



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**FACULTAD DE INGENIERÍA**

**TESIS**

**“REGULACIÓN ESTOCÁSTICA Y  
PROCESOS ADAPTATIVOS EN DATOS DEL  
SATÉLITE ENVISAT”**

**QUE PARA OBTENER EL TÍTULO DE  
“INGENIERO EN TELECOMUNICACIONES”**

**PRESENTA:**

Granda Sandoval Aldo Isbal

**DIRECTOR DE TESIS:**

Dr. Miguel Moctezuma Flores

Ciudad Universitaria, Junio 2015



## Agradecimientos

### **A mis padres y hermana**

Gracias por su apoyo incondicional. No tengo palabras para agradecer todo lo que han hecho por mí. Muchas gracias.

### **Al Dr. Miguel Moctezuma**

Gracias por su apoyo para lograr este trabajo.

### **A profesores y compañeros**

Por su apoyo, conocimiento y experiencias.

### **Amigos**

La amistad es algo que se siente y no se mide. Gracias amigos.

# Índice

<b>1. Introducción.....</b>	<b>1</b>
<b>1.1 Objetivo.....</b>	<b>2</b>
<b>1.2 Fundamentos .....</b>	<b>2</b>
<b>1.3 Definición del problema.....</b>	<b>3</b>
<b>1.4 Aportación del trabajo de Tesis.....</b>	<b>3</b>
<b>2. Satélite Ambiental, ENVISAT, Environmental Satellite.....</b>	<b>4</b>
<b>2.1 Generalidades del Sistema ENVISAT .....</b>	<b>5</b>
2.1.1 Panorama general del Sistema ENVISAT.....	5
2.1.2 Datos técnicos del Satélite ENVISAT .....	6
2.1.3 Instrumentos del Satélite ENVISAT .....	6
<b>2.2 Datos del Radar de Apertura Sintética, SAR, Synthetic Aperture Radar .....</b>	<b>9</b>
2.2.1 Principio del SAR .....	10
<b>2.3 El Instrumento de Radar de Apertura Sintética Avanzado ASAR, Advanced Synthetic Aperture Radar .....</b>	<b>14</b>
2.3.1 Características generales del ASAR.....	14
2.3.2 Modos de operación del ASAR.....	16
<b>3. Campos Aleatorios de Markov, MRF, Markov Random Fields .....</b>	<b>18</b>
<b>3.1 Fundamentos de la teoría de la probabilidad y variable aleatoria .....</b>	<b>19</b>
<b>3.2 Etiquetado visual .....</b>	<b>22</b>
3.2.1 Sitios y etiquetas .....	23
3.2.2 Asignación de etiquetas .....	24
3.2.3 Vecindad.....	25
3.2.4 Cliques.....	29
<b>3.3 Campos Aleatorios de Markov, MRF, Markov Random Fields.....</b>	<b>31</b>
3.3.1 Definición de los MRF .....	32
3.3.2 Campos Aleatorios de Gibbs, GRF, Gibbs Random Fields.....	32
3.3.3 Equivalencia Markov-Gibbs.....	33
3.3.4 Uso de los modelos MRF.....	35
3.3.4.1 Modelo de Ising .....	35
3.3.4.2 Modelo de Potts.....	37
3.3.4.3 La suavidad a priori.....	38
3.3.5 Etiquetado de Bayes de los MRF .....	40
3.3.5.1 Estimación de Bayes .....	41
3.3.5.2 Etiquetado MAP-MRF .....	42
3.3.5.3 Distribución a priori .....	43
3.3.5.4 Distribución a posteriori.....	44

<b>3.4 Ejemplos de Modelos de Markov</b> .....	<b>44</b>
3.4.1 Modelos MRF de bajo nivel .....	45
3.4.1.1 Modelos de Observación.....	45
3.4.1.2 Restauración de Imagen .....	46
3.4.1.3 Segmentación de Textura .....	48
3.4.1.4 Detección de Borde.....	50
3.4.1.5 Análisis de movimiento .....	51
3.4.2 Modelos MRF de alto nivel .....	52
3.4.2.1 Correspondencia bajo condiciones relacionadas .....	52
3.4.2.2 Correspondencia de objetos múltiples .....	55
<b>4. Optimización estocástica</b> .....	<b>58</b>
<b>4.1 Fundamentos</b> .....	<b>59</b>
<b>4.2 Segmentación con Markov</b> .....	<b>60</b>
<b>4.3 Esquema de Metropolis</b> .....	<b>63</b>
<b>4.4 Recocido Simulado</b> .....	<b>64</b>
<b>4.5 Optimización</b> .....	<b>66</b>
<b>5. Esquema de regularización adaptativa</b> .....	<b>69</b>
<b>5.2 Esquema adaptativo sobre el Recocido Simulado</b> .....	<b>71</b>
<b>6. Resultados</b> .....	<b>72</b>
<b>6.1 Resultados utilizando el esquema convencional</b> .....	<b>73</b>
<b>6.2 Imágenes con el esquema adaptativo de Recocido Simulado</b> .....	<b>79</b>
<b>7. Conclusiones</b> .....	<b>84</b>
<b>Referencias</b> .....	<b>87</b>

# 1. Introducción

En este capítulo se presentan el objetivo del trabajo de tesis, así como la definición del problema y la aportación que tiene este trabajo de tesis.

## 1.1 Objetivo

Desarrollo de un esquema para la segmentación de imágenes de radar del Radar de Apertura Sintética Avanzado, ASAR, Advanced Synthetic Aperture Radar, que estuvo a bordo del satélite ambiental ENVISAT, Environmental Satellite, utilizando el método probabilístico de los Campos Aleatorios de Markov con optimización estocástica aplicando el algoritmo de Recocido Simulado para comprobar que se tiene una mejor textura de la imagen.

## 1.2 Fundamentos

Con el objetivo de estudiar y obtener datos sobre la atmósfera y la superficie terrestre, la Agencia Espacial Europea, ESA, European Space Agency, lanzó el satélite ENVISAT. Algunos datos recabados por el satélite fueron imágenes generadas por las ondas de radio mediante el Radar de Apertura Sintética Avanzado, ASAR.

Las imágenes obtenidas del ASAR representan algunas regiones de la superficie terrestre, las cuales son de interés para este trabajo; ya que las imágenes por su naturaleza presentan ruido, y para diferenciar regiones dentro de las imágenes, estas se pueden procesar mediante métodos probabilísticos.

Dentro de la probabilidad y estadística se encuentra la teoría sobre los Campos Aleatorios de Markov, MRF, Markov Random Fields, en los cuales se define un proceso de etiquetado que se utiliza para diferenciar datos de una imagen (en este caso del ASAR), es decir, se define un conjunto de valores finito que se le asigna a cada pixel de la imagen. Sin embargo, el proceso de etiquetado de los MRF hace que el proceso computacional sea tardado y complicado, pero se resuelve con la optimización de los MRF de acuerdo con el teorema de Hammersley-Clifford (se define la equivalencia entre los Campos Aleatorios de Markov y los Campos Aleatorios de Gibbs) y el algoritmo de Recocido Simulado (Li, 1995).

La imagen a procesar será interpretada como varios conjuntos de pixeles llamados vecindades y subconjuntos llamados cliques (pixeles con el mismo valor dentro de una vecindad), que con la optimización de los MRF darán una solución a un mínimo global. Con el proceso anterior se aproximan los datos de la imagen original con la respuesta final del algoritmo de recocido simulado, teniendo como resultado una imagen con regiones diferenciadas y un contorno más detallado de las texturas.

### **1.3 Definición del problema**

Se hace el procesamiento de imágenes obtenidas por el SAR del derrame petrolero ocasionado por la compañía British Petroleum que ocurrió el 20 de abril de 2010 en el Golfo de México, el cual provocó la lamentable muerte de 11 trabajadores.

Debido al gran impacto ambiental, la ESA tomó imágenes de radar con el satélite ENVISAT. Estos datos se procesan para que con ayuda de los MRF y su optimización estocástica, se pueda diferenciar entre lo que fue el petróleo crudo derramado y el mar abierto para hacer una aproximación del área del derrame.

El procesamiento de las imágenes será mediante un proceso estocástico analizando cada pixel y comparándolo con sus vecinos. Se etiqueta cada pixel con una clase o etiqueta (nivel de gris, ya que la imagen está en escala de grises) y se analizará por medio de los MRF con la optimización del Recocido Simulado, en el cual se aplica el modelo de Metropolis y se llega a un mínimo global para contrarrestar el gasto computacional en el momento de procesar cada imagen.

Al final del procesado se podrá tener un tono dentro de la escala de grises que representará el petróleo crudo derramado y otro para el mar abierto.

### **1.4 Aportación del trabajo de Tesis**

La aportación de este trabajo de tesis es la corroboración del método de los MRF en la segmentación de imágenes mediante la modificación del algoritmo de recocido simulado para optimizar el proceso estocástico aplicado a imágenes obtenidas del satélite ENVISAT.

Se obtienen datos más precisos sobre el derrame petrolero, donde el tono de gris del petróleo en toda la imagen será único y se tendrá un dato más aproximado del área del derrame.

## 2. Satélite Ambiental, ENVISAT, Environmental Satellite

Para el estudio del medio ambiente, se han desarrollado diversos experimentos y sistemas que ayudan a comprender mejor el comportamiento de la naturaleza; es por esto que científicos de la Agencia Espacial Europea desarrollaron el satélite ENVISAT, *Environmental Satellite*, o Satélite Ambiental. En este capítulo, se describen características generales del satélite, pero con un enfoque sobre los datos ASAR, Advanced Synthetic Aperture Radar.

## **2.1 Generalidades del Sistema ENVISAT**

El Satélite Ambiental ENVISAT, Environmental Satellite, fue un satélite lanzado por la Unión Espacial Europea, ESA, European Space Agency. Contó con diez instrumentos para recopilar información sobre la atmósfera, la tierra, los océanos y el hielo. La duración de la misión de este satélite estaba prevista para ser de por lo menos cinco años en órbita polar después de su fecha de lanzamiento, sin embargo duró en órbita más de 10 años.

### **2.1.1 Panorama general del Sistema ENVISAT**

El 1 de marzo de 2002 el satélite ENVISAT fue lanzado por la Agencia Espacial Europea como el más grande satélite europeo de observación de la superficie de la Tierra, proporcionaba mediciones de la atmósfera, el océano, la tierra, el hielo y datos sobre el cambio climático (“European Space Agency”, ESA, 2015). Sin embargo, el 8 de abril de 2012 se perdió la comunicación entre la estación terrena y el satélite, por lo que el 9 de mayo de 2012 fue dada por terminada la misión del ENVISAT después de varios intentos para reestablecer la comunicación.

El satélite ENVISAT contó con 10 instrumentos ópticos y radares, donde el principal instrumento fue el Radar de Apertura Sintética Avanzada, ASAR, Advanced Synthetic Aperture Radar, utilizados para diferentes mediciones de la Tierra (“ESA, 2015”).

Los dos objetivos principales por los cuales el ENVISAT fue lanzado fueron:

1. Proporcionar a Europa una capacidad mejorada para la detección remota de la Tierra desde el Espacio.
2. El Sistema ENVISAT fue lanzado para el seguimiento de las observaciones hechas por el Satélite de Detección Remota Europea ERS, European Remote Sensing Satellite. Dicho satélite fue lanzado en 1991 para la observación de la Tierra portando instrumentos para la medición de la temperatura de la superficie terrestre y la nieve en el océano; además de tener un radar de apertura sintética (“ESA”, 2015).

## 2.1.2 Datos técnicos del Satélite ENVISAT

El ENVISAT fue lanzado por la empresa francesa Arianespace en un cohete Ariane5; se posicionó en una órbita polar sincrónica con el Sol, también conocida como heliosincrónica, con un periodo nominal de 100 minutos de repetición cada 35 días. Inclinado  $98.55^\circ$  con respecto al Ecuador a una altitud media de 800 km, con una desviación máxima de +/- 1 km de la pista de tierra y de +/- 5 minutos en el cruce del Ecuador ("ESA", 2015).

Las dimensiones del satélite fueron las siguientes: longitud 25 m; ancho 10 m; alto 7 m y; tuvo una masa total de 8 211 kg. Con respecto a la energía utilizada se presenta la Tabla 2.1.

Tabla 2.1. Presupuesto de energía

	Potencia media [W]	
	Luz del sol	Eclipse
Carga útil	1 841	1 886
Módulo de carga	860	527
Módulo de servicio	859	684
Carga total	3 560	3 097
Capacidad del sistema	3 847	3291

Dependiendo del instrumento que fuera utilizado se necesitaron diferentes frecuencias en operación.

## 2.1.3 Instrumentos del Satélite ENVISAT

Se desarrollaron diferentes instrumentos para estudiar específicamente cada fenómeno; ya sea sobre la atmósfera, el océano, el medio ambiente o cambio climático. El satélite contó con diez instrumentos para medición y detección remota, donde uno de los más importantes fue el ASAR. La Fig. 2.1 muestra un esquema de lo que fue el satélite ENVISAT.

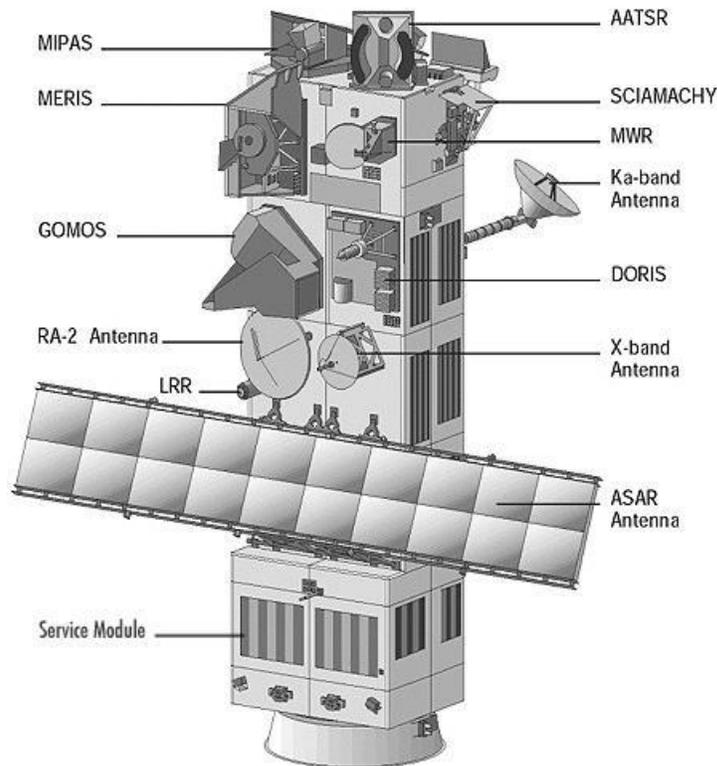


Fig. 2.1 Satélite ENVISAT marcando sus diferentes instrumentos utilizados durante su misión (“ESA”, 2015).

Los diez instrumentos utilizados en la misión del ENVISAT se numeran y se describen brevemente como:

1. **ASAR** (Advanced Synthetic Aperture Radar): Radar de Apertura Sintética Avanzada, operó en la banda C (4 a 8 GHz). ASAR asegura la continuidad con el modo de imagen (del SAR, Synthetic Aperture Radar) y el modo de onda AMI (Instrumento de Microondas Activo) del ERS, European Remote Sensing Satellite, y ERS-2, European Remote Sensing Satellite 2, (“The ERS-1 and ERS-2”, 2015).
2. **MERIS** (Medium-Spectral Resolution): Espectrómetro de Imágenes de Resolución Media, MERIS, fue un instrumento de formación de imágenes principalmente para el estudio del mar y la superficie de la tierra. El MERIS contó con cinco módulos ópticos los cuales trabajaron en el rango de frecuencias de 390 nm a 1 040 nm (“Center for Remote Imaging, Sensing & Processing”, 2015).
3. **AATSR** (Advanced Along Track Scanning Radiomete): El Radiómetro Explorador de Barrido Avanzado fue utilizado para la medición precisa de la

superficie del mar (“What is AATSR”, 2015). Las frecuencias utilizadas para la mediciones de la radiación reflejada y emitida fueron: 0.55  $\mu\text{m}$ , 0.66  $\mu\text{m}$ , 0.87  $\mu\text{m}$ , 1.6  $\mu\text{m}$ , 3.7  $\mu\text{m}$ , 11  $\mu\text{m}$  y 12  $\mu\text{m}$  (“ESA”, 2015).

4. **RA-2** (Radar Altimeter 2): Radar Altimetro 2 (RA-2) fue un instrumento para determinar el retraso de dos vías del eco de radar de la superficie de la Tierra en menos de un nanosegundo; además medía la alimentación y la forma de los impulsos de radar reflejados. Las frecuencias en las que trabajaba el Radar fueron de 13.575 GHz (banda Ku) y en la frecuencia de 3.2 GHz (banda S) como un canal secundario para estimar el retardo del rango ionosférico (“ESA”, 2015).
5. **MWR** (Microwave Radiometer): El Radiómetro de Microondas fue un radiómetro de dos canales operando en la frecuencia de 23.8 GHz y 36.5 GHz, donde el objetivo de este instrumento fue la medición de la humedad de la atmósfera (“ESA”, 2015).
6. **GOMOS** (Global Ozone Monitoring by Occultation of Stars): Monitoreo de Ozono Mundial por Ocultación de Estrellas, fue un espectrómetro de resolución media que trabajó en el rango de frecuencias de ultravioleta y luz visible en el rango de 250 nm a 950 nm, con la intención de observar la capa de ozono y de otros elementos, así como la temperatura y el vapor de agua. (“ESA”, 2015).
7. **MIPAS** (The Michelson Interferometer for Passive Atmospheric Sounding): El Interferómetro Michelson para el Sondeo Atmosférico Pasivo, MIPAS, fue un espectrómetro de transformada de Fourier para la detección de espectros de emisión en los límites de la atmósfera media y superior, operando en el rango de 4.15  $\mu\text{m}$  a 14.6  $\mu\text{m}$  (“ESA”, 2015).
8. **SCIAMACHY** (Scanning Imaging Absorption Spectrometer for Atmospheric Cartography): El Espectrómetro de Absorción de Imágenes de Exploración para la Cartografía Atmosférica fue un sensor atmosférico que consistía de un espectrómetro de imágenes. La misión principal de este instrumento fue llevar a cabo las mediciones globales de gases traza (componentes menores de la atmósfera) en la troposfera y en la estratosfera (“ESA”, 2015; “The World Data Center for Remote Sensing of the Atmosphere”, 2015). Las mediciones de los gases traza se grababan en el instrumento por la transmisión, la retrodispersión y la radiación reflejada en la atmósfera en las bandas desde ultravioleta al infrarrojo cercano que son del rango de 240 nm a 2 400 nm (“ESA”, 2015).

9. **DORIS** (Doppler Orbitography and Radio-positioning Integrated by Satellite): La Orbitografía Doppler y Radio-Posicionamiento fue un sistema de seguimiento por microondas, utilizando el fenómeno del desplazamiento de Doppler de una señal de radiofrecuencia transmitida desde una estación terrena y recibida en el satélite. Midió los desplazamientos usando una frecuencia de 2.03 625 GHz (banda S) y una de 401.25 MHz (VHS) para la corrección ionosférica del retardo de propagación. Su principal aplicación es el apoyo en la determinación de la órbita y compartir sus aplicaciones, también permite la estimación de la cantidad de electrones libres en la ionosfera (“ESA”, 2015).
  
10. **LRR** (Laser Retro-Reflector): El Láser Retro-Reflector fue un dispositivo pasivo con un arreglo de espejos que proporciona un objetivo para el láser de seguimiento mediante pulsos de alta potencia, ya que este dispositivo se utiliza como un reflector por las estaciones en tierra. Sus funciones son el apoyo al satélite que oscila y calibrar la altitud del instrumento RA-2 (“Radar Altimetry Tutorial”, 2015)

## **2.2 Datos del Radar de Apertura Sintética, SAR, Synthetic Aperture Radar**

El Radar de Apertura Sintética (SAR), es el antecesor del Radar de Apertura Sintética Avanzado (ASAR) que llevaba a bordo el satélite ENVISAT. El SAR fue diseñado para obtener imágenes de radar más claras y con una mayor área de cobertura.

El primer SAR llevado como carga útil a bordo de un satélite fue en junio de 1978 en el satélite SeaSat, Seafaring Satellite, el cual dio las primeras imágenes de radar de la superficie de la Tierra. Este sistema se caracterizó por obtener imágenes en cualquier momento, sin depender de las condiciones climáticas. Se tuvo una resolución geométrica independiente del sensor de altitud o longitud de onda y trabajó dentro de la banda C para disminuir la atenuación de la señal, ya que alrededor de esta banda existe una transmisión exitosa, casi del 100% como se muestra en la Fig. 2.2 (Curlander, 1991).

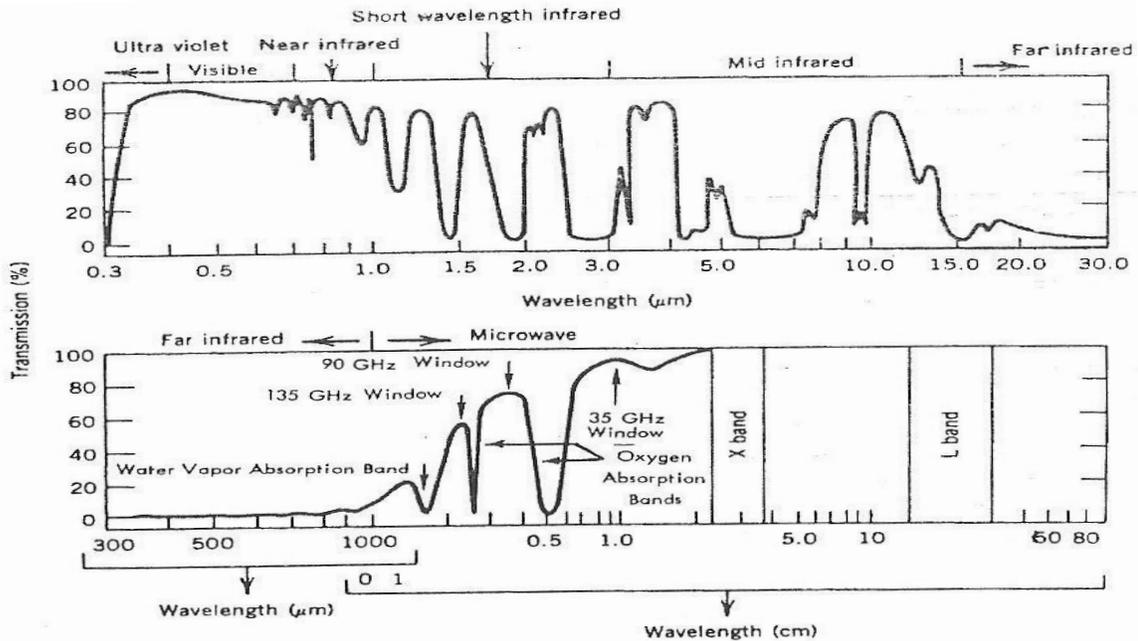


Fig. 2.2. Porcentaje de transmisión a través de la atmósfera para el rango de microondas del espectro electromagnético (Curlander, 1991, p. 5.).

