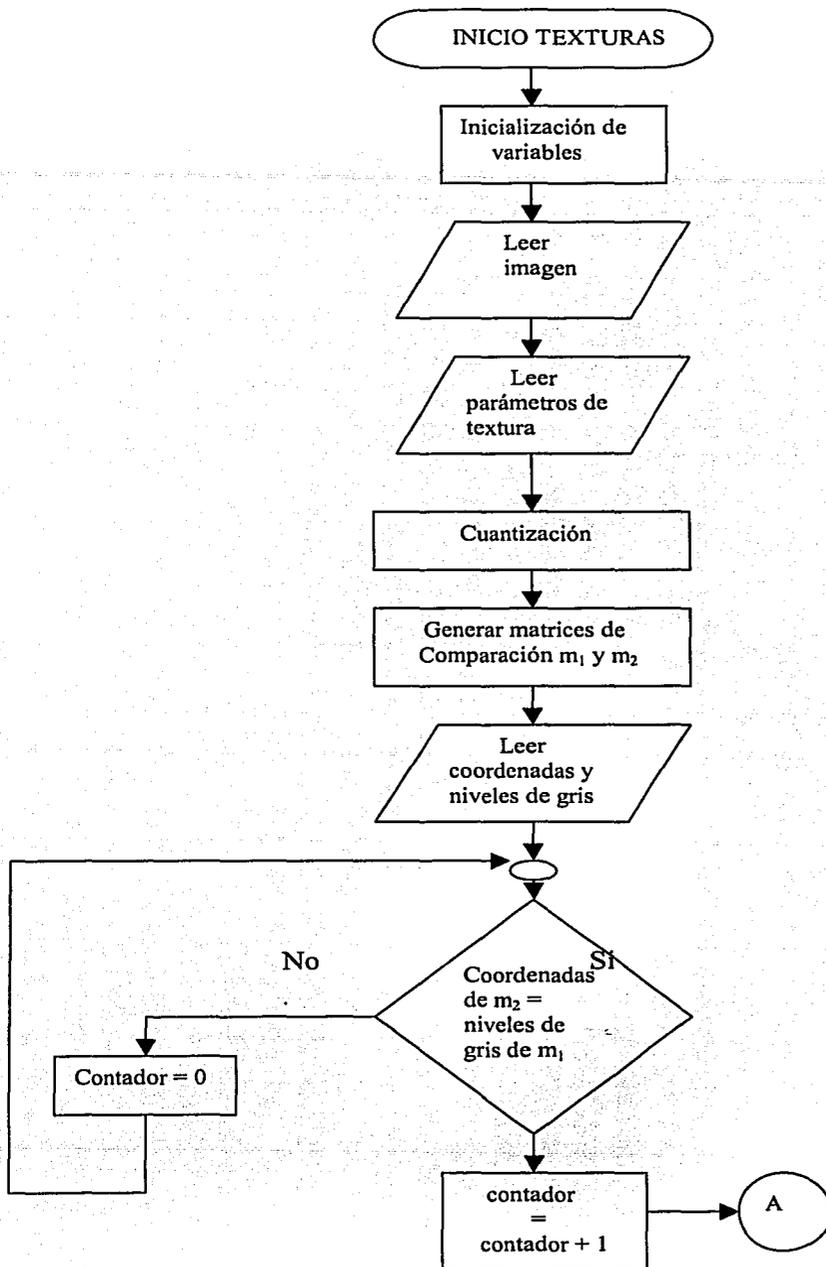
2. – PROGRAMACIÓN

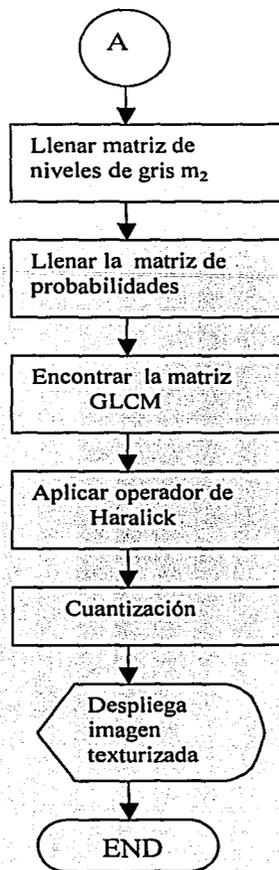
DIAGRAMA DE BLOQUES:

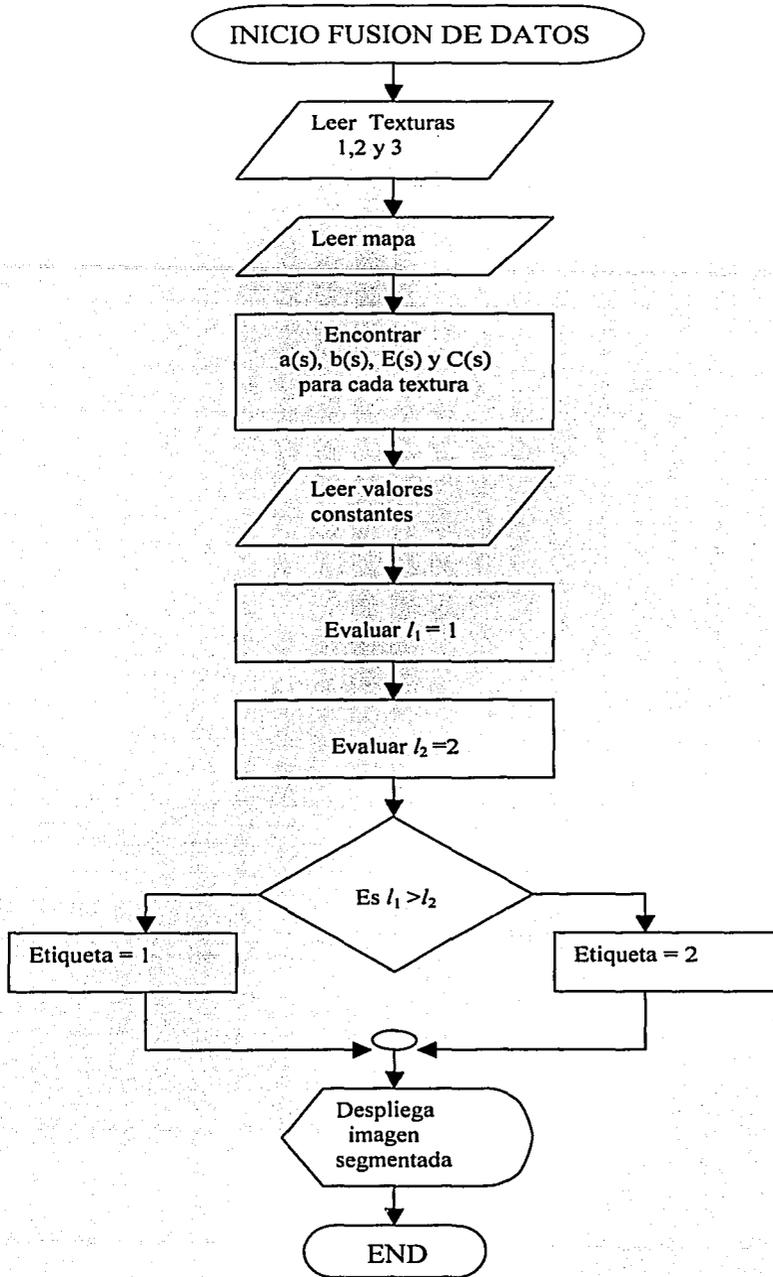


Como se muestra en el diagrama de bloques, nuestro primer bloque es la generación de texturas, la cual genera una imagen de dos clases, clase urbana y clase no urbana, de las once imágenes generadas por el bloque 1, se seleccionan tres, las cuales entran al proceso de fusión de datos, además de éstas, entra otra imagen llamada mapa, generando así otra imagen en dos clases, urbano y no urbano.

DIAGRAMAS DE FLUJO







El diagrama de flujo de texturas es solo para un descriptor de Haralick, aunque también se puede generar para varios descriptores dentro del mismo programa. Otra manera de generar dichas texturas es obteniendo primero, la matriz de co-ocurrencia y a esta aplicarle los descriptores de Haralick.

Una vez obtenidas las texturas de interés, estas se procesan en el siguiente bloque que es precisamente el diagrama de flujo de la fusión de datos. A este proceso entran tres imágenes aparte de un mapa rudimentario de la imagen original.

3. – ANÁLISIS DEL SOFTWARE

El software generado para la realización de esta tesis consiste en dos programas separados, uno para las texturas y otro para la fusión de datos, este software puede ser modificado de tal forma que pueda ejecutarse en menos tiempo, esto con el uso de el lenguaje C, o algún otro tipo de lenguaje, pero la estructura de programación será siempre básicamente la misma.

El tamaño de espacio en disco es de 10.000Kb a 15.000Kb para cada textura, dependiendo claro, de cada una. Para el programa de fusión de datos se ocupan 13Kb.

Los resultados que generan los programas son buenos para integrar la información parcial a la segmentación, pero se pueden mejorar aún más, si es necesario integrar otro tipo de información, esto significa que las ventanas de análisis pueden cambiar de tamaño, así como los valores de desplazamiento y los niveles de cuantización. Los programas son flexibles para este propósito.