El SAR es un sensor activo que transmite una señal con medidas y mapeo de energía reflejada en la banda de frecuencias de las microondas que dan información acerca de la superficie de la Tierra.

### 2.2.1 Principio del SAR

El sistema SAR se puede dividir en cinco principales grupos (ver Fig. 2.3), los cuales se describen brevemente (Skolnik, 1988):

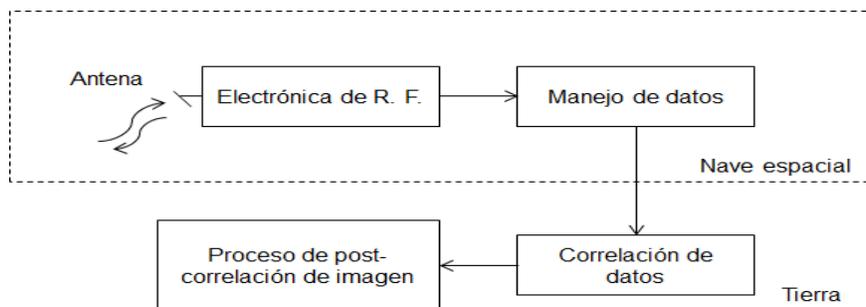


Fig. 2.3. Diagrama de bloques del sistema SAR (Skolnik, 1988, p. 326).

- La antena (Antenna): es la interfaz entre la electrónica de Radio Frecuencia (R. F.) y el espacio libre.
- Electrónica de R. F. (R. F. Electronics): consiste de los subsistemas de transmisión, el cual generan señales de radio frecuencia; y del subsistema receptor que detecta la energía devuelta. Este sistema convierte la energía devuelta para ser procesada por el elemento de manejo de datos.
- Elemento de manejo de datos (Data Handling): la salida de este elemento puede ser una cinta digital con una grabación digital del regreso de los ecos, o una película óptica con grabaciones analógicas de los ecos regresados.
- Correlación de datos (Data Correlation): este elemento transforma la señal grabada en una imagen, la cual es una representación de dos dimensiones de la superficie reflejada.
- Proceso de post-correlación de imagen (Image Post-Correlation Processing): se realiza por geometría y calibración radiométrica, además del registro con otros conjuntos de datos adquiridos con un similar o diferente sensor de imagen.

En la Fig. 2.4 se muestra la geometría simplificada de un radar de apertura real de visión lateral SLAR, Side-Looking Airbone Radar, que tiene la misma geometría que el SAR. El SLAR se lleva sobre una plataforma (aeronave o satélite) moviéndose a una velocidad  $V_s$  en línea recta con altitud constante. Se asume que el haz del radar es dirigido perpendicularmente a la trayectoria del vuelo del vehículo y hacía abajo sobre una superficie plana. La velocidad entre la plataforma y la tierra es  $V_{st}$ . Para esta geometría, el ángulo señalado  $\gamma$ , relativo a la vertical, es el mismo que el ángulo incidente  $\eta$ , que es el ángulo entre el haz del radar y la normal de la superficie de la tierra en el punto particular de interés. El radar transmite pulsos de energía electromagnética y los ecos regresados son muestreados para el proceso de la señal coherente o semejante en el tiempo (Curlander, 1991).

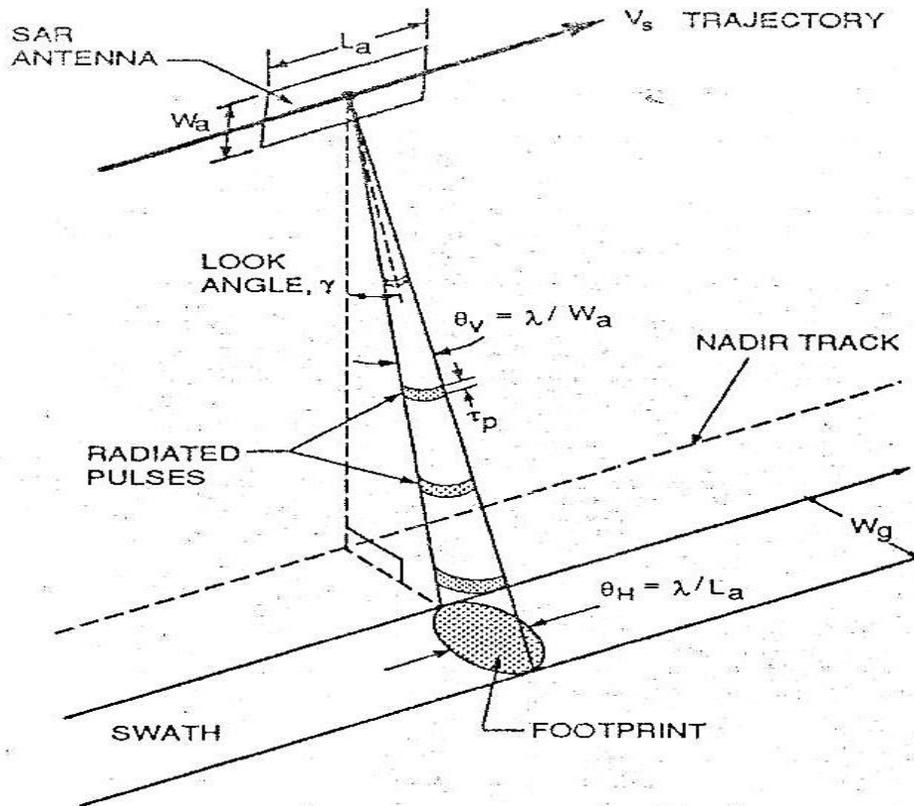


Fig. 2.4. Geometría simplificada de un radar de apertura real de visión lateral (Curlander, 1991, p.14).

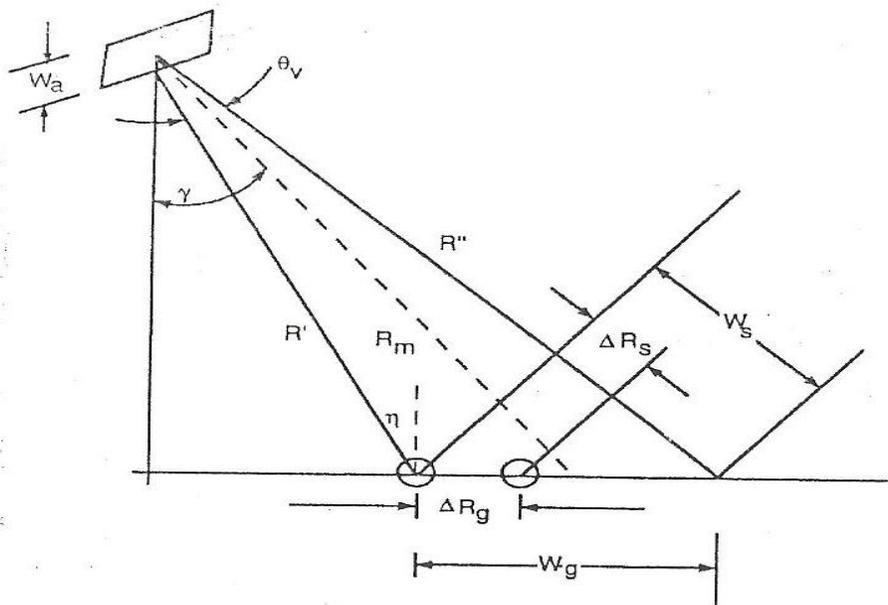


Fig. 2.5 Geometría del radar que muestra la franja de tierra,  $w_s$  en el ancho de haz del radar  $\theta_v$  (Curlander, 1991, p.15).

En la Fig. 2.5 se observa que el rango alcanzado  $W_g$  del haz del radar (es decir, el ancho de la franja de la superficie plana) se puede establecer con la altura de la antena  $W_a$ , que determina el ancho del haz vertical  $\theta_v = \lambda/W_a$ . Si  $R_m$  es el alcance (inclinación) del radar a media franja, entonces:

$$W_g \approx \frac{\lambda R_m}{W_a \cos \eta} \quad (2.1)$$

La resolución del radar en el rango (en tierra o superficie plana) en la Fig. 2.5 se define como la separación de rango mínimo de dos puntos que pueden ser distinguidos por separado por el sistema. Si la medida de tiempo del pulso del radar es  $\tau_p$ , la separación mínima de dos puntos resolubles es entonces:

$$\Delta R_g = \frac{\Delta R_s}{\sin \eta} = \frac{c \tau_p}{2 \sin \eta} \quad (2.2)$$

donde  $\Delta R_s$  es la resolución en el rango inclinado y  $c$  es la velocidad de la luz.

Una técnica de compresión de pulso se emplea para lograr tanto una resolución alta (con un pulso largo) y una alta relación señal a ruido, SNR. Con el apropiado proceso del pulso recibido, la resolución del rango obtenido es:

$$\delta R_g = \frac{c}{2 B_g \sin \eta} \quad (2.3)$$

donde  $B_g$  es el ancho de banda de frecuencia del pulso transmitido.

La resolución del rango del sistema de radar se puede determinar por el tipo de pulso transmitido y la trayectoria que éste toma de regreso para procesarlo; ya que todos los sistemas de radar convencionales, SLAR o SAR, resuelven objetivos en la dimensión del rango. Esto es la resolución del objetivo en la dimensión paralela de la línea de la plataforma del vuelo, es decir, el acimut o la dimensión a lo largo del rastro, que distingue un Radar de Apertura Sintética de otros sistemas de radar.

## **2.3 El Instrumento de Radar de Apertura Sintética Avanzado ASAR, Advanced Synthetic Aperture Radar**

El instrumento ASAR tiene varias similitudes con el SAR, ya que también trabaja en el banda de frecuencias C (5.331GHz) con cinco modos de polarización: VV, HH, VV/HH, HV/HH o VH/VV; además que asegura la continuidad del modo de imagen SAR, así como el modo de onda AMI del ERS y ERS-2. También, incorpora la capacidad de dirigir el haz de imagen en diferentes posiciones e hileras (“ESA”, 2015).

Con las mejoras del ASAR se tiene mayor capacidad de cobertura, gama de ángulos de incidencias, polarización y modos de operación. Las mejoras son proporcionadas debido a los cambios en el diseño de los instrumentos como los son: un arreglo de antenas activas totalmente equipada con módulos de transmisión/recepción distribuidos, los cuales proporcionan distintos haz de luz de transmisión y recepción; generación de una forma de onda digital para la generación del pulso *chirp*, es decir, se genera un pulso que tiene una variación de incremento o decremento con el tiempo; un esquema de cuantificación adaptativo de bloque y un modo de operación ScanSAR (“Radar de Apertura Sintética con una cobertura de franja, en el rango de inclinación, que es más amplio que los SAR convencionales” (Monti, Andrea; Prati, Claudio, 2015)) por barrido de haz en elevación (“ESA”, 2015).

### **2.3.1 Características generales del ASAR**

El ASAR es un radar de red en fase con módulos de T/R, Transmisión/Recepción, organizados a través de la antena, de tal manera que, mediante los ajustes de fase del módulo individual y la ganancia, los haz de transmisión y recepción pueden ser configurados y dirigidos (“ESA”, 2015).

Este instrumento se compone de dos principales grupos funcionales: el subconjunto de antena ASA, Antenna Subassembly, y el subconjunto de la electrónica central CESA, Central Electronics Subassembly. La antena activa tiene 20 piezas con 16 módulos de T/R cada uno. El ASAR se regula por su control de instrumentos ICE, Instrument Control Equipment, además de que mantiene y gestiona la base de datos de los parámetros en operación (“ESA”, 2015).

El ASAR consiste en un arreglo activo en fase coherente de SAR, es decir, elementos de transmisión y recepción distribuidos sobre el eje "Y" que se montan sobre el eje a lo largo de la antena alineada con la dirección del vuelo del satélite. El objetivo del sistema SAR es proporcionar una representación de dos dimensiones de la reflectividad de la escena en alta resolución, con ejes definidos en el rango y en la dirección del acimut. Ya que el contenido de las imágenes del patrón de haz de una tira de tierra hacia el lado derecho de la trayectoria del vuelo tiene un contenido principalmente ilimitado en la dirección de movimiento (dirección del acimut), pero delimitado en la dirección ortogonal (dirección de alcance) por el ancho de haz de elevación de la antena ("ESA", 2015).

La resolución de alcance de un radar se limita principalmente por el ancho de banda del pulso transmitido; por lo que para preservar la resolución radiométrica el ASAR utiliza una técnica que consiste en generar un impulso largo con una modulación en frecuencia lineal o chirp. La duración del pulso se define para ser coherente con la relación señal a ruido, mientras que el ancho de banda del chirp está definido por la resolución de rango requerido ("ESA", 2015).

Para un objetivo determinado, debido a que la distancia entre el objetivo y el radar cambia continuamente (distancia oblicua), la fase de los cambios de la señal reflejada se basa en la geometría de observación. Debido a que esta ley es determinística, es posible que la fase de las señales de regreso con respecto a otra de un efecto neto, que es equivalente a que todas las señales se hayan recibido simultáneamente por una antena de igual longitud a la longitud de la trayectoria sobre cada señal del radar que se van recogiendo. De esta manera, la antena sintetizada puede pasar como un número de elementos radiantes independientes (apertura real) donde la separación está establecida por la Frecuencia de Repetición de Pulsos PRF, Pulse Repetition Frequency, y la velocidad de la plataforma ("ESA", 2015).

El cambio de fase con respecto al tiempo es la frecuencia angular Doppler. La resolución de acimut está determinada por el ancho de banda Doppler de la señal recibida. Además, en el ASAR, el ancho de banda de los pulsos de los retornos del objetivo, en acimut, debe ser mayor que el ancho de banda de acimut para satisfacer el criterio de Nyquist ("ESA", 2015).

### 2.3.2 Modos de operación del ASAR

La “ESA” (2015) define cinco modos exclusivos de operación del ASAR que son los siguientes:

- **Modo de Monitoreo Mundial** (MM, Global Monitoring Mode): proporciona imágenes de baja resolución sobre una franja de 405 km en polarización HH o VV. En este modo se tiene una baja velocidad de datos debido a una proporción ligeramente reducida a lo largo de la pista de radio y el uso del filtrado digital.
- **Modo de Onda** (WM, Wave Mode): en este modo se miden los cambios en la retrodispersión de la superficie de los océanos debido a la acción de las olas. Se generan viñetas donde el tamaño mínimo es de 5 km por 5 km.
- **Modo de Imagen** (IM, Image Mode): en este modo se generan productos de alta resolución espacial (30 m).
- **Modo de Polarización Alterna** (AP, Alternating Polarization Mode): usando este modo se tienen productos de alta resolución en cualquier franja como un IM pero con el cambio de polarización de subapertura a subapertura dentro de la apertura sintética. El resultado se encuentra en dos imágenes de la misma escena en diferentes combinaciones de polarización (HH/VV o HH/HV o VV/VH) con aproximadamente 30 m de resolución.
- **Modo de Franja Ancha** (WS, Wide Swath Mode): en el modo WS se utiliza el ScanSAR produciendo imágenes de una franja más amplia (405 km) con resolución media (150 m) en polarización HH o VV.

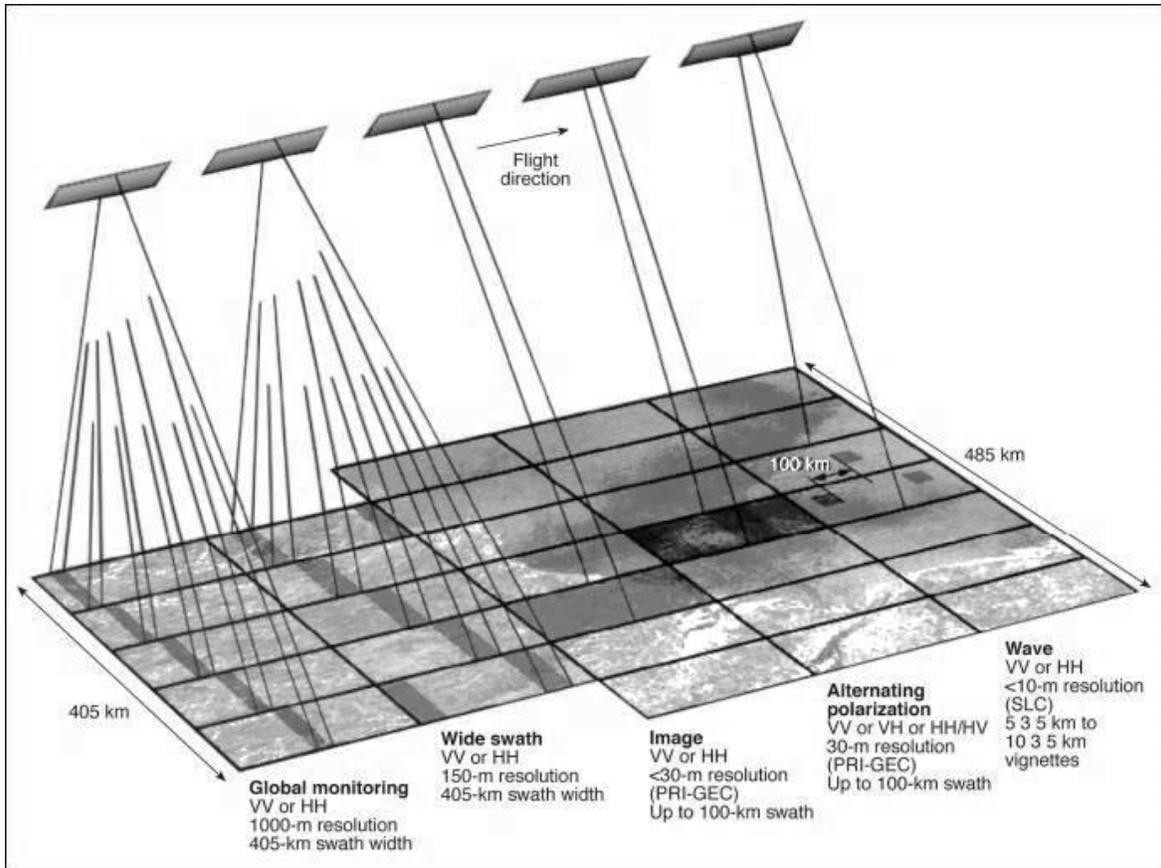


Figura 2.6. Ilustración de los modos de operación del ASAR (“Sharing Earth Observation Resource”, 2015).

# **3. Campos Aleatorios de Markov, MRF, Markov Random Fields**

Las características que componen a los Campo Aleatorios de Markov hacen posible su uso en el procesamiento de imágenes, ya que son procesos estocásticos donde una variable aleatoria de un espacio muestral tiene una probabilidad de ser el valor de un pixel de alguna imagen mediante el esquema de Bayes. Esto es posible debido a las propiedades que tienen las imágenes digitales como: intensidad, color, textura y tamaño de la imagen en pixeles. En este capítulo se abarca el estudio de manera general de los Campos Aleatorios de Markov.

### 3.1 Fundamentos de la teoría de la probabilidad y variable aleatoria

“En la probabilidad, un experimento es cualquier proceso de ensayo y observación. Un experimento cuyo resultado es incierto antes de que se lleve a cabo se llama experimento aleatorio” (Ibe, 2009, p. 1). Al realizar un experimento bajo las mismas condiciones no siempre produce el mismo resultado. La variabilidad de los resultados se estudia mediante la teoría de la probabilidad y variables aleatorias. Ya que los resultados del experimento pueden ser descritos por números, las variables aleatorias pueden considerarse como los posibles resultados de dicho experimento.

“Una variable aleatoria  $X$  es una función de valor real de los elementos de un espacio muestral  $S$ ” (Larson, 1978). Esta función se puede tomar como una regla para asignar un valor numérico a cada elemento de un conjunto; además, se tiene que el dominio de la variable aleatoria es un espacio muestral y su rango es el conjunto de los números reales. Las variables aleatorias pueden ser: discretas, continuas o mixtas.

Un *espacio muestral* es el conjunto de todos los posibles resultados diferentes cuando se realiza un experimento aleatorio y cada uno de estos elementos del espacio se considera como un punto del espacio muestral. Por otro parte, un *suceso* o *evento* es un subconjunto del espacio muestral, ya que es la ocurrencia de cualquier resultado prescrito o de cualquiera de las posibles salidas o resultados del experimento.

Un proceso estocástico es una sucesión de variables aleatorias  $X_1, X_2, X_3, \dots, X_n$ . Estas variables aleatorias pueden ser un número contable, donde el proceso estocástico recibe el nombre de *proceso discreto de parámetros*, por ejemplo la sucesión diaria de temperaturas. Mientras que, si el proceso consiste de un número no contable de variables aleatorias se conoce como *proceso continuo de parámetros*, como el caso del monitoreo continuo de una máquina (Estrada, 2009).

Se define la función de distribución y densidad que son dos características asociadas a las variables aleatorias:

#### **Función de Distribución**

Dada una variable aleatoria  $X$ , la función:

$$F_X(x) = P\{X \leq x\} \quad (3.1)$$

es llamada *función de distribución* de  $X$ , definida para cualquier número  $x \in (-\infty, \infty)$  donde  $F(-\infty) = 0$  y  $F(\infty) = 1$

### Función de Densidad

Dada una variable aleatoria  $X$ , se puede derivar la distribución  $F_X(x)$ :

$$f(x) = \frac{dF(x)}{dx} \quad (3.2)$$

y se obtiene *la función de densidad* de  $X$ .

Debido a que  $F(x)$  no cambia en su orden se tiene que  $f(x) \geq 0$  y:

$$\int_{-\infty}^{\infty} f(x)dx = F(\infty) - F(-\infty) = 1 \quad (3.3)$$

Ya que  $F(x)$  puede no tener una derivada para cada  $x$ , se puede asociar una función de densidad a todas las variables aleatorias.

De acuerdo con las definiciones anteriores, que tienen su base en la Teoría de Bayes, Obregon (1975) define los siguientes postulados:

#### Probabilidad condicional

Sean  $A$  y  $B$  dos sucesos cualesquiera asociados con un espacio muestral, con la única condición de que  $P(B) \neq 0$

Se define la probabilidad condicional de  $A$  dado  $B$ , que se escribe como  $P(A|B)$ :

$$P(A|B) = \frac{P(AB)}{P(B)} \quad (3.4)$$

#### Sucesos independientes

Se dice que  $A$  y  $B$  son sucesos independientes si la probabilidad condicional de  $A$  dado  $B$  es igual a la probabilidad de  $A$ , esto es, si:

$$P(A|B) = P(A) \quad (3.5)$$

O que es lo mismo, si:

$$P(AB) = P(A)P(B) \quad (3.6)$$

### Probabilidad de una intersección

La ecuación (3.6) sólo es válida para sucesos independientes. Sin embargo, sean o no independientes dos sucesos A, B, siempre se puede afirmar que:

$$P(AB) = P(A|B)P(B) = P(B|A)P(A) \quad (3.7)$$

### Principio de expansión

Sea A un suceso y  $B_1, B_2, \dots, B_n$  una serie de sucesos disyuntos [que no tienen elementos en común] tales que la ocurrencia de A va necesariamente acompañada por la ocurrencia de uno de ellos, esto es:

$$A(B_1 + B_2 + \dots + B_n) = A \quad (3.8)$$

Por lo tanto se puede escribir:

$$P(A) = \sum_{i=1}^n P(AB_i) \quad (3.9)$$

Ya que los sucesos  $AB_i$  son disyuntos.

Aplicando (3.7) [probabilidad de intersección] a cada uno de los términos de la sumatoria anterior se obtiene el principio de expansión:

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i) \quad (3.10)$$

### Teorema de Bayes

Sea A un suceso y  $B_1, B_2, \dots, B_n$  una serie de suceso disyuntos tales que la ocurrencia de A va necesariamente acompañada por la ocurrencia de uno de ellos.

Se conocen las probabilidades de los  $B_i$  como también las probabilidades condicionales de A dado cada uno de los  $B_i$ . Sabiendo ahora que A ocurrió, se desea calcular la probabilidad de cada  $B_i$  dada esta información.

[Se tiene] que:

$$P(B_i|A) = \frac{P(B_iA)}{P(A)} = \frac{P(A|B_i)P(B_i)}{P(A)} \quad (3.11)$$

Expandiendo el denominador por medio [del principio de expansión] y [aplicando la probabilidad de intersección] al numerador, se tiene que:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)} \quad (3.12)$$

La expresión anterior constituye el teorema de Bayes. (pp. 58-60).

El teorema de Bayes se puede utilizar como en la siguiente situación: sean  $B_1, B_2, \dots, B_n$  causas cualesquiera que dieron lugar a la ocurrencia de  $A$ , así que, la probabilidad que se asigna a la ocurrencia de la causa  $B_i$  es  $P(B_i)$ , que se conocen como *probabilidades a priori* o *a priores*. Mientras que, posteriormente, al conocer que  $A$  ha ocurrido y se quiere conocer las probabilidades asignadas a cada una de las causas, estas probabilidades se denominan *probabilidades a posteriori* o *posteriores* que están dadas por el teorema de Bayes (Obregon, 1975).

### **Campo aleatorio**

Uno de los conceptos más importantes para comprender qué es un MRF, es el *campo aleatorio*.

Un campo aleatorio es esencialmente un proceso estocástico definido en el conjunto de nodos espaciales (o sitios). Específicamente, sea  $S = \{1, \dots, N\}$  un conjunto finito y sea  $\{X(s), s \in S\}$  una colección de variables aleatorias en el espacio muestral  $\Omega$ . Entonces el evento conjunto  $x = \{x_1, \dots, x_m\}$  es llamado una *configuración* de  $X(s)$ , que corresponde a una realización de campo aleatorio (Ibe, 2009, p. 381).

## **3.2 Etiquetado visual**

Para el estudio de los MRF el etiquetado de los píxeles es una representación natural, es decir que, para cada píxel de una imagen se le puede asignar un valor de algún conjunto de etiquetas, estas etiquetas pueden representar colores, posiciones, texturas, entre otros. Entonces, la asignación de valores a cada píxel

de alguna imagen puede plantearse como un problema de etiquetado, donde la solución es un conjunto de etiquetas que son asignadas a cada pixel

### 3.2.1 Sitios y etiquetas

Para enfrentar el problema de etiquetado se especifica un conjunto de sitios y un conjunto de etiquetas. Se define un conjunto discreto de  $N$  sitios como  $S = \{1, \dots, N\}$  donde  $1, \dots, N$  son índices.

Los sitios en una imagen pueden representar un punto o una región en el espacio Euclidiano tal como los pixeles de la imagen, alguna característica, esquinas, segmentos de línea, entre otros. En un enrejado o malla los sitios son considerados espacialmente periódicos. Para el caso de una imagen 2-D, que consiste de un enrejado rectangular con tamaño  $n \times n$ , puede ser descrita como (Li, 1995):

$$S = \{(i, j) | 1 \leq i, j \leq n\} \quad (3.13)$$

Los elementos del conjunto de sitios anterior corresponden a los lugares donde la imagen es muestreada. En el caso de sitios irregulares (sitios que no son rectangulares), las características obtenidas de la imagen son a nivel más abstracto, como la detección de esquinas y líneas.

Para los MRF normalmente los sitios se tratan como ordenados, ya que se considera una imagen de  $n \times n$  donde el pixel  $(i, j)$  puede ser referido a un solo número  $k$  el cual tomará valores en  $1, 2, \dots, m$  con  $m = n \times n$  (Li, 1995). La relación que existe entre sitios se establece por un *sistema de vecindades* (conjunto de pixeles de cierta dimensión dentro de la imagen).

Los sitios pueden sufrir eventos y estos eventos pueden ser la asignación de etiquetas; así que, un conjunto de etiquetas  $L$  puede ser categorizado como continuo o discreto. En el caso continuo, las etiquetas pueden corresponder al conjunto de los números reales  $\mathbb{R}$  o un intervalo del conjunto

$$L_c = [X_l, X_h] \subset \mathbb{R} \quad (3.14)$$

Para el conjunto de etiquetas continuas también es posible que tomen un valor de un vector o matriz, por ejemplo  $L_c = \mathbb{R}^{a \times b}$  donde  $a$  y  $b$  son las dimensiones (Li, 1995).

En el caso discreto de etiquetado, que se toma como base para este trabajo, una etiqueta toma un valor discreto de un conjunto de  $M$  etiquetas

$$L_d = \{l_1, \dots, l_M\} \quad (3.15)$$

o

$$L_d = \{1, \dots, M\} \quad (3.16)$$

Una de las características del conjunto de etiquetas es el ordenamiento de éstas. Por ejemplo, para el caso del conjunto de etiquetas continuo, el conjunto de los reales  $\mathbb{R}$ , se puede ordenar como “el menor que” o “mayor que”; mientras que para un conjunto discreto, como el  $\{0, \dots, 255\}$ , que representa los valores cuantificados de intensidades, se tiene un conjunto ordenado de intensidades  $0 < 1 < 2 < \dots < 255$ . Un conjunto desordenado se puede aplicar cuando se determinen, por ejemplo, 256 símbolos diferentes como los tipos de textura (Li, 1995).

### 3.2.2 Asignación de etiquetas

Un ejemplo del problema de etiquetado de un sitio es la asignación de una etiqueta  $f_i$  del conjunto  $L = \{borde, no - borde\}$  a un sitio  $i \in S$  donde los elementos pertenecientes a  $S$  indican los pixeles de la imagen.

El conjunto  $f = \{f_1, \dots, f_m\}$  es llamado un *etiquetado* de los sitios en  $S$  en términos de las etiquetas en  $L$ . Se puede considerar una función con dominio  $S$  e imagen  $L$ , cuando a cada sitio se le asigna una única etiqueta  $f_i = f(i)$ ; porque esto es un mapeo de  $S$  a  $L$  (Li, 1995):

$$f: S \rightarrow L \quad (3.17)$$

En el caso donde todos los sitios tienen el mismo conjunto de etiquetas  $L$ , el espacio de configuración, es decir, el conjunto de todos los posibles etiquetados, es el producto cartesiano ( $m$  es el tamaño de  $S$ ) (Li, 1995)

$$\mathbb{F} = L \times L \dots \times L = L^m \quad (3.18)$$

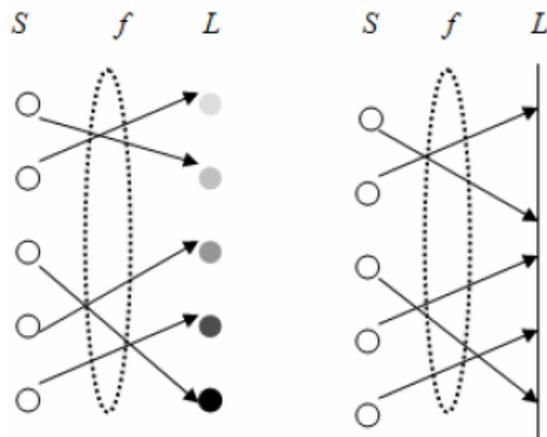


Fig. 3.1. Ejemplo de etiquetado de sitios considerado como un mapeo de un conjunto de sitios  $S$  a un conjunto de etiquetas  $L$ . A la izquierda se tiene un mapeo con un conjunto de etiquetas discretas y a la derecha se tiene un mapeo con un conjunto continuo.

### 3.2.3 Vecindad

Un *sistema de vecindad* se compone de sitios definidos dentro del conjunto  $S$ , los cuales están relacionados uno con el otro, donde cada sitio estará más influenciado por sus vecinos más cercanos que los más alejados (Chellapa, 1993; Li, 1995). Así que, un sistema de vecindad se define como:

$$N = \{N(i), i \in S\} \quad (3.19)$$

done  $N(i)$  es el conjunto de sitios que son vecinos de  $i, i \notin N(i)$ . Por otra parte, la relación que existe entre vecinos o sistema de vecindad debe cumplir las siguientes propiedades:

- 1) Un sitio no es un vecino a sí mismo:  $i \notin N_i$
- 2) La relación de vecinos debe ser mutua o simétrica:  $i \in N_{i'} \Leftrightarrow i' \in N_i$

Un sistema de vecindades se puede representar por una gráfica mediante sitios conocidos como nodos y bordes que hace referencia a dos sitios conectados si y sólo si ellos son vecinos, esto es  $i \in \Gamma(j)$ . Las gráficas pueden ser lineales o de dos dimensiones, 2D, que representan las imágenes de tamaño  $m \times n$  pixeles. Para las gráficas 2D se les nombra enrejado, ya que consisten en un conjunto de nodos que pertenecen a la imagen de tamaño  $m \times n$  pixeles. Así que, para un conjunto  $S$  o enrejado regular, el conjunto de vecinos de  $i$  o del sitio a estudiar (pixel interior) se define como el conjunto de alrededor en un radio  $r$ , es decir:

$$N_i = \{i' \in S \mid [dist(pixel_{i'}, pixel_i)]^2 \leq r, i' \neq i\} \quad (3.20)$$

donde  $dist(A, B)$  denota la distancia Euclidiana entre  $A$  y  $B$ , mientras que  $r$  toma un valor entero (Li, 1995).

Sin embargo, para un conjunto  $S$  irregular, el conjunto de vecinos  $N_i$  de  $i$  está definido de la misma manera que cuando se trata de un conjunto regular, ya que comprende sitios cercanos dentro de un radio  $r$

$$N_i = \{i' \in S \mid [dist(característica_{i'}, característica_i)]^2 \leq r, i' \neq i\} \quad (3.21)$$

Para este caso la función  $dist(A, B)$  necesita estar definida para características no puntuales, ya que en general el conjunto irregular de vecinos tienen formas y tamaños diferentes (Li, 1995). Este tipo de sistemas se muestran en la Fig. 3.2.

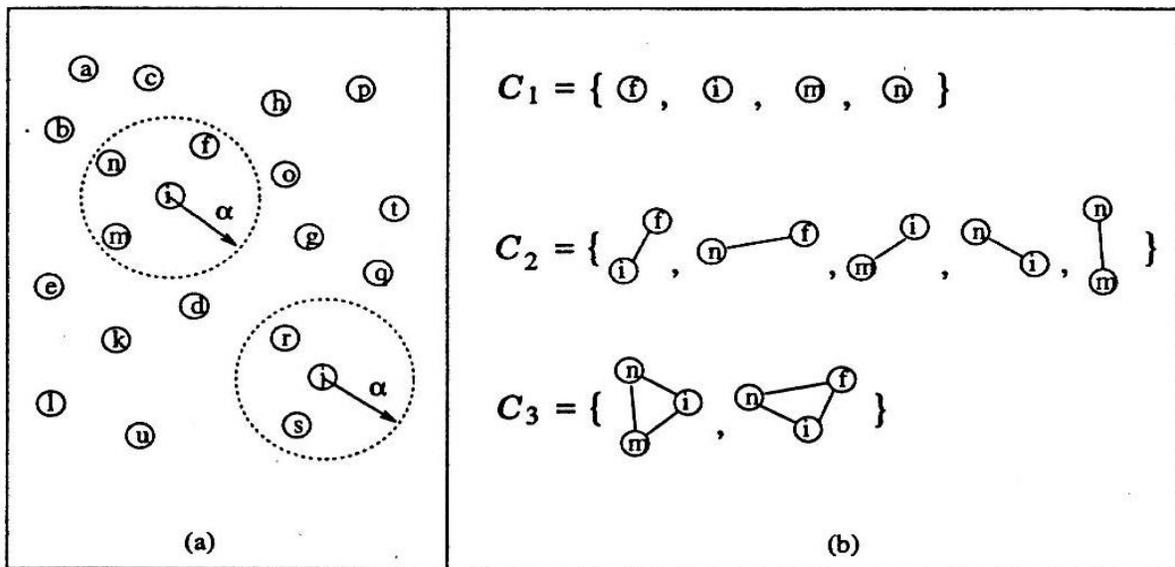


Fig. 3.2. Muestra de vecindades (a) y cliques (b) en un conjunto irregular de sitios.

Sin embargo, debido a que el objeto de estudio para este trabajo es una imagen rectangular, que se puede representar como una superficie rectangular  $L$  y dimensión  $M \times N$  se toma el conjunto de vecinos para un enrejado rectangular normal.

Un ejemplo de vecindad de primer orden para una gráfica lineal es ilustrado en la Fig. 3.3 y caracterizada por la relación (Li, 1995):

$$P[X_i | X_j, j \notin i] = P[X_i | X_{i-1}, X_{i+1}] \Rightarrow X_{N(i)} = \{X_{i-1}, X_{i+1}\} \quad (3.22)$$

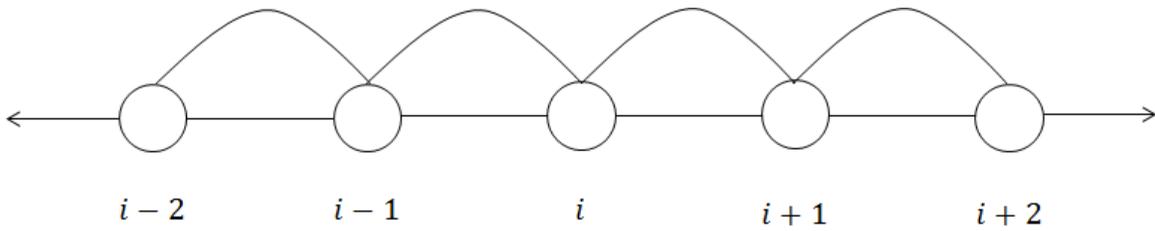


Fig. 3.3. Gráfica lineal para un sistema de vecindad de primer orden.

Para un enrejado rectangular regular con sitios que pertenecen a  $S = \{(i,j) | 1 \leq i, j \leq n\}$ , y que corresponden a los pixeles de una imagen  $n \times n$  de una gráfica 2D, se pueden considerar varios tipos de sistemas de vecindad como de primer orden, segundo orden, n-ésimo orden. En una gráfica completa todos los nodos son vecinos de todos los otros nodos, en la Fig. 3.4 se ilustran algunos ejemplos.

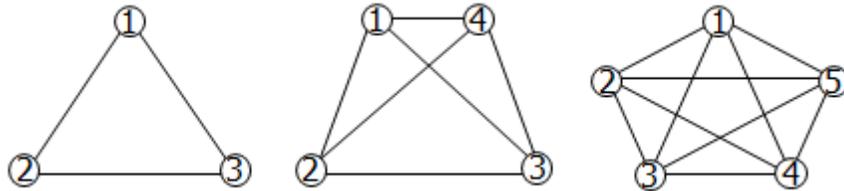


Fig. 3.4. Ejemplo de gráficas completas.

Otro ejemplo es el sistema de vecindad de primer orden 2D, que también es conocido como sistema de cuatro vecindades, se compone de un sitio en el interior con cuatro vecinos como se muestra en la Fig. 3.5 (a), donde  $x$  es el sitio en el interior a estudiar, y que es relativo al nodo  $(i,j)$ , mientras que los 0's representan los cuatro vecinos; de modo que la vecindad de primer orden del nodo  $(i,j)$  está dado por:

$$\mathfrak{N}_1(i,j) = \{(i-1,j), (i,j-1), (i+1,j), (i,j+1)\} \quad (3.23)$$

Por lo tanto, si el nodo  $(i,j)$  no está en el límite del enrejado, este tiene cuatro vecinos, como se muestra en la Fig. 3.6.

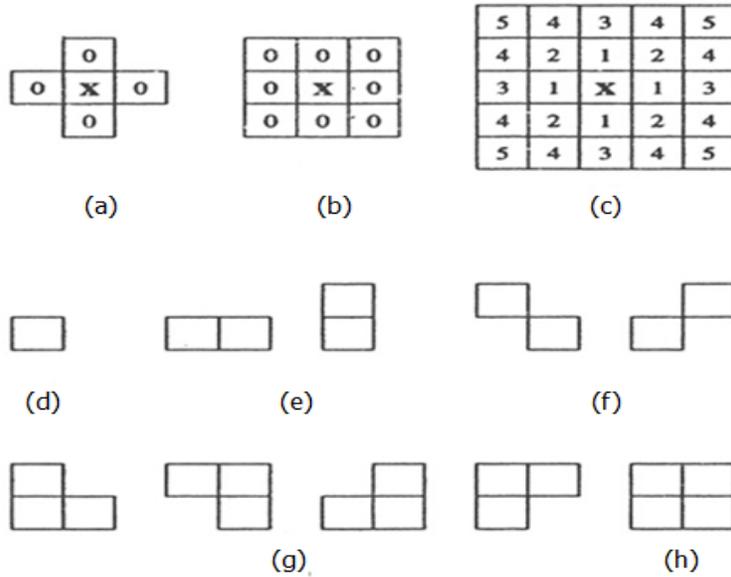


Fig. 3.5. Vecindades y cliques en una red de sitios regulares.

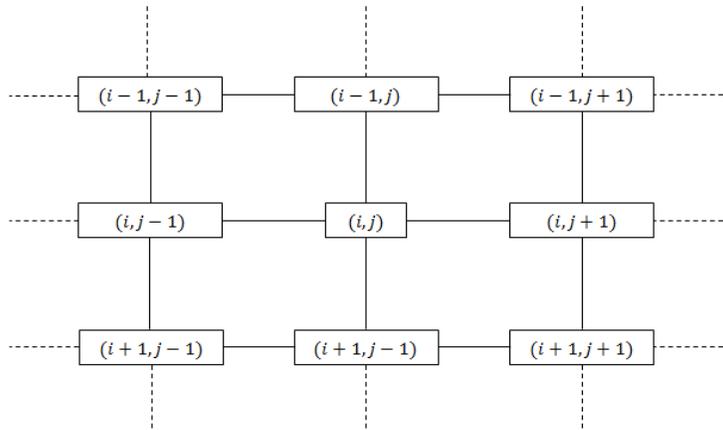


Fig. 3.6. Ejemplo de un sistema de vecindad de primer orden.

Otro ejemplo de sistema de vecindades común es el de segundo orden o sistema de 8 vecindades, ya que hay 8 vecinos para cada sitio que se encuentra en el interior para su estudio. En la Fig. 3.5 (b) se tiene una ilustración de un sistema de vecindades de segundo orden donde los vecinos son representados por 0's o también se puede apreciar en la Fig. 3.7 que el nodo interior cuenta con 8 vecinos, a menos que este en la orilla; mientras que en la Fig. 3.5 (c) se indican los vecinos más alejados en un sistema de n-ésimo orden donde los números del 1 al 5 indican los vecinos más alejados. Entonces, para un sistema de vecindad de segundo orden del nodo  $(i, j)$  está dado de la siguiente manera (Li, 1995):

$$\mathfrak{N}_2(i, j) = \{(i - 1, j), (i - 1, j - 1), (i, j - 1), (i + 1, j - 1), (i + 1, j), (i + 1, j + 1), (i, j + 1), (i - 1, j + 1)\} \quad (3.24)$$

De forma general, para un sistema de vecindades de orden  $c$  para el nodo  $(i, j)$  está dado de la siguiente manera (Li, 1995):

$$N_c(i, j) = \{(k, l) | 0 < (k - i)^2 + (l - j)^2 \leq c\} \quad (3.25)$$

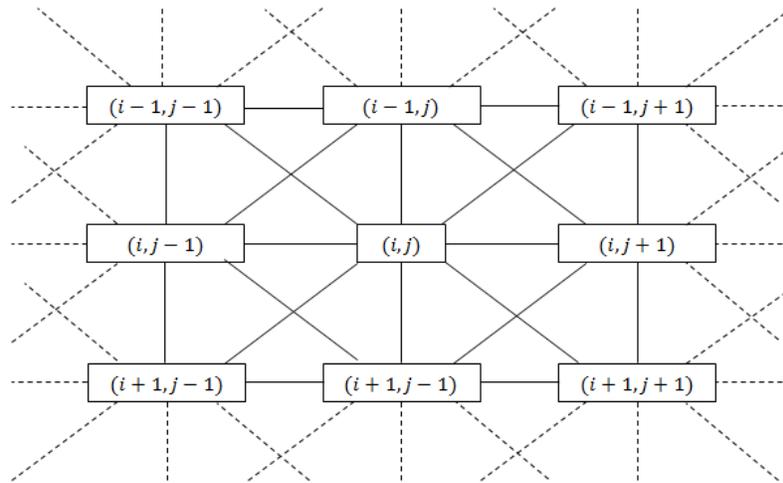


Fig. 3.7 Ejemplo de un sistema de vecindad de segundo orden con gráfica completa.

### 3.2.4 Cliques

Sitios ubicados en un conjunto de vecindad, por su estrecha relación en sus características, se pueden separar en subconjuntos que son conocidos como cliques. Estos cliques son subconjuntos de  $S$ ,  $C \subseteq S$  o  $C \subseteq L$ , donde  $L$  es la superficie rectangular y con dimensión  $M \times N$  que representa a la imagen. Entonces, los cliques abarcan nodos o sitios que tienen características semejantes, si y sólo si en este subconjunto cuenta con el nodo en estudio o interno. Los conjuntos de cliques se indican con una  $C$ .

Con frecuencia los conjuntos de cliques se escriben como  $\langle s, t \rangle$  si  $s$  y  $t$  son vecinos (Li, 1995). En la Fig. 3.8 se muestran algunos ejemplos de cliques.

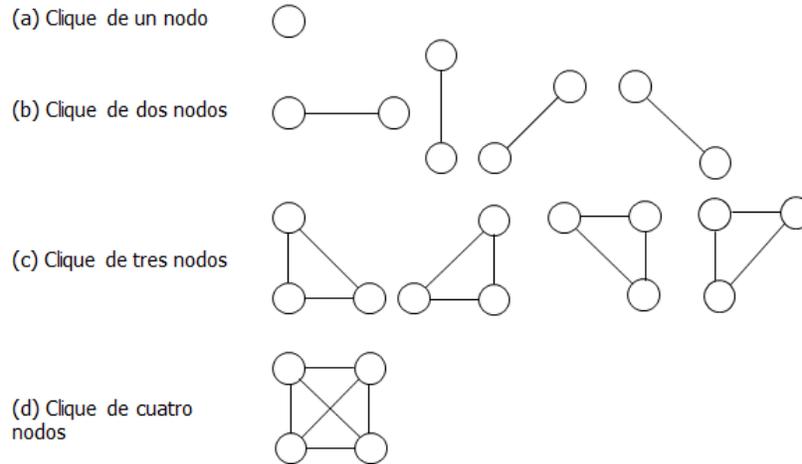


Fig. 3.8 Ejemplos de cliques.

Para el caso de un conjunto irregular de estados o sitio  $S$ , el conjunto de vecinos tiene tamaño y dimensiones variables. Así que el gráfico  $(S, N)$  es donde  $S$  contiene los nodos y  $N$  determina el enlace o las conexiones entre los nodos según sea la relación de vecindad. Este gráfico también es aplicable para el enrejado regular o sitios de  $S$  regular, como es el caso de una imagen.

Entonces el clique de un enrejado regular para  $(S, N)$  se determina por su tamaño, forma y orientación; se define a  $C$  como un subconjunto de  $S$ .  $C$  consiste en un solo sitio  $c = \{i\}$ , un par de sitios vecinos  $c = \{i, i'\}$ , o de un triplete de vecinos  $c = \{i, i', i''\}$  y así sucesivamente, los cuales son indicados por  $C_1$ ,  $C_2$  y  $C_3$  respectivamente (Li, 1995).

$$C_1 = \{i | i \in S\} \quad (3.26)$$

$$C_2 = \{\{i, i'\} | i' \in N_i, i \in S\} \quad (3.27)$$

y

$$C_3 = \{\{i, i', i''\} | i, i', i'' \in S \text{ son vecinos el uno al otro}\} \quad (3.28)$$

$$C = C_1 \cup C_2 \cup C_3 \dots \quad (3.29)$$

De las configuraciones de cliques de las ecuaciones anteriores (3.26, 3.27, 3.28 y 3.29) los sitios en un clique están de manera ordenada, lo que quiere decir es que  $\{i, i'\}$  no es el mismo clique que  $\{i', i\}$ , así como  $\{i', i''\}$  no es lo mismo que  $\{i'', i'\}$ . Entonces la colección de todos los cliques para un conjunto  $(S, N)$  es como en la ecuación 3.29.

Por otra parte, los cliques que pertenecen a sitios irregulares se representan por el número de nodos involucrados. Por ejemplo, se considera el caso presentado en la Fig. 3.2 (a) (circulo con línea punteada en la parte superior), donde se aprecia que  $m$  y  $n$  son vecinos uno del otro, así como  $n$  y  $f$ , mientras que en la Fig. 3.2 (b) se muestran los cliques de sitio único ( $C_1$ ), par ( $C_2$ ) y triple ( $C_3$ ) asociado con el conjunto de la Fig. 3.2 (a); debido a las circunstancias anteriores el conjunto  $\{m, i, f\}$  no pueden formar un clique porque  $f$  y  $m$  no son vecinos (Li, 1995). Se puede notar que mientras sea más grande el sistema de vecindades más grande serán los cliques.

### 3.3 Campos Aleatorios de Markov, MRF, Markov Random Fields

“Los Campos Aleatorios de Markov originalmente fueron utilizados en mecánica estadística para modelar sistemas de partículas en un enrejado bidimensional o tridimensional.” (Ibe, 2009, p. 382). Geman and Geman introdujo los Campos Aleatorios de Markov MRF, Markov Random Field, en la segmentación de imágenes, las cuales se modelan como datos organizados en enrejados y representadas como pixeles (Ibe, 2009).

Las características que componen a una imagen como la intensidad, color y textura provocan que haya una relación espacial entre pixeles vecinos; así que, los MRF son apropiados para el procesamiento de imágenes, ya que el concepto detrás de los MRF es que, cuando un proceso está en una ubicación en particular se tiene una mayor influencia de los eventos en puntos relativamente más cercanos a este que los más alejados (Ibe, 2009).

Los MRF se consideran como campos aleatorios porque están definidos por un conjunto de nodos espaciales o sitios, los cuales son representados por los pixeles de la imagen. Los sitios que se ubican en un punto en particular estarán más influenciados por los pixeles más cercanos que los alejados; así que, los sitios pueden ser agrupados por vecindades y dependiendo de sus características se pueden definir cliques (pixeles, con características muy semejantes como en color, textura o intensidad). Por lo tanto, en los MRF, el estado del pixel analizado depende de los vecinos más cercanos a este (Herrera, 2015).

### 3.3.1 Definición de los MRF

Para definir un MRF se considera un conjunto de variables aleatorias  $F = \{F_1, \dots, F_m\}$  (campo aleatorio) definidas en un conjunto  $S$  en el que cada variable aleatoria  $F_i$  toma valores  $f_i$  en  $L$  (conjunto de etiquetas). El evento conjunto,  $F_i = f_i$ , denota el evento de que  $F_i$  toma un valor  $f_i$  y la notación  $(F_1 = f_1, \dots, F_m = f_m)$ ; por simplicidad, se abrevia como  $F = f$  donde  $f = \{f_1, \dots, f_m\}$  que es una configuración de  $F$ , lo cual corresponde a la realización del campo. Para un conjunto discreto de etiquetas  $L$ , la probabilidad para que una variable aleatoria  $F_i$  tome el valor  $f_i$  se define como  $P(F_i = f_i)$  o se abrevia como  $P(f_i)$ . La probabilidad conjunta se denota  $P(F = f) = P(F_1 = f_1, \dots, F_m = f_m)$  y se abrevia como  $P(f)$ .  $F$  se dice que es un MRF en  $S$  con respecto al sistema de vecindades  $N$  si satisfacen las siguientes condiciones (Li, 1995).

$$P(f) > 0 \quad \forall f \in \mathbb{F} \quad (3.30)$$

$$P(f_i | f_{S-\{i\}}) = P(f_i | f_{N_i}) \quad (3.31)$$

donde  $S - \{i\}$  es la diferencia de conjuntos,  $f_{S-\{i\}}$  denota el conjunto de etiquetas en los sitios  $S - \{i\}$  y  $f_{N_i} = \{f_{i'} | i' \in N_i\}$  representa el conjunto de etiquetas en los sitios vecinos de  $i$ .

La primera condición, dada por la ecuación 3.30, se llama propiedad de positividad, la cual asegura que todas las configuraciones  $f$  tienen una oportunidad de ocurrir. La segunda propiedad dada por la ecuación 3.31 es llamada *propiedad Markoviana* o *Markovianidad*, la cual establece las características locales de  $f$ ; es decir, que los sitios vecinos tienen interacción o son dependientes el uno del otro (Li, 1995); así que, “la probabilidad de un suceso local a  $L$  condicionado a los sucesos restantes es equivalente a aquella condicionada a los sucesos vecinos de  $L$ .” (Encinas, 2015).

### 3.3.2 Campos Aleatorios de Gibbs, GRF, Gibbs Random Fields

Para definir un Campo Aleatorio de Gibbs, GRF, Gibbs Random Fields, se requiere, al igual que en los MRF's, un conjunto de variables aleatorias  $F = \{F_1, \dots, F_m\}$  (llamado campo aleatorio) definidas en el conjunto  $S$ . Se dice que “ $F$  es un campo aleatorio de Gibbs en  $S$  con respecto a  $N$  si y sólo si su configuración obedece a la distribución de Gibbs” (Li, 1995, p. 13). Donde  $N$  es un sistema de vecindad.

Li (1995) define la distribución de Gibbs de la siguiente forma:

$$P(f) = Z^{-1} \times e^{-\frac{1}{T}U(f)} \quad (3.32)$$

donde:

$$Z = \sum_{j \in \mathbb{F}} e^{-\frac{1}{T}U(f)} \quad (3.33)$$

$Z$  es una constante de normalización que es llamada función de partición,  $T$  es la constante de temperatura y  $U(f)$  es la función de energía. La energía:

$$U(f) = \sum_{c \in \mathcal{C}} V_c(f) \quad (3.34)$$

es una sumatoria de todos las *potenciales cliques*,  $V_c(f)$ , sobre todos los posibles cliques  $\mathcal{C}$ . El valor de  $V_c(f)$  depende de la configuración local en el clique  $c$ . (p. 13).

$P(f)$  es la probabilidad de ocurrencia de una configuración en particular o “patrón”,  $f$ . Las configuraciones que cuentan con menor energía tienen mayor probabilidad de ocurrencia, mientras que, la temperatura  $T$  controla la nitidez de la distribución. Cuando  $T$  es grande, toda la configuración tiende a distribuirse por igual, mientras que, con una  $T$  cercana a cero, la distribución se concentra alrededor de la energía mínima alcanzada (Li, 1995).

### 3.3.3 Equivalencia Markov-Gibbs

Los MRF se caracterizan por su propiedad local o la *Markovianidad* mientras que los GRF se caracterizan por su propiedad global o distribución de Gibbs. La equivalencia que existe entre estas dos propiedades se establece por el Teorema de Hammersley-Clifford (Li, 1995). El teorema establece que, un conjunto de variables aleatorias  $F$  es un MRF en un conjunto de sitios  $S$  con respecto a un sistema de vecindades  $N$  si y sólo si  $F$  es un GRF en  $S$  con respecto a  $N$ .

Para demostrar que un GRF es un MRF, Li (1995) postula:

Sea  $P(f)$  una distribución de Gibbs en  $S$  con respecto al sistema de vecindades  $N$ . Se considera la probabilidad condicional.

$$P(f_i | f_{S-\{i\}}) = \frac{P(f_i, f_{S-\{i\}})}{P(f_{S-\{i\}})} = \frac{P(f)}{\sum_{f'_i \in L} P(f')} \quad (3.35)$$

donde  $f' = \{f_1, \dots, f_{i-1}, f'_i, \dots, f_m\}$  es cualquier configuración que está de acuerdo con  $f$  en todos los sitios posibles excepto en  $i$ . Escribiendo  $P(f) = Z^{-1} \times e^{-\sum_{c \in C} V_c(f)}$  se obtiene:

$$P(f_i | f_{S-\{i\}}) = \frac{e^{-\sum_{c \in C} V_c(f)}}{\sum_{f'_i} e^{-\sum_{c \in C} V_c(f')}} \quad (3.36)$$

Se divide a al conjunto de cliques  $C$  en dos conjuntos  $A$  y  $B$ ; el conjunto  $A$  consiste de cliques que contienen a  $i$  y el conjunto  $B$  son cliques que no contienen a  $i$ . Así que, la ecuación anterior puede escribirse como

$$P(f_i | f_{S-\{i\}}) = \frac{[e^{-\sum_{c \in A} V_c(f)}][e^{-\sum_{c \in B} V_c(f)}]}{\sum_{f'_i} \{[e^{-\sum_{c \in A} V_c(f'_i)}][e^{-\sum_{c \in B} V_c(f'_i)}]\}} \quad (3.37)$$

Haciendo  $V_c(f) = V_c(f')$  para cualquier clique  $c$  que no contenga a  $i$ ,  $e^{-\sum_{c \in B} V_c(f)}$  se cancela tanto en el numerador como en el denominador. Por lo tanto, la probabilidad dependerá sólo de los potenciales cliques que contengan a  $i$

$$P(f_i | f_{S-\{i\}}) = \frac{e^{-\sum_{c \in A} V_c(f)}}{\sum_{f'_i} e^{-\sum_{c \in A} V_c(f')}} \quad (3.38)$$

es decir, esto depende de las etiquetas en el sistema de vecindades de  $i$ . (pp. 14-15).

Lo anterior demuestra que el campo aleatorio de Gibbs es un campo aleatorio de Markov. Sin embargo la prueba de que un MRF es un GRF es mucho más complicada y no se estudia en este trabajo.

Estas funciones potenciales determinan la forma de la distribución de Gibbs; ya que cuando todos los parámetros involucrados en la función potencial son especificados, la distribución de Gibbs está completamente definida (Li, 1995).

### 3.3.4 Uso de los modelos MRF

Algunos de los usos de los modelos de MRF son para el modelado de las propiedades de la imagen como regiones o texturas. Para el modelado de estas características se toma en cuenta la distribución condicional y conjunta, así como la correspondiente función de energía.

#### 3.3.4.1 Modelo de Ising

El modelo de Ising o Auto Modelo surge del problema de modelar materiales ferromagnéticos. La hipótesis es que el campo magnético generado por un cuerpo continuo es el resultado de la combinación de muchos momentos magnéticos dipolares atómicos o moleculares dentro del material (De Samaniego, 2007). Este modelo postula que existe una rejilla con dipolos magnéticos (campo magnético producido por algún elemento) o *spins* (propiedad física de las partículas subatómicas en la que se establece que cada partícula tiene un momento angular intrínseco de valor fijo) en cada sitio.

Así que, en cada sitio, al dipolo magnético de le puede asignar valores ya sea de 0 y 1 o +1 y -1. Se asume que hay  $N$  sitios en un enrejado finito o red cuadrada (rejilla) finita y los vecinos de  $s \in S$  son los sitios con distancia Euclidiana (con  $C = 1$  para este caso) etiquetados como  $i = 1, \dots, N$ , y sea una variable aleatoria  $\varphi_i(x)$  que se asigna a cada sitio  $i$ . Los *spins* o giros serán las variables aleatorias  $\varphi_i(x)$  las cuales pueden tomar dos valores  $\varphi_i(x) = \pm 1$  que representan los estados de cada sitio.

Para la configuración  $x = \{x_1, x_2, \dots, x_N\}$  los valores de la variable aleatoria que se toman como (Ibe, 2009):

$$\varphi_i(x) = \begin{cases} 1 & x_i = + \\ -1 & x_i = - \end{cases} \quad (3.39)$$

Por lo tanto, existen  $2^N$  configuraciones posibles,  $\{\varphi_1(x), \varphi_2(x), \dots, \varphi_N(x)\}$ , del sistema. En el caso más simple la función de energía está dada por:

$$H(x) = -\beta \sum_{\langle s,t \rangle} x_s x_t \quad (3.40)$$

donde  $\langle s,t \rangle$  indica que  $s$  y  $t$  son vecinos. Así que  $H$  corresponde a la función de energía de un potencial par de vecinos, y  $\beta$  representa la constante de acoplamiento que define la interacción entre spins (Winkler, 1995). Las configuraciones constantes de los estados  $-1$  y  $+1$  representan la configuración de un mínimo de energía.

Este tipo de modelos se codifican en la función de energía de Gibbs como un par de sitios de cliques potenciales; entonces, la función de energía para cliques de primer orden y de segundo orden está definida por Li (1995, p. 17) como:

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'}) \quad (3.41)$$

La expresión se conoce como energía de segundo orden porque se calcula para cada clique formado por dos estados. Para especificar un MRF o GRF es necesario que se especifique los valores de  $V_1$  y  $V_2$ .

Cuando el  $V_1(f_i) = f_i G_i(f_i)$  y  $V_2(f_i, f_{i'}) = \beta_{i,i'} f_i f_{i'}$ , donde  $G_i(\cdot)$  son funciones arbitrarias y  $\beta_{i,i'}$  son constantes que reflejan la interacción entre el par de sitios  $i$  e  $i'$ , la función de energía es:

$$U(f) = \sum_{\{i\} \in C_1} f_i G_i(f_i) + \sum_{\{i,i'\} \in C_2} \beta_{i,i'} f_i f_{i'} \quad (3.42)$$

La expresión anterior se conoce como automodelo y para que este sea un modelo autologístico los  $f_i$ 's toman un valor en el conjunto discreto de etiquetas  $L = \{0,1\}$  o  $L = \{-1,+1\}$ , la energía correspondiente tiene la forma:

$$U(f) = \sum_{\{i\} \in C_1} \alpha_i f_i + \sum_{\{i,i'\} \in C_2} \beta_{i,i'} f_i f_{i'} \quad (3.43)$$

donde  $\beta_{i,i'}$  puede verse como un coeficiente de interacción. Cuando  $N$  se considera como un sistema de vecindad que consiste de cuatro vecinos en una malla o rejilla 2D, el modelo autologístico recibe el

nombre de modelo de Ising. La probabilidad condicional toma la siguiente forma:

$$\begin{aligned}
 P(f_i | f_{N_i}) &= \frac{e^{\alpha_i f_i + \sum_{i' \in N_i} \beta_{i,i'} f_i f_{i'}}}{\sum_{f_i \in \{0,1\}} e^{\alpha_i f_i + \sum_{i' \in N_i} \beta_{i,i'} f_i f_{i'}}} \\
 &= \frac{e^{\alpha_i f_i + \sum_{i' \in N_i} \beta_{i,i'} f_i f_{i'}}}{1 + e^{\alpha_i + \sum_{i' \in N_i} \beta_{i,i'}}}
 \end{aligned} \tag{3.44}$$

Cuando se tiene una distribución homogénea, se tiene que  $\alpha_i = \alpha$  y  $\beta_{i,i'} = \beta$  sin importar  $i$  ni  $i'$ . (Li, 1995, pp. 17-18).

### 3.3.4.2 Modelo de Potts

El modelo de Potts es una generalización del modelo de Ising que, también es conocido como *Multi-Level Logistic Model*, *MLL* o modelo generalizado de Ising; ya que existen más de dos estados para cada sitio; es decir que hay más dos etiquetas discretas,  $M > 2$ , en el conjunto de etiquetas  $L = \{1, \dots, M\}$ , donde el potencial clique depende del número de sitios que lo forman así como su orientación en la que se encuentran (Li, 1995).

Para los cliques que tienen más de un sitio, se definen las funciones de potencial para cada cliques que tiene la siguiente forma:

$$V_c(f) = \begin{cases} -\zeta_c & \text{Si todos los estados de } c \text{ tiene la misma etiqueta} \\ \zeta_c & \text{en otro caso} \end{cases} \tag{3.45}$$

$\zeta_c$  es el potencial para los cliques de tipo  $c$ , para los cliques de un solo sitio dependerá de la etiqueta asociada a este  $V_c(f) = V_c(f_i) = \alpha_I$  si  $f_i = I \in L_d$  donde  $\alpha_I$  es el potencial para la etiqueta  $I$ .

Un modelo MLL será de segundo orden, si y sólo si  $\alpha$  (para cliques de un solo estado) y  $\beta$  (para cliques de dos estados) son parámetros distintos de cero. Así que, la función potencial clique para clique de dos estados es (Li, 1995):

$$V_c(f) = \begin{cases} -\beta_c & \text{Si todos los estados del clique } \{i, i'\} = c \in C_2 \text{ tiene la misma etiqueta} \\ \beta_c & \text{en otro caso} \end{cases} \tag{3.46}$$

$\beta_c$  es el parámetro  $\beta$  para los cliques de tipo  $c$  y  $C_2$  es el conjunto de cliques de dos estados. Para un sistema de vecindades de 4 vecinos existen cuatro tipos de

cliques de dos estados y cuatro diferentes  $\beta_C$ . Cuando el modelo es isotrópico los cuatro toman el mismo valor. Debido a la simplicidad del modelo de Potts, este es usado ampliamente para modelar regiones y texturas (Li, 1995).

Así que para un modelo isotrópico, la probabilidad condicional puede ser expresada de la siguiente manera:

$$P(f_i = I | f_{N_i}) = \frac{e^{-\alpha_I - \beta n_i(I)}}{\sum_{I=1}^M e^{-\alpha_I - \beta n_i(I)}} \quad (3.47)$$

donde  $n_i(I)$  es el número de estados de  $N_i$  que tienen etiquetas  $I$ . El caso particular para el cual las etiquetas son  $\{0,1\}$  pasa a ser el modelo de Ising (Li, 1995).

Para más de dos estados la función de energía está dada por:

$$H(x) = -\beta \sum_{\langle s,t \rangle} 1_{\{x_s = x_t\}} \quad (3.48)$$

Este es llamado el modelo de Potts. También cada término en la sumatoria puede ser individualmente ponderado, es decir:

$$H(x) = \sum_{\langle s,t \rangle} a_{st} x_s x_t + \sum_s a_s x_s \quad (3.49)$$

donde  $x_s = \pm 1$ . Si  $a_{st} = 1$  entonces  $x_s = x_t$  es favorable y, a la inversa,  $a_{st} = -1$  establece  $x_s = -x_t$

En el modelo de Ising se asume que cada sitio en el enrejado puede tener uno de dos spins, sin embargo en el modelo de Potts cada sitio del enrejado puede ser en uno de los estados distintos  $q \geq 2$ . El modelo Ising corresponde al modelo de Potts con  $q = 2$  (Li, 1995).

### 3.3.4.3 La suavidad *a priori*

La suavidad asume que las propiedades físicas en una zona de espacio o en un intervalo de tiempo presentan coherencia y no hay cambios abruptos. La suavidad ha sido una suposición *a priori* en la visión de bajo nivel, que se ha desarrollado en un marco general, llamado regularización (Li, 1995).

La limitante de suavidad se expresa a veces por la probabilidad *a priori* o equivalentemente a un término de energía  $U(f)$  que mide el grado por el cual la suavidad asumida se altera por  $f$ . Hay dos formas básicas de las cuales los términos de suavidad corresponden a situaciones con etiquetas discretas y continuas.

[Del apartado del modelo de Potts o MLL en la ecuación 3.45 y 3.46] con una negativa  $\zeta$  y un coeficiente  $\beta$  proporciona un método para construir términos de suavidad para etiquetas discretas desordenadas. Siempre que todas las etiquetas  $f_c$  en un clique tomen el mismo valor será una solución  $f$  localmente suave; por lo que se incurre en un potencial clique negativo (costo); de otra manera, si no son los mismos se incurre en un potencial positivo.

Para MRF's continuos [...] el potencial clique en un punto es de la forma  $[f^{(n)}(x)]^2$ . El orden  $n$  determina el número de sitios en el clique involucrado; por ejemplo,  $[f^1(x)]^2$  donde  $n = 1$  corresponde a un par de sitios potenciales suaves. Diferente orden corresponde a diferente clase de suavidad.

Se toma como ejemplo la restauración o reconstrucción de una superficie sin textura, donde  $f = \{f_1, \dots, f_m\}$  es el mapeo de una superficie  $f(x)$  en  $x \in [a, b]$  en una superficie unidimensional por simplicidad. La distribución de Gibbs  $P(f)$  o energía  $U(f)$  depende de la superficie  $f$  que se quiere reconstruir. Se asume una superficie plana con  $-a$  *a priori*, la cual tiene una ecuación  $f(x) = a_0$  y debería tener una derivada de primer orden igual a cero,  $f'(x) = 0$ ; por lo tanto se debe de escoger una energía *a priori* como:

$$U(f) = \int [f'(x)]^2 dx \quad (3.50)$$

la cual es llamada una *cadena* (string). La energía toma el valor mínimo cero si y sólo si  $f$  es completamente plana o toma un valor positivo en otro caso. Así que, la superficie que minimiza únicamente la ecuación 3.50 tiene una altura constante (valores de gris para una imagen).

En el caso discreto donde la superficie tiene puntos discretos  $a \leq z_i \leq b, i \in S$ . Se usa la diferencia de primer orden para aproximar

la primera derivada y se usa una sumatoria para aproximar la integral, por lo que la energía es:

$$U(f) = \sum_i [f_i - f_{i-1}]^2 \quad (3.51)$$

donde  $f_i = f(x_i)$ .

El equivalente para 2D, llamada como *membrana*, es:

$$\iint \{ [f_x(x, y)]^2 + [f_y(x, y)]^2 \} dx dy \quad (3.52)$$

Similarmente, la energía *a priori* se puede designar para una superficie plana o cuadrática. Con una superficie plana,  $f(x) = a_0 + a_1x$  se tiene una segunda derivada igual a cero  $f''(x) = 0$  y se tiene una energía:

$$U(f) = \int [f''(x)]^2 dx \quad (3.53)$$

llamada barra. La superficie que minimiza sólo tiene gradientes constantes. Para el caso discreto se utiliza una diferencia de segundo orden para poder aproximar la derivada de segundo orden, así que se tiene una energía:

$$U(f) = \sum_i [f_{i+1} - 2f_i + f_{i-1}]^2 \quad (3.54)$$

Un término suave  $U(f)$  se utiliza usualmente en conjunto con otros términos de energía, ya que la superficie que minimiza una de las energías suaves *a priori* únicamente tiene ya sea niveles de grises, gradiente o curvatura constantes. (Li, 1995, pp. 21-22).

### 3.3.5 Etiquetado de Bayes de los MRF

Tanto en la distribución *a priori* y la función de verosimilitud de un patrón conocido, lo que mejor se puede estimar de esta información es el etiquetado de Bayes. La

solución del *máximo a posteriori* (MAP) es un caso especial en el marco de Bayes (Li, 1995).

### 3.3.5.1 Estimación de Bayes

En las estimaciones de Bayes un riesgo es minimizado para obtener la estimación óptima. Este riesgo de Bayes de estimación  $f^*$  se define como:

$$R(f^*) = \int_{f \in \mathbb{F}} C(f^*, f) P(f|d) df \quad (3.55)$$

donde  $d$  es la observación,  $C(f^*, f)$  es una función de costo y  $P(f|d)$  es la distribución *posteriori*. Se necesita calcular la distribución de la probabilidad y de la distribución *a priori*. De acuerdo con las reglas de Bayes, la probabilidad posterior que se puede calcular mediante la siguiente función:

$$P(f|d) = \frac{p(d|f)P(f)}{p(d)} \quad (3.56)$$

donde  $P(f)$  es la probabilidad a priori del etiquetado  $f$ ,  $p(d|f)$  es función de densidad de probabilidad condicional de la observación  $d$ , también llamada función de *verosimilitud* de  $f$  para  $d$  fijo, y  $p(d)$  es la densidad de  $d$  la cual es una constante cuando  $d$  está dada.

La función de costo  $C(f^*, f)$  determina el costo de la estimación de  $f$  cuando  $f^*$  es verdadera. Esta función de costo se define de acuerdo a la preferencia. Dos elecciones de la función comúnmente usadas son las siguientes:

Función de costo cuadrática:

$$C(f^*, f) = \|f^* - f\|^2 \quad (3.57)$$

donde  $\|a - b\|$  es la distancia entre  $a$  y  $b$ , y la función de costo  $\delta(0 - 1)$

$$C(f^*, f) = \begin{cases} 0 & \text{si } \|f^* - f\| \leq \delta \\ 1 & \text{en otro caso} \end{cases} \quad (3.58)$$

donde  $\delta > 0$  es para cualquier constante pequeña. (Li, 1995, pp. 31-32).

### 3.3.5.2 Etiquetado MAP-MRF

La estimación MAP se define como:

$$f^* = \arg \max_{f \in \mathbb{F}} P(f|d) \quad (3.59)$$

debido que la probabilidad de  $p(d)$  es una constante para una  $d$  fija, mientras que la  $P(f|d)$  es proporcional a la distribución conjunta

$$P(f|d) \propto P(f, d) = p(d|f)P(f) \quad (3.60)$$

Por lo tanto la estimación MAP tiene el siguiente equivalente:

$$f^* = \arg \max_{f \in \mathbb{F}} \{p(d|f)P(f)\} \quad (3.61)$$

En el etiquetado MAP-MRF,  $P(f|d)$  es la distribución *a posteriori* de un MRF. Para el etiquetado de los MAP-MRF se usa una formulación simple de la restauración de los MRF como un ejemplo.

Se asume una superficie subyacente o inferior plana; la distribución *a priori* conjunta es:

$$P(f) = \frac{1}{Z} e^{-U(f)} \quad (3.62)$$

donde  $U(f)$  es la energía *a priori* dada en la ecuación 3.52. Asumiendo que la observación es verdadera más el ruido Gaussiano independiente,  $d_i = f_i + e_i$ , donde  $e_i \sim N(\mu, \sigma^2)$ , entonces la densidad de probabilidad es:

$$p(d|f) = \frac{1}{\prod_{i=1}^m \sqrt{2\pi\sigma_i^2}} e^{-U(d|f)} \quad (3.63)$$

donde:

$$U(d|f) = \sum_{i=1}^m (f_i - d_i)^2 / 2 \sigma_i^2 \quad (3.64)$$

es la energía probabilidad. Y ahora la probabilidad *a posteriori* es:

$$P(f|d) \propto e^{-U(f|d)} \quad (3.65)$$

$$\begin{aligned} U(f|d) &= U(d|f) + U(f) \\ &= \sum_{i=1}^m (f_i - d_i)^2 / 2 \sigma_i^2 + \sum_{i=1}^m (f_i - f_{i-1})^2 \end{aligned} \quad (3.66)$$

es la *energía a posteriori*. La estimación MAP se encuentra equivalente al reducir al mínimo la energía *a posteriori*  $f^* = \arg \min_f U(f|d)$ .

Cuando los parámetros  $\sigma_i$  se dan de tal manera que  $U(f|d)$  está totalmente especificado la solución del etiquetado MAP-MRF está completamente definida. (Li, 1995, pp. 32-33).

### 3.3.5.3 Distribución *a priori*

Como se mencionó, las imágenes están compuestas por características similares, es decir, los pixeles que conforman a la imagen tienen intensidad, texturas, si son bordes o no, color, etc. Estas características son mejor representadas en términos de variables aleatorias locales.

Para definir la distribución *a priori* se considera que sea un supuesto que  $X$  es un MRF con un sistema de vecindad  $\mathcal{V}' = \{\mathcal{V}'_s : s \in S\}$  y distribución:

$$P(X = \omega) = \frac{1}{Z} e^{-U'(\omega)} \quad (3.67)$$

$$U'(\omega) = \sum_{c \in \mathcal{C}'} V'_c(\omega) \quad (3.68)$$

donde  $U'(\omega)$  es la función de energía. La ecuación anterior representa un MRF *a priori* que es la representación de Gibbs a través de potenciales cliques, con los cuales es más conveniente usarlos que trabajar con probabilidades. (Zoltan, 1994, p.47).

### 3.3.5.4 Distribución *a posteriori*

Para poder definir una distribución *a posteriori* se debe tener en cuenta el *modelo de degradación*. La observación está relacionada al proceso de etiquetado a través de un *modelo de degradación* el cual modela la relación entre el *campo de etiquetas*  $\mathcal{X}$  y el *proceso de observación*  $\mathcal{F}$ :

$$\mathcal{F} = \psi(H(\mathcal{X}), N) \quad (3.69)$$

y

$$\forall r \in \mathcal{R}: \mathcal{F}_r = \psi(H_r(\mathcal{X}_{\psi(r)}), N_r) \quad (3.70)$$

donde  $\psi(a, b)$  es una función invertible en  $a$ .  $H_r$  es una función local definida con una pequeña parte de  $\psi(r)$  de  $S$  tal que  $\psi(r) \in S$ ,  $|\psi(r)| \ll |S|$  y  $\psi^{-1}(s) = \{r \in \mathcal{R} | s \in \psi(r)\}$ . (Zoltan, 1994, p.47).

## 3.4 Ejemplos de Modelos de Markov

Debido a que los MRF han sido de gran utilidad en el procesamiento de imágenes, se ha logrado un gran avance agregando variables a los MRF. Existen modelos de bajo nivel y alto nivel, siendo los de bajo nivel los más comunes.

Algunos ejemplos de modelos de Markov de bajo nivel son: Restauración de Imagen, Detección de Borde, Segmentación de Textura, Detección de Movimiento, etc. Mientras que para los de alto nivel se tienen que el procesamiento de la imagen es con base en sus propias características, como la detección de algún objeto o reconocimiento en la imagen.

Para los modelos de Markov, un modelo de imagen de MRF se plantea con una formulación matemática, por lo que, sea  $R = \{r_1, r_2, \dots, r_M\}$  un conjunto de sitios y  $\mathcal{F} = \{\mathcal{F}_r: r \in R\}$  un conjunto de datos de la imagen (u observaciones) de estos

sitios y el conjunto de todas estas observaciones posibles  $f = (f_1, f_2, \dots, f_{rM})$  es descrito por  $\Phi$ . Por otro lado se tiene un conjunto de sitios  $S = \{s_1, s_2, \dots, s_N\}$  donde cada uno de estos sitios toma una etiqueta del conjunto  $\Lambda = \{0, 1, \dots, L - 1\}$ . Así que, el conjunto del etiquetado discreto global  $\omega = (\omega_{s_1}, \dots, \omega_{s_N}), \omega_s \in \Lambda$  es un espacio de configuraciones  $\Omega$ . En los sitios  $R$  y  $S$  se puede tener partes en común que también puede referirse a un conjunto común de sitios.

El objetivo es modelar las etiquetas y la observación con un campo aleatorio conjunto  $(x, \mathcal{F}) \in \Omega \times \Phi$ . El campo  $X = \{X_s\}_{s \in S}$  es llamado el *campo de etiquetas* y  $\mathcal{F} = \{F_r\}_{r \in R}$  es llamado el *campo de observación*.

### 3.4.1 Modelos MRF de bajo nivel

Para este tipo de modelos de MRF el procesamiento de bajo nivel se realiza sobre un enrejado regular de imágenes. Se establece un conjunto de sitios  $S = \{1, \dots, m\}$  los cuales indican los pixeles de la imagen sobre un plano de la imagen. Mientras que un conjunto  $L$  contiene etiquetas continuas y discretas para los pixeles (Li, 1995).

#### 3.4.1.1 Modelos de Observación

Li (1995) define que en este tipo de modelos se tiene

...una observación  $d = \{d_1, \dots, d_m\}$  que es un arreglo rectangular de valores de pixeles, en el cual cada pixel toma un valor  $d_i$  en el conjunto  $D$ .

En la observación  $d$  puede ser considerada como una trasformada y una versión degradada de un MRF de la realización de  $f$ . Un modelo de observación general se puede expresar como:

$$d = \varphi(B(f)) \odot e \quad (3.71)$$

donde  $B$  es un efecto de borrosidad en la imagen;  $\varphi$  es una transformación la cual puede ser lineal o no lineal, determinística o no determinística;  $e$  es el ruido del sensor y  $\odot$  es un operador de adición o multiplicación. Cada valor del pixel observado es asumido

para ser la suma de todos los valores de grises reales y el ruido Gaussiano independiente

$$d_i = \varphi(f_i) + e_i \quad (3.72)$$

donde  $\varphi(\cdot)$  es una función lineal y  $e_i \sim N(0, \sigma_i^2)$ . La distribución de probabilidad de  $d$  condicional en  $f$ , o la probabilidad de  $f$ , es:

$$p(d|f) = \frac{1}{\prod_i^m \sqrt{2\pi\sigma_i^2}} e^{-U(d|f)} \quad (3.73)$$

donde:

$$U(d|f) = \sum_{i \in S} \frac{(\varphi(f_i) - d_i)^2}{[2\sigma_i^2]} \quad (3.74)$$

es la energía de probabilidad. La función  $\varphi(\dots)$  mapea una etiqueta  $f_i$  a un valor de gris real donde  $f_i$  puede ser numérico o simbólico, continuo o discreto. La función  $\varphi(\dots)$  puede ser considerada que un único valor numérico para una etiqueta  $f_i$  e indica  $d_i = \varphi(f_i) + e_i$  simplemente como  $d_i = f_i + e_i$ . Entonces la energía de probabilidad llega a ser:

$$U(d|f) = \sum_{i \in S} \frac{(f_i - d_i)^2}{[2\sigma_i^2]} \quad (3.75)$$

para un ruido Gaussiano independiente e idénticamente distribuido. (p. 38).

### 3.4.1.2 Restauración de Imagen

En una imagen se pueden presentar anomalías en sus características como regiones borrosas, por lo cual pierde sus características originales, es decir que la imagen presenta ruido. Con ayuda de los MRF's es posible una restauración de la imagen (Zoltan, 1994).

Se toma una imagen  $\mathcal{F}$  borrosa con ruido que se quiere restaurar. Se considera la siguiente función:

$$\mathcal{F} = \psi(H(\mathcal{X}), N) \quad (3.76)$$

Para este caso,  $N$  es un ruido Gaussiano blanco con media  $\mu$  y varianza  $\sigma^2$ .  $H$  es una matriz de borrosidad de cambio invariante. El proceso de observación consiste de los valores de grises para cada pixel; estos pixeles son sitios que pertenecen al conjunto  $\mathcal{R}$ . Por otro parte, en el *proceso de etiquetado*  $\mathcal{X}$ , debido a la complejidad del mismo, se va a dividir en dos subprocesos: un *proceso de pixel* y un *proceso de línea*. Zoltan define el proceso de etiqueta y de línea como:

Se define un proceso de línea  $\mathcal{X}^l$ , el cual se asume como un MRF sobre un sistema de vecindades homogéneo  $\mathcal{V}$  en  $\mathcal{R}$  y  $\mathcal{X}^p$  es también considerado un MRF sobre un sistema de vecindades como el mostrado en la Fig. 3.9, donde los vecinos son los sitios marcados con negro y gris. Por otra parte  $\mathcal{X}$  tiene una distribución *a priori* de la siguiente forma:

$$P(\mathcal{X}^p = \omega^p, \mathcal{X}^l = \omega^l) = \frac{1}{Z} e^{-U'(\omega^p, \omega^l)} = \frac{1}{Z} e^{-\sum_{c \in \mathcal{C}} V_c(\omega)} \quad (3.77)$$

donde  $\omega = (\omega^p, \omega^l)$ ,  $\omega^p$  toma valores entre los niveles de gris disponibles y  $\omega^l$  entre los estados de línea. Si se escoge un sistema de vecindades  $\mathcal{V}$  suficientemente largo para incluir la dependencia causada por la borrosidad de  $H$ , la distribución *a priori* define un MRF con la siguiente función de energía:

$$U(\omega^p, \omega^l) = U'(\omega^p, \omega^l) + \frac{\|\vec{\mu} - \Psi^{-1}(H(\omega^p), f)\|^2}{2\sigma^2} \quad (3.78)$$

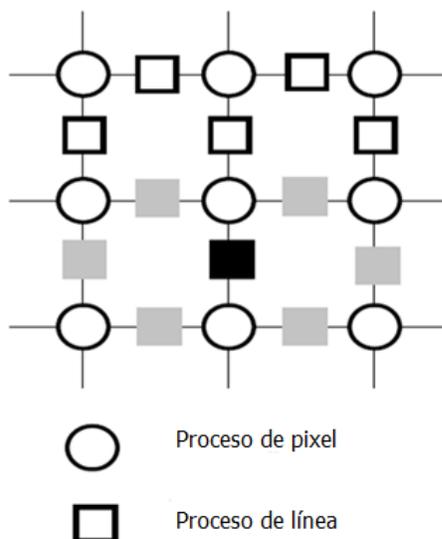


Fig. 3.9 Modelo de restauración de imagen de German

Para un etiquetado óptimo  $\hat{\omega}$  es necesaria una minimización de estimación MAP a la función de energía anterior; así que, la imagen restaurada es dada por el *proceso de pixel*  $\hat{\omega}^P$ . (Zoltan, 1994, pp. 49-50).

### 3.4.1.3 Segmentación de Textura

En la segmentación de textura se divide una imagen en regiones de acuerdo a las texturas de las regiones. La segmentación de textura puede ser supervisada o sin supervisión. Para la segmentación de textura supervisada, se asume que para todos los parámetros de las texturas y para el ruido (si este se presenta), están especificados; esta segmentación divide la imagen en términos de las texturas donde las funciones de distribución han estado completamente especificadas (Li, 1995).

Para la segmentación sin supervisión, además de hacer la partición o división en regiones, también se necesita estimar los parámetros involucrados.

La estimación se debe realizar mediante el uso de la realización de una simple clase de MRF, es decir, un solo tipo de textura, ya sea con ruido o libre de este; sin embargo esta dependerá de los parámetros de la textura fundamental. Una estrategia para resolver esta problemática es la aplicación de un algoritmo iterativo entre segmentación y estimación. (Li, 1995, p. 57).

La segmentación de textura se presenta como otro problema de etiquetado. Se define la energía a posteriori  $P(f)$  de acuerdo con un MRF de alto nivel, debido a la segmentación de textura MAP con el modelo de textura jerárquica (Li, 1995).

La segmentación  $f$  parte a  $S$  en regiones en las que cada uno de los cuales es asignado a un tipo de textura del conjunto de etiquetas  $L$ ; el campo de bajo nivel contribuye a una función de probabilidad  $P(d|f)$  donde  $d$  son los datos de la imagen compuestos de texturas múltiples.

Se define el conjunto de etiquetas  $L \in \{1, \dots, M\}$  mientras que  $f$  representa una segmentación en la cual  $f_i \in L$  es el indicador del tipo de textura para el pixel  $i$ . Indicamos el conjunto de todos los sitios etiquetados  $I$  por:

$$S^{(I)}(f) = \{i \in S | I = f_i\} \quad (3.79)$$

entonces:

$$S = \bigcup_{I \in L} S^{(I)}(f) \quad (3.80)$$

y

$$S^{(I)}(f) \cap S^{(J)}(f) = \phi, \quad I \neq J \quad (3.81)$$

La función de energía de probabilidad puede ser expresada como:

$$U(d|f) = \sum_{c \in C} V_c(d|f) = \sum_{I \in L} \sum_{c \subset S^{(I)}} V_c^{(I)}(d|f) \quad (3.82)$$

donde  $V_c^{(I)}(d|f)$  es la función potencial para los datos  $d$  en  $c$  etiquetados como tipo  $I$ .

Suponer que el tipo de textura  $I$  es modelado como un Modelo Logístico Multinivel con parámetros:  $\theta^{(I)} = \{\alpha^{(I)}, \beta^{(I)}, \dots\}$

Entonces de acuerdo a  $V_c(f) = V_c(f_i) = \alpha_I$  si  $f_i = I \in L_d$ , los potenciales cliques de un solo sitio son:

$$V_c^{(I)}(d|f) = \alpha^{(I)} \quad (3.83)$$

donde  $I = f_i$  y de acuerdo a los potenciales cliques de modelo *logístico multinivel*, de la ecuación 3.46, los potenciales cliques de múltiples sitios son:

$$V_c^{(I)}(d|f) = \begin{cases} \beta_c^{(I)} & \text{si todo } d_i, i \in c, \text{ son lo mismo} \\ \beta_c^{(I)} & \text{en otro caso} \end{cases} \quad (3.84)$$

La expresión anterior corresponde para los cliques en el interior de  $S^{(I)}$ . En los límites de la región así como en estos un clique  $c$  podría pasar a través de dos o más  $S^{(I)}$ 's. En este caso, la siguiente regla puede ser usada para determinar el tipo de textura para la generación de información: Si  $c$  se encuentra en su mayoría en  $S^{(I)}$  entonces escogemos los parámetros del modelo  $\theta^{(I)}$ ; si estos se sitúan por igual en todos los involucrados  $S^{(I)}$ 's, escogemos una  $I$  al azar de las etiquetas involucradas. (Li, 1995, pp. 57-58).

#### 3.4.1.4 Detección de Borde

Para la detección de borde se utilizan modelos de MRF que son a veces *campos aleatorios de Markov-Gauss* (CGMRF, por sus siglas en inglés. Así que, las características para los CGMRF están dadas como:

$$P(f_s | f_r, r \in S) = \frac{1}{\sqrt{2\pi\zeta}} e^{\left(\frac{1}{2\zeta^2}(f_s - \mu_m - \sum_{r \in \mathcal{V}_s} g_r (f_s - \mu_m)^2)\right)} \quad (3.85)$$

donde  $\zeta$  es la derivación,  $\mu_m$  es la media y  $g_r$  es el parámetro de modelo... Las observaciones  $\mathcal{F}$  se consideran corrompidas por un ruido Gaussiano aditivo con media cero y varianza  $\sigma^2$ . El conjunto de etiquetas se trata de un proceso mixto que contiene tanto un *proceso de pixel*  $X^p$  y un proceso de línea... Se asume un sistema de vecindades de primer orden y se marca un proceso de línea horizontal y vertical por  $X^h$  y  $X^v$  respectivamente. Una forma posible para la función de energía está dada como:

$$\begin{aligned}
U(\omega, f) = \frac{1}{2\sigma^2} \sum_{s=(i,j) \in S} & ((f_{i,j} - \omega_{i,j}^p)^2 \\
& + \beta^2(1 - 2(\mathcal{G}_h + \mathcal{G}_v))f_{i,j}^2 \\
& + \mathcal{G}_h (\beta^2(f_{i,j} - f_{i-1,j})^2 (1 \\
& - \omega_{i,j}^h) + \alpha\omega_{i,j}^h) \\
& + \mathcal{G}_v (\beta^2(f_{i,j} - f_{i,j-1})^2 (1 \\
& - \omega_{i,j}^v) + \alpha\omega_{i,j}^v)
\end{aligned} \tag{3.86}$$

con  $1 - 2(\mathcal{G}_h + \mathcal{G}_v) > 0$ .  $\mathcal{G}_v$  y  $\mathcal{G}_h$  son los parámetros del modelo para los cliques verticales y horizontales.  $\beta^2$  corresponde a un término de regularización reflejando la coincidencia en los datos... La estimación del *proceso de línea* está hecha por una aproximación del Campo Medio o Promedio. (Zoltan, 1994, pp. 51-52).

### 3.4.1.5 Análisis de movimiento

Para el análisis de movimiento se presenta un MRF donde el proceso de observación se define tanto en el enrejado de la imagen  $S$  como en el eje del tiempo  $t$  (Zoltan, 1994).

La detección de objetos en movimiento depende del análisis de la variación de la distribución de la intensidad en el tiempo. En cada pixel, se tiene un vector de observación de dos elementos

$$\vec{f}_s^1(t) = |y_s(t) - y_s(t - dt)| \tag{3.87}$$

$y_s(t)$  representa el valor de la intensidad en el pixel  $s$  en el tiempo  $t$ .  $\vec{f}^2$  es un mapa lógico de cambios temporales entre el tiempo  $t$  y  $t - dt$ . Es igual a 1 si un cambio temporal de la intensidad es válido en el sitio  $s$  y 0 en otro caso. Así que el *proceso de etiquetado* es un valor binario. Para la función de energía  $U$  se tiene tres términos. Dos de ellos relacionados a la observación y etiquetas simultáneamente y uno de ellos es usado para reconstruir la máscara de un objeto móvil en un tiempo dado

$$U_1^2(\omega, \vec{f_s^2}) = \sum_{s \in S} V_1(\omega_s(t), \vec{f_s^2}(t), \vec{f_s^2}(t + dt)) \quad (3.88)$$

La otra expresión consiste entre la ocurrencia del etiquetado y la variación de la intensidad:

$$U_1^2(\omega, \vec{f_s^2}) = \sum_{s \in S} \left( \frac{1}{2\sigma^2} (\vec{f_s^1}(t) - \mu\omega_s(t))^2 + (\vec{f_s^1}(t + dt) - \mu\omega_s)^2 \right) \quad (3.89)$$

donde  $\mu$  y  $\sigma$  son parámetros del modelo. El tercer término de la función de energía  $U$  corresponde a  $U_2$  con potenciales que favorecen máscaras homogéneas. (Zoltan, 1994, p. 52).

### 3.4.2 Modelos MRF de alto nivel

La coincidencia de objetos, así como la estimación de reconocimiento y planteamiento se representan mediante las características extraídas de la imagen. Mientras que en la coincidencia visual, la transformación es de la escena (imagen) al modelo del objeto considerado o viceversa. De lo modelos derivados se tiene como resultado un conjunto de características del objetos o de la imagen.

#### 3.4.2.1 Correspondencia bajo condiciones relacionadas

En los modelos de los MRF de alto nivel, el análisis se basa en el tratamiento de las características de la imagen, así como puntos críticos, los parches de línea y superficie, donde este tipo de características son más abstractas que los pixeles de la imagen. Estas características en una escena no sólo se atribuyen por sus propiedades, sino también por la relación que existe entre ellas; ya que una escena o un objeto se presenta por características limitadas debido a las propiedades y relaciones.

#### Representación Estructural Relacional

Una Estructura Relacional, ER, se refiere a un conjunto de las características, propiedades y relaciones de una imagen. El ER describe una escena del modelo de un objeto. Así que el problema de reconocimiento de un objeto se basa en la correspondencia ER (Li, 1995).

Para la ER se toma como bases una escena, de la cual se asume que hay  $m$  características de esta. Las características son incluidas en un índice por un conjunto de sitios  $S = \{1, \dots, m\}$  que constituyen los nodos de la ER. Cada nodo está asociado con un vector compuesto de un número de  $K$  *propiedades unitarias* o *relaciones unitarias*, estas unidades unitarias pueden ser el color de una región, el tamaño de un área, o la longitud de una línea. Cada par de nodos son relacionados unos con otros por un vector compuesto de un número de relaciones binarias (puede ser distancia entre dos puntos o un ángulo entre dos líneas). Entre  $n$  características que pertenecen a  $S$ , habría un vector de relaciones de  $n$ -arreglos  $K_n$ . El alcance de la dependencia relacional puede ser determinado por un sistema de vecindades  $N$  en  $S$ . Ahora la RS para la escena está definida de un triplete (Li, 1995):

$$G = (S, N, d) \quad (3.90)$$

donde  $d = \{d_1, d_2, \dots, d_H\}$  y  $H$  es el orden más alto. Para  $H = 2$ , el RS es también llamado *relación de gráfica* (RG, por sus siglas en inglés). La ER para un objeto de modelo es similarmente definido como:

$$G' = (L, N', D) \quad (3.91)$$

En este caso, el conjunto de etiquetas  $L$  reemplaza el conjunto de sitios. Cada elemento en  $L$  indica una de las características del modelo  $M$ . Un sistema de vecindades para  $L$  se define para consistir de todos los otros elementos (Li, 1995).

$$N'_i = \{I' | \forall I' \in L, I' \neq I\} \quad (3.92)$$

“Esto significa que cada característica del modelo es relacionada a todas las otras características.” (Li, 1995, p. 103). Un modelo ER describe un modelo general o una parte de este. Cuando un RS es usado para describir una parte, por ejemplo una vista de un objeto, el objeto en general puede ser descrito por varios ER y estos pueden ser relacionados por algunas restricciones inter-ER (Li, 1995).

Por otro lado se tiene el concepto de modelo virtual compuesto de un simple nodo  $L_0 = \{0\}$ . Esto es llamado modelo nulo o NULL, el cual representa todo lo no modelado por  $G'$ , así como las características debido a todos los otros objetos de modelo y el ruido. Así el conjunto de etiqueta actual en correspondencia de la escena con el modelo más NULL contiene  $M + 1$  etiquetas. Estos es:

$$L^+ = \{0, 1, \dots, M\} \quad (3.93)$$

Los modelos basados en correspondencia pueden ser considerados como el descubrimiento de un mapeo óptimo de una ER de imagen a una modelo ER (o viceversa). Tal como un mapeo de un RS a otro es llamado *morfismo* escrito como:

$$f: G = (S, N, d) \rightarrow G' = (L, N', D) \quad (3.94)$$

el cual mapea de cada nodo en  $S$  a otro nodo en  $L$

$$f: S \rightarrow L \quad (3.95)$$

y por lo tanto, la relación de mapeo  $f: d_n \rightarrow D_n$

Un morfismo puede llamarse como *isomorfismo* si esto es uno a uno y sobre o encima de. Esto es llamado un *monomorfismo* si esto es uno a uno pero no sobre. Y es llamado un *homomorfismo* si es varios a uno (Li, 1995, pp. 104-105).

### Correspondencia basada en MRF

Para el etiquetado de una escena en términos de un modelo de objeto, se define a  $f = \{f_i \in L^+ | i \in S\}$  donde los elementos en  $S$  indican características de la imagen y estas en  $L$  hacen el modelo de las características del objeto más el NULL; además, se asume que  $f$  es una realización de un MRF con respecto a un sistema de vecindades  $N$  (Li, 1995).

Se describe una Energía y Probabilidad *a posteriori*, en el que todos los casos, un  $N_i$  de un sistema de vecindades  $N$  en  $S$  puede consistir de todos los otros sitios  $i' \neq i$ . Cuando la escena es muy larga,  $N_i$  incluye estos de los otros sitios que son dentro de una distancia espacial  $r$  de  $i$  (Li, 1995).

Li (1995) define la siguiente función de energía *a priori*:

$$U(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'}) \quad (3.96)$$

donde el simple sitio potencial (refiriéndose a un clique) está definido como:

$$V_1(f_i) = \begin{cases} v_{10} & \text{si } f_i=0 \\ 0 & \text{en otro caso} \end{cases} \quad (3.97)$$

donde  $v_{10}$  es una constante. Si  $f_i$  es la etiqueta NULL, esto incurre en una penalidad de  $v_{10}$ ; o ninguna penalidad en otro caso. El potencial del par de sitios se define como:

$$V_2(f_i, f_{i'}) = \begin{cases} v_{20} & \text{si } f_i=0 \text{ ó } f_{i'}=0 \\ 0 & \text{en otro caso} \end{cases} \quad (3.98)$$

donde  $v_{20}$  es una constante. Y al igual que en  $v_{10}$  ocurre la misma penalidad (p.109).

Por otro lado la función de densidad de probabilidad condicional de los datos observados  $d$ , también llamada la *función de probabilidad*, cuando es vista como una función de  $f$  dado  $d$  fija, tiene las siguientes características:

1. Esto es condicionado en puras correspondencias no NULL  $f_1 \neq 0$ .
2. Esto es independiente de un sistema de vecindades  $N$ , y
3. Esto depende en cómo el modelo del objeto es observado en la escena que a su vez depende en la transformación fundamental y el ruido.

Se tiene una energía *a posteriori* de la siguiente forma:

$$\begin{aligned} U(f|d) = & \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'}) \\ & + \sum_{i \in S: f_i \neq 0} V_1(d_1(i)|f_i) \\ & + \sum_{i \in S: f_i \neq 0} \sum_{i' \in S - \{i\}: f_{i'} \neq 0} V_2(d_2(i, i')|f_i, f_{i'}) \end{aligned} \quad (3.99)$$

En esta energía *a posteriori* se encuentran varios parámetros: el ruido, varianza y las penalizaciones *a priori* (Li, 1995).

### 3.4.2.2 Correspondencia de objetos múltiples

La correspondencia de objetos múltiples se trata del etiquetado óptimo de la escena en términos de un modelo de objeto, ya que la configuración de MAP de  $f^*$

deriva la ecuación anterior, ecuación 3.99, en el mapeo óptimo de la escena al modelo del objeto en consideración (Li, 1995).

Se supone que hay modelos potenciales de objeto  $L$ , entonces la solución MAP de  $L$ ,  $f^{(1)}, \dots, f^{(L)}$ , se puede obtener después de la correspondencia de la escena para cada uno de los modelos a su vez. Sin embargo, cada característica de la escena puede tener sólo una correspondencia del modelo de características. Para resolver este inconveniente se utiliza un método de minimización.

Se reescribe  $E(f) = U(f|d)$  en la ecuación [3.99] de la siguiente forma:

$$E(f) = \sum_{i \in S} E_1(f_i) + \sum_{i' \in S, i' \neq i} E_2(f_i, f_{i'}) \triangleq \sum_{i \in S} E(f_i | N_i) \quad (3.100)$$

obteniendo así la siguiente energía posterior local de primer y segundo orden:

$$E_1(f_i) = \begin{cases} v_{10} & \text{si } f_i = 0 \\ V_1(d_1(i) | f_i) & \text{en otro caso} \end{cases} \quad (3.101)$$

$$E_2(f_i, f_{i'}) = \begin{cases} v_{20} & \text{si } i' \in N_i \text{ y } (f_i = 0 \text{ ó } f_{i'} = 0) \\ V_1(d_1(i, i') | f_i, f_{i'}) & \text{si } i' \neq i \text{ y } f_i \neq 0 \text{ y } f_{i'} \neq 0 \\ 0 & \text{otro caso} \end{cases} \quad (3.102)$$

son energía *a posteriori* local formada de dos energías de primer y segundo orden.  $E(f_i | N_i) = E_1(f_i) + \sum_{i' \in S} E_2(f_i, f_{i'})$  es el costo incurrido por una correspondencia local  $i \rightarrow f_i$  que proporciona el resto de las correspondencia. Esto se puede utilizar como la base para la selección de los mejores objetos correspondientes para  $i$  en la correspondencia a los modelos múltiples de objeto. La característica de la imagen  $i$  es considerada un objeto  $l_i$  si la decisión incurre en el último costo

$$l_i = \arg \min_{l \in \{1, \dots, L\}} E(f_i^{(l)} | f_{N_i}^{(l)}) \quad (3.103)$$

El etiquetado final para  $i$  es el número de características  $f_i^{(l)}$  del objeto  $l_i$ . Se nota que la regla anterior no maximiza la posteriori desde por lo menos las funciones de partición que son diferentes de la correspondencia de diferentes objetos.

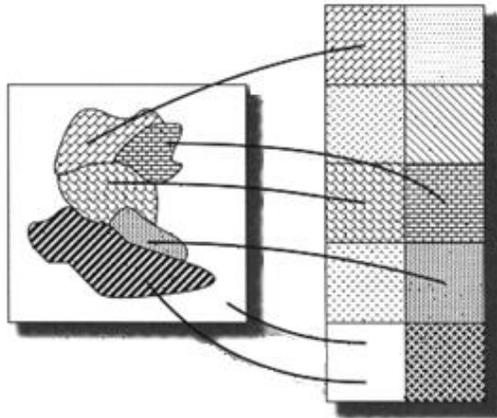


Fig. 3.10 Mapeo de una escena a múltiples modelos de objeto. Con diferentes texturas se tienen diferentes estructuras. Las líneas gruesas representan el sub-mapeo. El fondo y la estructura sin modelo son mapeados como la estructura NULL (Li, 1995).

Aplicando la ecuación [3.100] a cada  $i$  se produce una asignación o mapeo global de la escena a los modelos, compuestos de varios sub-mapeos, como en la Fig. [3.10]. (Li, 1995, pp. 111-112).

A la derecha de la imagen, los cuadrados con diferentes texturas representan las estructuras del modelo en estudio al que la escena debe ser correspondida. El cuadro negro representa el modelo NULL. A la izquierda se encuentra la estructura donde cada región corresponde a una sub-parte de una estructura del modelo que se superponen (sin separación). El reconocimiento de la escena superpuesta es la partición de la estructura de esta escena en partes, tales que, cada parte es debida a un objeto simple y a la correspondencia entre características en cada parte y éstas en la correspondencia del modelo del objeto. En las partes de la escena correspondiente al trasfondo y objetos sin modelado debería ser asignadas a una etiqueta NULL (Li, 1995).

# 4. Optimización estocástica

En este capítulo se presenta la equivalencia entre los MRF y los GRF, de acuerdo al teorema de Hammersley-Clifford, en un marco de la teoría de la probabilidad de Bayes. Además, se estudia el algoritmo de Recocido Simulado que se utiliza para la segmentación de la imagen a procesar para que la función de energía llegue a un mínimo global, y así solucionar el problema de etiquetado de los MRF.

## 4.1 Fundamentos

El proceso de optimización estocástica consiste en optimizar la función de energía *a posteriori*. Este se puede lograr mediante algoritmos clasificados como locales y globales (López, 2008).

Una de las características de los algoritmos globales es que son capaces de salir de los mínimos locales y son inicializados en puntos alejados de la solución global; ya que la solución mínima es usualmente definida como una solución global o una de estas cuando hay mínimos globales múltiples (López, 2008).

El mínimo local y global pueden ser estudiados mediante el análisis de convexidad porque se establece una función real  $E(f)$  definida en  $\mathbb{F} = \mathbb{R}^n$ . La función se dice que es convexa en  $\mathbb{F}$  si:

$$E(\lambda x + (1 - \lambda)y) \leq \lambda E(x) + (1 - \lambda)E(y) \quad (4.1)$$

para dos puntos  $x, y \in \mathbb{F}$  y cualquier número real  $\lambda \in [0,1]$ .

...Un mínimo local es también un mínimo global si  $E(f)$  es convexo en  $\mathbb{F}$ . Este es el mínimo global único si  $E(f)$  es estrictamente convexo. Por lo tanto, cuando  $E(f)$  es convexo, se puede encontrar un mínimo global mediante la búsqueda de un mínimo local empleando el método descendente de gradiente. (Li, 1995, p.207).

Para la minimización global se requiere encontrar todos los mínimos locales y demostrar que no haya más; ya que si no se hace lo anterior sería un procesamiento muy lento y arduo, por lo que se deben encontrar un mínimo global exacto para encontrar aproximaciones a este con menor tiempo de procesamiento. Para afrontar el problema del mínimo local se tienen métodos comúnmente usados como la búsqueda aleatoria y de recocido. Para la búsqueda aleatoria, para obtener un mínimo local en el mínimo global, se toma como estrategia que

...una nueva configuración no siempre hace la energía descendente y un aumento en la energía ocasionalmente se permite. El cambio para generar una configuración es proporcional a  $e^{-E(f)T}$  donde  $T$  es un parámetro de control. Por lo tanto, este método encuentra una configuración de energía menor con una probabilidad grande. (Li, 1995, p. 208)

En el método del recocido se incorpora el método de búsqueda local para superar el problema del mínimo local. Esto se logra con la disminución de algún parámetro, como la temperatura en la distribución de Gibbs, ya que va de un valor alto a uno bajo durante una minimización iterativa.

## 4.2 Segmentación con Markov

Retomando el teorema de Bayes como una probabilidad condicional, se genera una imagen segmentada de la original; ya que se define una variable  $Y$  que representa la imagen segmentada de la imagen original  $X$  con  $i = 1, 2, \dots, n$ , la probabilidad condicional de la imagen segmentada tal que la imagen original se define como:

$$P(Y_i|X) = \frac{P(X|Y_i)P(Y_i)}{P(X)} \quad (4.2)$$

donde:

$P(Y_i|X)$  es la probabilidad *a posteriori*

$P(Y_i)$  es la probabilidad *a priori* de la clase  $Y_i$

$P(X|Y_i)$  es la probabilidad conjunta de  $X$  dado  $Y_i$

$P(X)$  es la probabilidad total de  $X$

En el teorema de Hammersley-Clifford se establece la equivalencia de un MRF con un GRF; así que la probabilidad a priori de la clase  $P(Y_i)$ , define un MRF y por lo tanto  $P(Y_i)$  puede ser expresada como una distribución de Gibbs:

$$P(Y_i) = Z^{-1} \times e^{-U(Y_i)} \quad (4.3)$$

$Z$  no proporciona información discriminante por lo que no se considera (Li, 1995). Mientras que las funciones de energía  $U(Y_i)$  son la suma de los potenciales cliques  $V_c(f)$  sobre todos los posibles cliques  $c$ , es decir:

$$U(Y_i) = \sum_{c \in C} V_c \quad (4.4)$$

Se consideran los cliques de segundo orden debido a su sencillez y bajo costo computacional (menos tiempo de procesamiento). La función potencial de clique queda definida como:

$$V_c = -\beta\delta(y_i - y_j) \quad (4.5)$$

donde el clique  $c$  contiene los sitios  $i$  y  $j$ , además que  $y_i$  y  $y_j$  son los niveles de gris asignados a cada sitio (López, 2008).

Para calcular la energía *a posteriori* se emplea la aproximación de la regla de Bayes:

$$P(Y_i|X) \approx P(X|Y_i)P(Y_i) \quad (4.6)$$

y bajo el esquema MAP se maximiza la distribución a posteriori  $P(Y_i|X)$  o se minimiza una función equivalente de energía mediante una configuración  $\psi$  (López, 2008)

$$\psi = \arg \max P(Y = \psi|X = x) = \arg \min U(\psi|x) \quad (4.7)$$

Entonces, la ecuación 4.6 toma la forma

$$U(Y_i|X) = U(X|Y_i) + U(Y_i) \quad (4.8)$$

Las funciones de energía  $U(X|Y_i)$  se calculan con base en las aproximaciones de la Fig. 4.1 (López, 2008; Moctezuma, 2014). Se debe tomar en cuenta que para cada función de energía conjunta,  $U(X|Y_i)$ , que depende del número de clases (referente al valor medio de gris elegido de un conjunto de pixeles de la imagen original),  $\omega_i$  con  $i = 1, 2, \dots, n$  elegidas, que hace alusión al pixel analizado, ya pertenece a una clase o etiqueta.

Se necesita una media para cada clase elegida, es decir, para cada rango de grises elegidos se necesita el valor promedio de ese conjunto, esto para hacer la segmentación de la imagen. Sobre el eje de las abscisas se tiene el nivel de grises desde 0 (aunque realmente se empiece con 1) al nivel 256 de gris; en este eje se tendrán especificadas las medias de las clases, separadas una distancia  $S = m_s - m_a$ , donde  $m_s$  es la media siguiente y  $m_a$  es la media actual, mientras que en el eje de las ordenadas, se tiene el valor de que corresponde a la energía que se trata de una probabilidad, por lo que su valor mínimo será 0 y el máximo 1; dependiendo de la clase se tendrá una pendiente positiva o una negativa (López, 2008).

- Para la pendiente positiva se tiene la siguiente función:

$$U = \frac{2}{S}X - \frac{2}{S}m - \frac{1}{2} \tag{4.9}$$

- Para la pendiente negativa se tiene:

$$U = -\frac{2}{S}X + \frac{2}{S}m + \frac{3}{2} \tag{4.10}$$

donde

$U$ : es la energía condicional

$X$ : es el pixel (con nivel de gris) analizado

$m$ : es la media de la clase

$S$ : es la separación de las medias

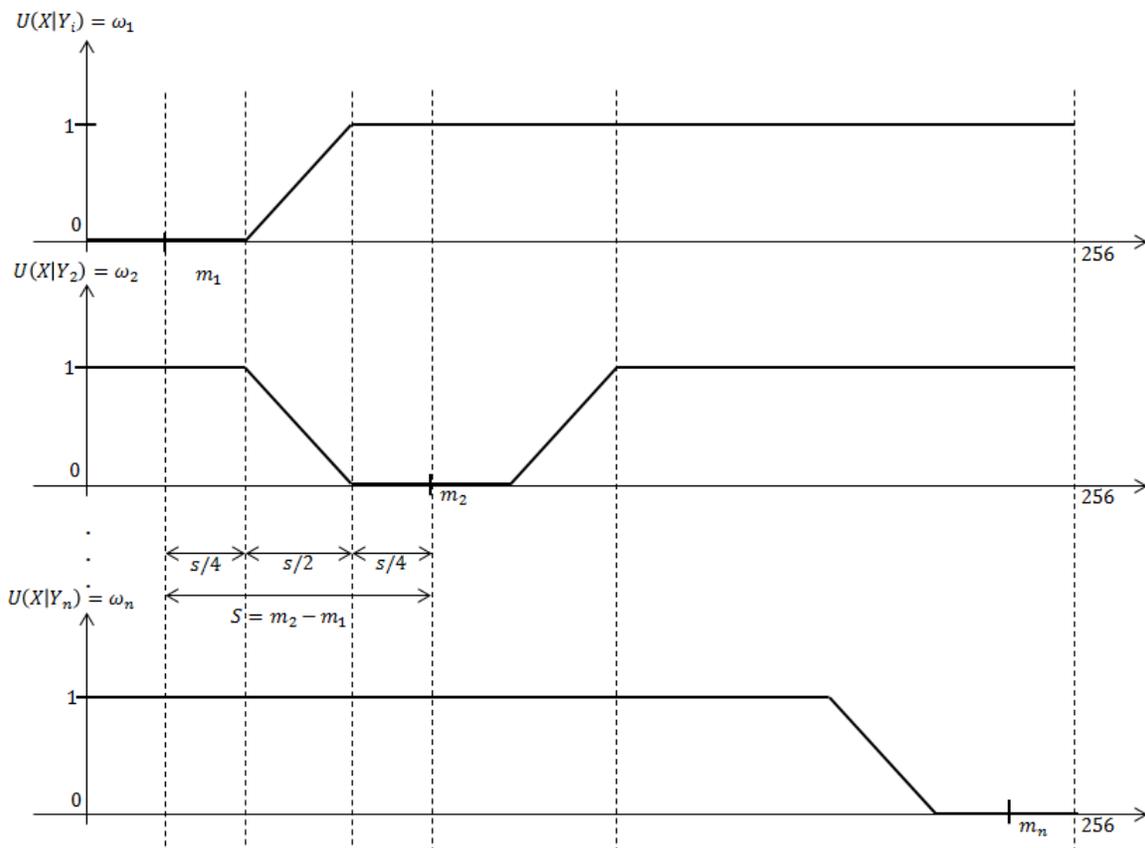


Fig. 4.1. Aproximaciones para las funciones de energía para las diferentes clases.

Para la energía de la clase  $U(Y_i)$  se aplica el modelo de Potts con un sistema de vecindad de 8-conectividad o de segundo orden, donde se aplica la sumatoria:

$$U(Y_i) = -\beta \sum_{(a,e) \in c} \delta(y_a, y_e) \quad (4.11)$$

la cual quiere decir que se suman todos aquellos sitios con la misma etiqueta, es decir, misma clase, con un coeficiente de iteración  $\beta$ , ya sea del pixel analizado o de algún otro elegido aleatoriamente con distribución uniforme.

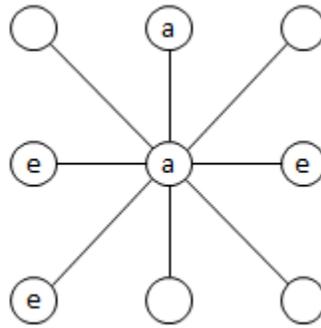


Fig. 4.2. Sistema de vecindad de 8-conectividad asociada al pixel analizado (centro de la imagen)

### 4.3 Esquema de Metropolis

El algoritmo de Metropolis inicialmente fue utilizado para “...simular el comportamiento de un sistema de partículas en equilibrio térmico a una temperatura  $T$ ...” (Li, 1995, p. 209). Li define el algoritmo de la siguiente manera (1995, p. 209):

*Inicialización de  $f$ ;*

*Repetir:*

*Generar una  $f' \in N(f)$ ;*

*$\Delta E \leftarrow E(f') - E(f)$ ;*

*$P = \min\{1, e^{-\Delta E/T}\}$ ;*

*Si aleatoriamente del conjunto  $[0,1) < P$  entonces  $f \leftarrow f'$ ;*

*Hasta que se alcanza el equilibrio*

*Regresar a  $f$*

El algoritmo anterior produce una cadena de Markov, ya que genera una secuencia de configuración donde en cada siguiente configuración,  $f'$  es escogida aleatoriamente de un sistema de vecindades  $N(f)$  de  $f$  aplicando una alteración como el cambio de etiquetas de una  $f_i$ .

Aleatoriamente es escogido un número del conjunto  $[0,1)$  con distribución uniforme. Si el número escogido anteriormente es menor a  $P$ , el cambio de la configuración es aceptado; por lo que

*...una nueva configuración  $f'$  para la cual  $\Delta E \leq 0$  es aceptada con seguridad, mientras que para  $\Delta E > 0$  es aceptado con probabilidad  $P = e^{-\Delta E/T}$ . [Por otro lado]...el muestreador de Gibbs genera la siguiente configuración basada en una probabilidad condicional en lugar de cambiar la energía. Un candidato  $f'_i$  para el siguiente  $f_i$  es aleatoriamente tomado de una distribución condicional  $P = (f'_i | f_{S-\{i\}})$  [esto se demuestra por una  $T$  fija, ya que el algoritmo de muestreo tiene una distribución de Gibbs como equilibrio]. (Li, 1995, p. 209).*

#### 4.4 Recocido Simulado

El Recocido Simulado, RS, es un algoritmo estocástico para la optimización combinatoria que, para este caso, funciona como una solución óptima al problema de segmentación en los MRF, ya que es necesario definir funciones de potencial clique como minimizar la función de energía *a posteriori*, esto quiere decir que minimiza la energía *a posteriori* para encontrar la imagen  $Y_n$ , con lo cual se encuentra la solución al problema de etiquetado.

El nombre de este algoritmo se debe a que simula el proceso de recocido de un sólido, es decir, el proceso de calentamiento a una temperatura determinada seguida de un descenso controlado de la temperatura, lo cual busca una configuración de baja energía.

El RS es usado para la búsqueda aleatoria controlada por una secuencia de temperatura descendente  $T$ . Con una  $T$  alta, el aumento en la energía  $E$  puede ser aceptado; mientras que para una  $T$  baja, solo un pequeño incremento puede ser aceptado y, cerca del punto de congelación de  $T$ , ningún incremento se acepta. Esto permite que el RS no sea un mínimo local sino un mínimo global.

En el algoritmo de recocido simulado se toma cualquier  $f$  del espacio de configuraciones  $\mathbb{F}$  con la siguiente probabilidad:

$$P_T = [P(f)]^{1/T} \quad (4.12)$$

donde  $T > 0$  es la temperatura del sistema. En el límite cuando  $T \rightarrow \infty$ ,  $P_T$  es una distribución uniforme en  $\mathbb{F}$ ; para  $T = 1$ ,  $P_T(f) = P(f)$  y cuando  $T \rightarrow 0$ ,  $P_T(f)$  es concentrado en el pico o picos de  $P(f)$  (Li, 1995; Winkler, 1995).

Li define el algoritmo del recocido simulado de la siguiente forma (1995, p 210):

*Inicializar en  $T$  y  $f$ ;*  
*Repetir*  
     *Muestras aleatorias  $f$  de  $N(f)$  bajo  $T$ ;*  
     *Decremento de  $T$ ;*  
*Hasta ( $T \rightarrow 0$ );*  
*Regresar  $f$ ;*

El RS aplica un algoritmo de muestreo sucesivamente, además que aplica valores decrecientes de la temperatura  $T$ . Debido a que este algoritmo sigue el proceso de recocido, se empieza por una  $T$  de un valor muy alto y se toma un conjunto de una configuración aleatoria  $f$ . Para una  $T$  fija, el muestreo se realiza de acuerdo a la distribución de Gibbs. Después que el algoritmo de Metropolis converge al equilibrio en el momento  $T$ ,  $T$  decrece. Lo anterior será repetido hasta que  $T$  tome un valor muy cercano a cero o cuando el sistema es congelado o cercano al mínimo de  $E(f)$ .

Para lograr que la función de energía *a posteriori* se minimice, es necesario que la temperatura baje por cada configuración, además con esto se asegura la convergencia a un mínimo global.

Se deben seguir las siguientes condiciones:

1. Referente a la convergencia de Metropolis, el cual menciona que si “...a cada configuración es visitada infinitamente a menudo, la distribución de la configuración generada es garantizada para converger en la distribución de Boltzman...” (Li, 1995, p. 210).
2. Referente al RS, se establece que si la secuencia decreciente satisface:

$$\lim_{t \rightarrow \infty} T^{(t)} = 0 \quad (4.13)$$

y:

$$T^{(t)} \geq \frac{m \times \Delta}{\ln(1 + t)} \quad (4.14)$$

donde  $\Delta = \max_f E(f) - \min_f E(f)$ , por lo que *el sistema converge a mínimo global independientemente de la configuración inicial  $f^{(0)}$*  (Li, 1995, p. 210).

## 4.5 Optimización

Para la optimización de la función de energía, con ayuda del recocido simulado, se considera el esquema de la Fig. 4.3 en la cual se aprecia que,  $X$  corresponde a la imagen original, ya en una escala de grises, mientras que  $Y$  es la imagen segmentada o, dicho de otra manera, es la imagen de clases desde  $Y^{(1)}$  hasta  $Y^{(n)}$ .

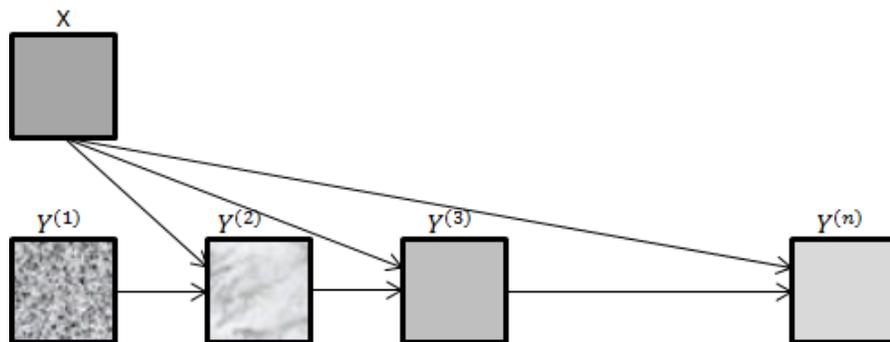


Fig. 4.3. Representación de la imagen original  $X$  y la imagen segmentada  $Y$ .

El algoritmo para la segmentación de la imagen con MRF utilizando la función de energía *a posteriori* optimizado por el recocido simulado se define por Moctezuma (2014):

### 1. Inicialización:

- Se determina aleatoriamente la solución  $Y^{(1)}$ , es decir que, se determina una clase para cada pixel de forma aleatoria con distribución uniforme en el rango del número de clases.
- Escoger una temperatura suficientemente alta,  $T = 2.5$

### 2. Ciclos de iteración que van desde *iteración* = 1 hasta *iteración* = $n$

#### 2.1 Se barre la imagen secuencialmente:

- En cada etapa  $k$  se escoge un pixel o sitio de coordenadas  $(i, j)$ , clase *actual* que su valor es  $Y_g^{(k)}$  con una vecindad  $V_g^{(k)}$  (como el mostrado en la Fig. 4.2).

- Se hace el sorteo de una variable aleatoria con distribución uniforme, que será la clase *elección*  $\varepsilon \in \Omega$ .

2.2 Se hace el criterio de Metrópolis y el Modelo de Potts para determinar la energía de la clase  $U(Y_i)$ :

- Se hace un análisis de 8-conectividad para el pixel en la imagen  $Y^{(k)}$

Tabla 4.1. Inicio del proceso del modelo de Potts.

Clases	Energía (inicialización)
actual elección	$E_{actual}=0$ $E_{elección}=0$
$U(Y) = -\beta \sum_{(a,e) \in \mathcal{C}} \delta(y_a, y_e)$	

- Si *actual* =  $x$  :  $E_{actual} = E_{actual} - \beta$
- Si *elección* =  $x$  :  $E_{elección} = E_{elección} - \beta$

$x$ : es el pixel analizado, ya sea el *actual* o *elección*.

2.3 Se calcula la variación de la energía local asociada al pixel analizado  $g$ .

- De la imagen original,  $X$ , en las coordenadas  $(i, j)$  se toma el nivel de gris del pixel analizado  $g$  y se proyecta en las funciones descritas en la Fig. 4.1, tanto de la clase *actual* como *elección*. Posteriormente sumar la contribución a la de Potts del punto 2.2.

- $E_{actual} = E_{actual} + Potencial_{actual}(i, j)$
- $E_{elección} = E_{elección} + Potencial_{elección}(i, j)$

donde  $Potencial_{actual}$  representa la energía condicional  $U(X|Y)$

- Después se hacen las siguientes condiciones:
  - Si  $E_{elección} < E_{actual}$  entonces:
    - $Y_{(i,j)} = elección$
  - En caso contrario:
    - $p = e^{-((E_{elección} - E_{actual})/T)}$

- Se sortea una variable aleatoria  $\varepsilon_2$ , uniformemente distribuida en el rango de  $[0,1]$
- Si  $\varepsilon_2 < p$  entonces:
  - $Y_{(i,j)} = \text{elección}$

3. Reinicio del proceso al punto 2.1, ya que se debe se hacer el barrido para cada uno de los pixeles

4. Decremento de la temperatura:

$$T_n = T_{n-1} \times 0.95$$

5. Se regresar al punto 2 para terminar el número de iteraciones fijadas.

# 5. Esquema de regularización adaptativa

En este capítulo se presenta un esquema de regularización adaptativa con el que se comprueba que, con la actualización del sistema de vecindades de segundo orden a uno de primer orden y con la elección de nuevas medias después de cierto número de iteraciones en el algoritmo de recocido simulado, se tiene una imagen segmentada con una mejor definición de la textura; por lo tanto se obtiene un mejor detalle sobre el mar abierto y las manchas del petróleo derramado.

## 5.1 Esquema general

Se demuestra que con el esquema de mejora propuesto sobre el convencional, visto en el Capítulo 4 de este trabajo, se tiene una mejor detección de la textura de la imagen, es decir, un mejor contorno sobre las manchas de petróleo generadas en el mar abierto y tener una mejor aproximación sobre la cantidad derramada. Esto se logra realizando una regularización estocástica donde se actualiza en un cierto número de ciclos del algoritmo el sistema de vecindades para la optimización de los MRF, además de la actualización de las medias.

Debido a que el criterio de MAP (Máximo A Posteriori) es un *estimador* para el procesamiento de imágenes, éste se emplea para la optimización de las funciones de probabilidad en los MRF (Moctezuma, Parmiggiani, 2014; Zoltan, 1994).

De acuerdo a la propiedad que tienen los MRF sobre tener un campo medible, es decir, que los datos son regulares en las iteraciones hechas por el algoritmo de recocido simulado, las medias y valores de gris de los pixeles siempre están en un conjunto medible, así que, los parámetros del modelo Gaussiano se pueden definir de una *ventana de formación* de la que se tienen como resultado medias  $\mu_i$  y con una varianza de cada clase aproximada como (Moctezuma, Parmiggiani, 2014):

$$\sigma_j^2 \approx (\mu_j - \mu_i/2)^2 \quad (5.1)$$

donde  $\mu_j > \mu_i$ .

Con la aproximación convencional de los parámetros de estimación,  $X$  representa la imagen original y con ruido, mientras que  $Y$  es la imagen sin ruido, es decir, que  $Y$  es el campo de etiquetas obtenido del proceso de segmentación con el vector de parámetros  $\theta$ .

Idealmente las ventanas de formación deben ser asignadas a las regiones del campo de etiquetas o valores de los grises, de donde se quiere tener una media en específico; pero la desventaja es que se "...utilizan datos cuyas funciones de distribución son la suma de varias funciones de distribución..." (Moctezuma y Parmiggiani, 2014, p. 5). Estas ventanas de formación son un conjunto de variables aleatorias, por lo que se da una representación estadística parcial sobre la clase que se desea determinar (Moctezuma y Parmiggiani, 2014).

Así que, se debe determinar un campo de etiquetado  $W$  con una aproximación MAP:

$$(\hat{W}, \hat{\theta}) = \arg \max_{W, \theta} p(W, X | \theta) \quad (5.2)$$

$(\hat{W}, \hat{\theta})$  son los parámetros de estimación (Moctezuma y Parmiggiani, 2014). Algunas soluciones a la ecuación 5.2 es utilizando algoritmos de maximización de expectativa y maximización de expectativa estocástica; sin embargo se proponen una solución simple a la ecuación 5.2 como:

$$\hat{W} = \arg \max_W p(W | \hat{\theta}) \quad (5.3)$$

$$\hat{\theta} = \arg \max_{\theta} p(\hat{W} | \theta) \quad (5.4)$$

La propuesta para la regularización estocástica es la actualización de los parámetros  $\hat{\theta}$  del modelo Gaussiano, ya que en las imágenes obtenidas del SAR las manchas generadas por el petróleo derramado difieren del mar abierto. Como resultado se obtiene un campo  $\hat{W}$  a partir de una segmentación adaptativa, es decir, que con cada actualización de los parámetros de las medias se tendrá un nuevo campo de etiquetas (Moctezuma y Parmiggiani, 2014).

## 5.2 Esquema adaptativo sobre el Recocido Simulado

Para dar solución al campo de etiquetado con una aproximación MAP se realizan modificaciones en el algoritmo de Recocido Simulado. Se tiene el siguiente pseudocódigo para optimizar las ecuaciones 5.3 y 5.4 (Moctezuma y Parmiggiani, 2014).

1. Se definen los parámetros de iniciación
  - a. Se estable la ventana de formación
  - b. Definición del vector  $\hat{\theta} = \{\mu_i, \sigma_i^2\}$  de la ventana de formación
  - c. Sistema de vecindad de segundo orden (8 vecinos)
  - d. Número de iteraciones  $n = 20$
2. Hacer las  $n$  iteraciones
3. Actualizar el vector  $\hat{\theta}$  del campo de etiquetado como resultado de la iteración  $n$ .
4. Se actualiza el sistema de vecindades
  - a. De segundo orden a uno de primer orden (4 vecinos)
  - b. Hacer  $n = n_2$
5. Se repiten los pasos 2 y 3 hasta la convergencia de  $\hat{\theta}$

## 6. Resultados

Se presentan los resultados aplicando tanto el modelo convencional como el propuesto sobre la regulación estocástica para dos imágenes del mismo lugar pero de diferente fecha.

## 6.1 Resultados utilizando el esquema convencional

Se aplica el algoritmo de recocido simulado sobre dos imágenes de los datos ASAR WR del ENVISAT con diferente fecha.

Para el procesamiento de las imágenes se utiliza el esquema convencional con las medias iniciales  $\mu_i = 71$  y  $\mu_j = 134$  para la Fig. 6.1 las cuales representan valores en la escala de grises que van de 1 a 256; además la varianza con valor  $\sigma_t^2 = 992.25$ . Se aplican 60 iteraciones debido a que considera una mejor convergencia para estas imágenes que con 50, es decir, se tiene una mejor textura y definición de la imagen; sin embargo, el procesamiento toma más tiempo.

Se establece una temperatura  $T$  de 2.5 con una disminución gradual en cada iteración de la forma  $T^n = qT^{n-1}$  con un valor de  $q = 0.95$  (para una convergencia más rápida) de acuerdo al criterio de Boltzman para el equilibrio térmico y una beta de 0.35 (Ludwin, 2008; Moctezuma, 2014); se utiliza una computadora TOSHIBA con procesador AMD Turion (tm II P520) Dual-Core Processor 2.30 GHz; memoria RAM de 3.74 GB utilizable; sistema operativo Windows 7 Home Premium de 64 bits y software MATLAB 7.8.0 (R2009a).

En la Fig. 6.1 se muestra la imagen original de nombre “mw0509.bmp”, la cual fue tomada el 9 de mayo de 2010.

Imagen original mw0509.bmp

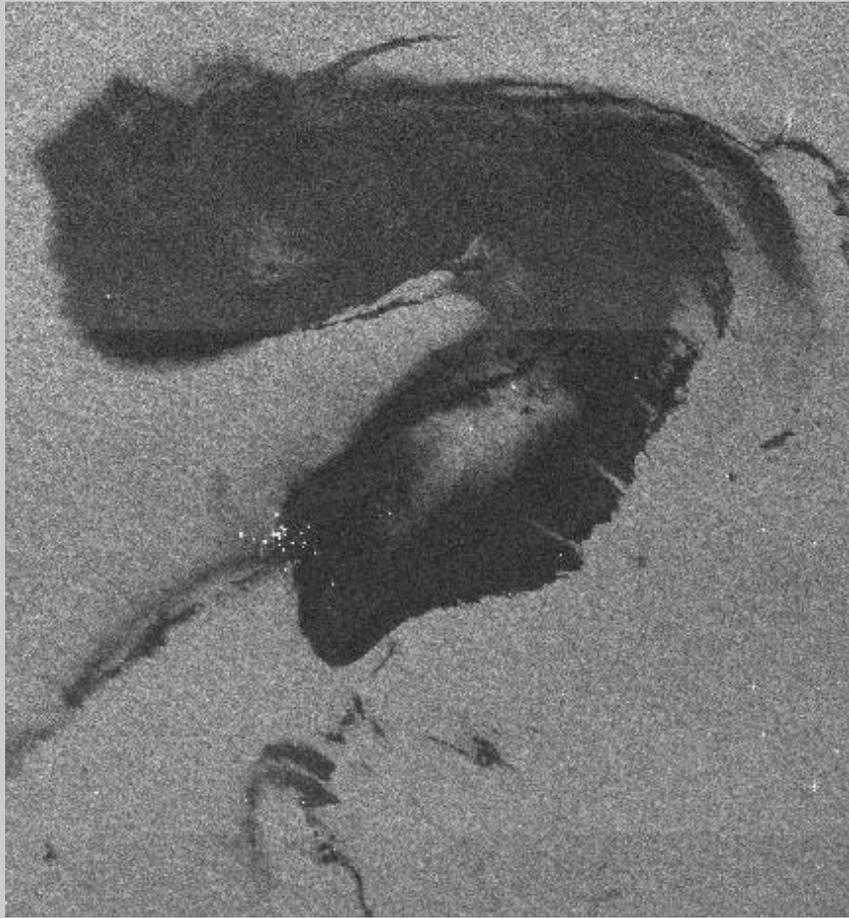


Figura 6.1. Imagen original mw0509.bmp (proporcionada por el Dr. Flavio Parmiggiani mediante el proyecto European Space Agency, Project C1F 1069).

En el modelo propuesto se hace una actualización en la iteración 20 del sistema de vecindades de uno de segundo orden (ocho vecinos) a uno de primer orden (cuatro vecinos), además de tomar dos nuevas medias (valores promedios de los grises de la imagen) para el procesamiento, por tal motivo, en la Fig. 6.2 se muestra la imagen mw0509.bmp (Fig. 6.1) segmentada en la iteración 20.

Imagen segmentada en la iteración n=20

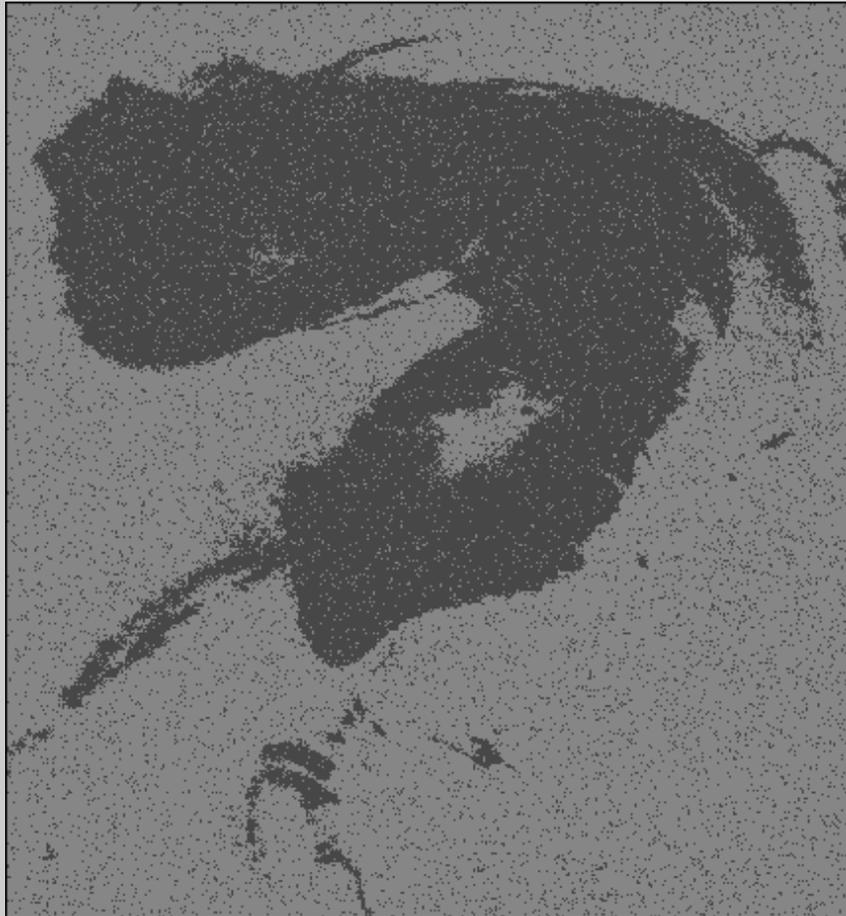


Fig. 6.2. Imagen segmentada en la iteración 20.

En la Fig. 6.3 se muestra el procesamiento de la imagen en la iteración 50, donde se pueden apreciar zonas con ruido, es decir, en el contorno de la mancha negra (representación del petróleo derramado) se puede apreciar puntos que provocan que no se tenga una buena definición del contorno de la mancha. Debido a lo anterior se decide realizar diez iteraciones más para obtener una mejor convergencia y tener una buena definición del contorno de la mancha de petróleo.

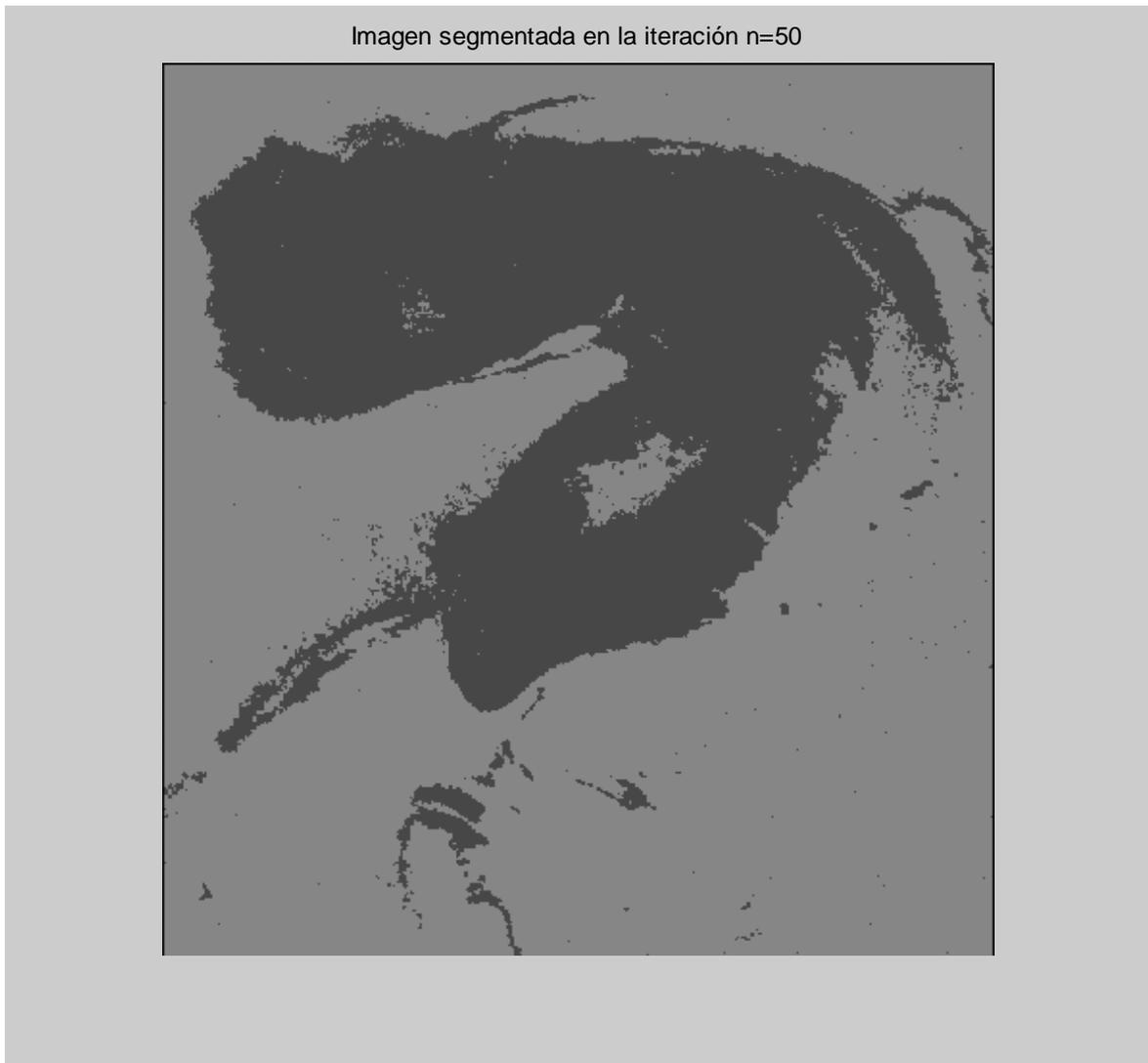


Fig. 6.3. Imagen segmentada en la iteración 50.

En la Fig. 6.4 se muestra la imagen segmentada hasta la iteración 60 con el modelo convencional de Recocido Simulado para dar solución al MRF. Se tiene como resultado una imagen con dos tonos de grises, el más oscuro representa el petróleo y el claro es el mar abierto.

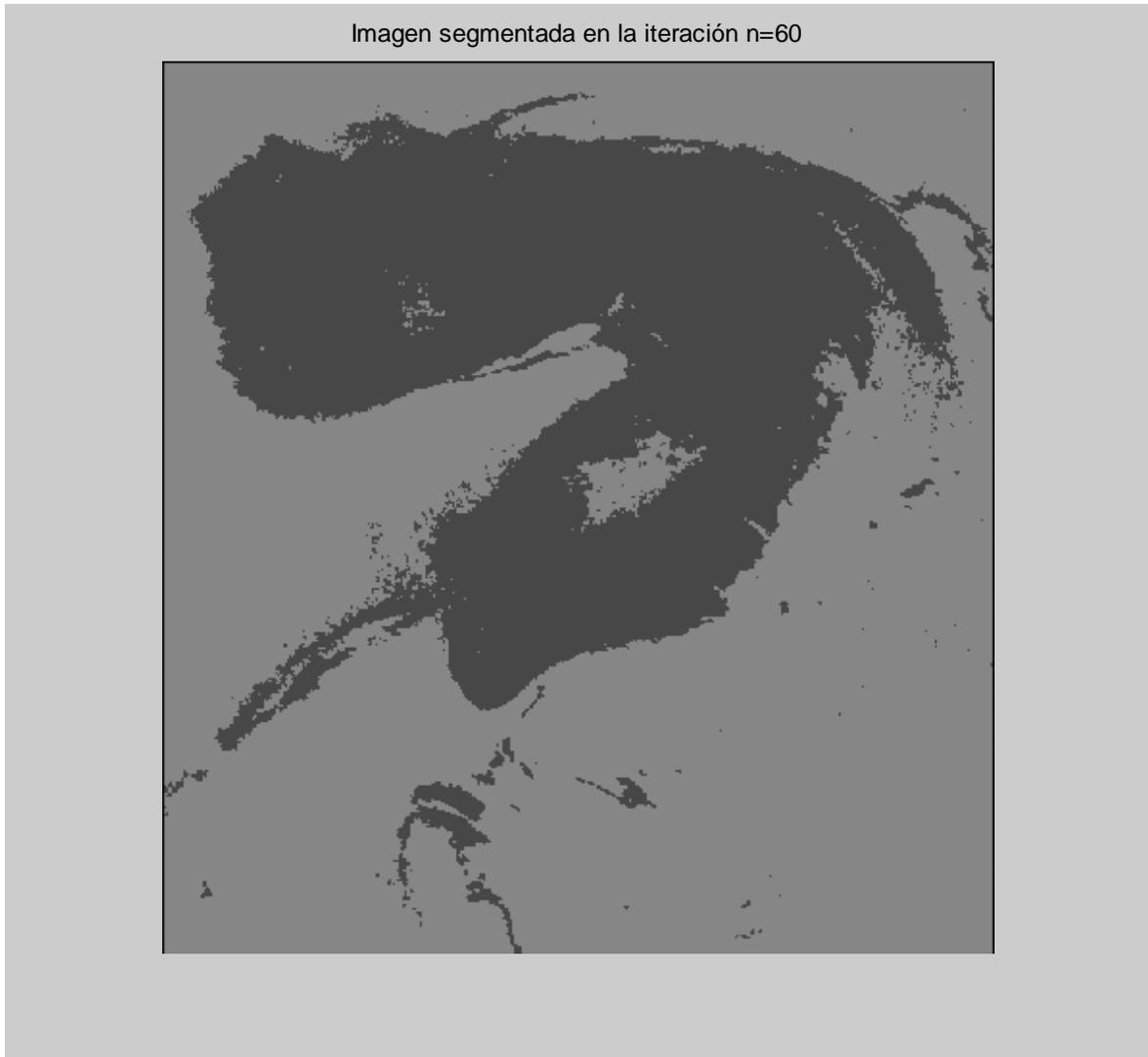


Fig. 6.4 Imagen segmentada en la iteración 60. Imagen resultado.

De la Fig. 6.4 se tiene los resultados esperados de acuerdo a lo que se requiere de la imagen original, que es una imagen segmentada con dos medias que separen lo que es mar abierto del petróleo derramado.

Se hace el proceso de segmentación para la imagen “mw0512.bmp” que se muestra en la Fig. 6.5 la cual se tomó el 12 de mayo de 2010.

Al igual que para la imagen de la Fig. 6.1 se hace la actualización en la iteración 20 mostrada en la Fig. 6.6. Para la imagen de la Fig. 6.5 se tienen las medias iniciales de  $\mu_i = 2$  y  $\mu_j = 55$  con varianza  $\sigma_i^2 = 702.25$ .



Fig. 6.5. Imagen mw0512.bmp original de la misma zona con diferente fecha (proporcionada por el Dr. Flavio Parmiggiani mediante el proyecto European Space Agency, Project C1F 1069).

Las medias iniciales fueron tomadas después de determinar el valor promedio de cada región mediante el software "ImageJ".

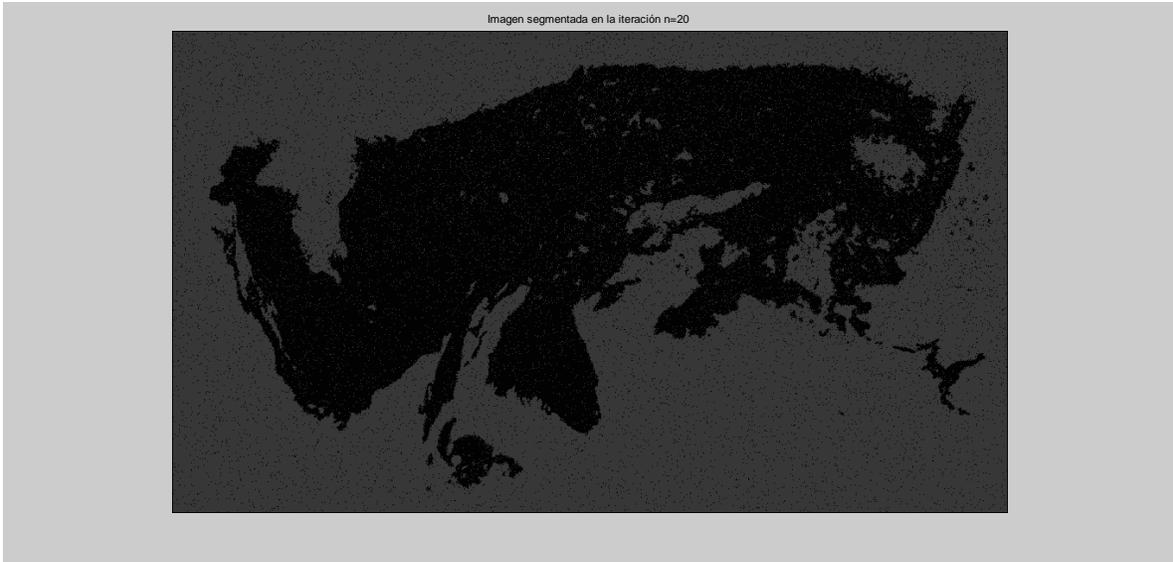


Fig. 6.6. Imagen segmentada en la iteración 20 con el modelo propuesto.

Ejecutando el proceso convencional, utilizando el algoritmo de recocido simulado con un sistema de vecindades de segundo orden y un proceso de 60 iteraciones, tenemos como resultado la imagen de la Fig. 6.7.

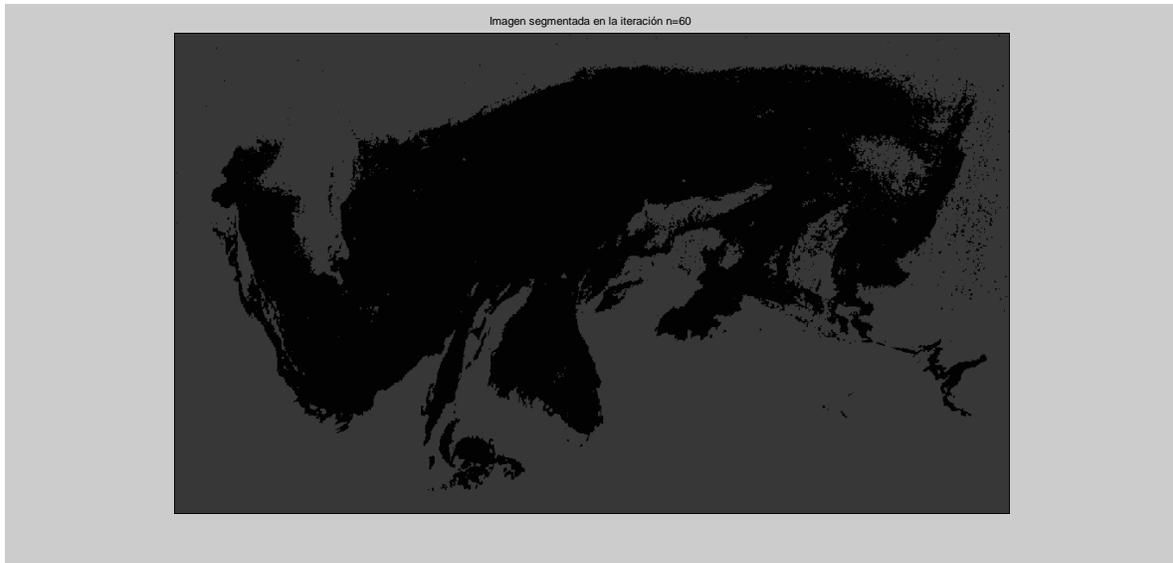


Fig. 6.7. Imagen segmentada con el criterio convencional de los MRF.

Hasta la iteración 60, aún se tienen zonas con muy pocos ruidos entre las manchas de petróleo y el mar abierto, haciendo referencia a los puntos que se aprecian a la derecha de la imagen.

## 6.2 Imágenes con el esquema adaptativo de Recocido Simulado

Para la iteración 20, tanto en el modelo convencional como en el modelo propuesto (esquema adaptativo) el resultado es el mismo, ya que se utilizan las medias iniciales y el sistema de vecindades es con ocho vecinos.

Después de la iteración 20 se utiliza la función *ginput* de Matlab (se elige un pixel de la imagen con ayuda del cursor) para elegir dos pixeles de la imagen, uno donde la zona oscura predomina y el otro donde el tono más claro de gris es el dominante, con el fin de definir la ventana formación teniendo al final una matriz de 20 por 20 pixeles; se hace el promedio de ambas matrices de pixeles para obtener dos nuevas medias.

Después de la actualización de la imagen *mw0509.bmp* se tiene los siguientes datos: medias  $\mu_i' = 80$  y  $\mu_j' = 163$  y varianza de  $\sigma_i'^2 = 1722.25$ . Se procede nuevamente a las iteraciones del algoritmo para converger en la iteración 60 con resultado mostrado en la Fig. 6.8.



Fig. 6.8. Imagen segmentada con el esquema adaptativo. Iteración 60.

Se hace una comparación de la imagen original (Fig. 6.1) con la segmentada mediante el modelo convencional (Fig. 6.4) y con el esquema adaptativo (Fig. 6.8) en la Fig. 6.9.

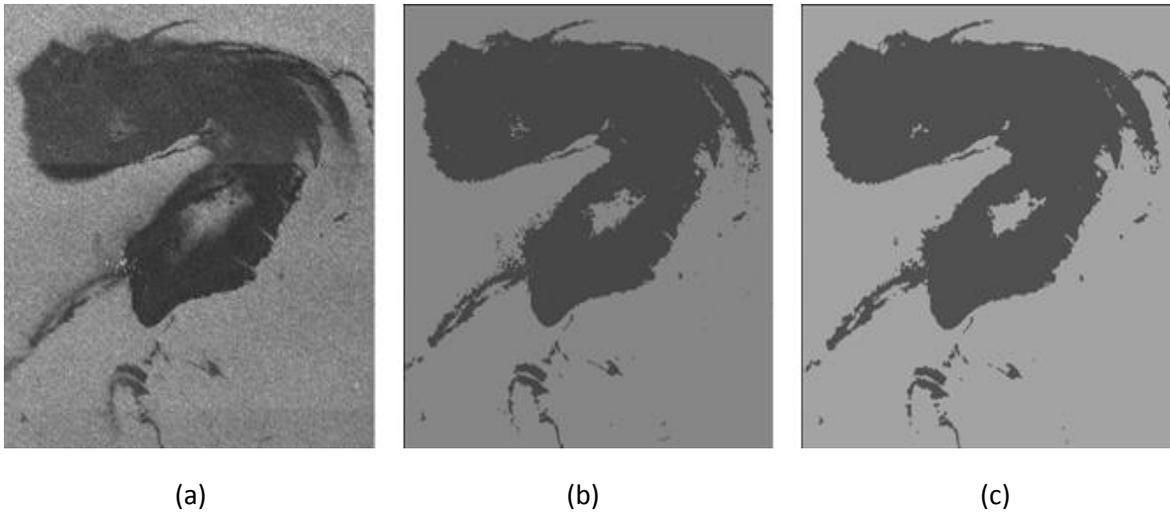


Fig. 6.9. (a) Imagen Original, (b) imagen segmentada con RS convencional e (c) imagen segmentada con esquema adaptativo.

En la Fig. 6.9 (c) se ve un cambio en los tonos debido a la actualización de las medias; además se puede apreciar un mejor contorno en la textura del petróleo, ya que en la Fig. 6.9 (b) algunas zonas de la franja entre en el mar abierto y el petróleo, por ejemplo, en la parte derecha superior, se aprecian algunos puntos del mismo tono que la mancha de petróleo y en la Fig. 3.9 (c) estos puntos desaparecen y se tiene un mejor contorno del derrame. El resultado es una mejor distinción entre el mar y el petróleo.

Se hace el comparativo del valor del petróleo derramado (píxeles con valor de la media menor) por metro cuadrado, tomando en cuenta que el valor de pixel es de 22500 m<sup>2</sup>. En la tabla 6.1 se muestran los resultados de la imagen “mw0509.bmp”.

Tabla 6.1. Resultados de la segmentación de la imagen mw0509.bmp del 9 de mayo de 2010.

Imagen mw0509.bmp		
Esquema	Cantidad de petróleo [km <sup>2</sup> ]	Diferencia [km <sup>2</sup> ]
Convencional	3 554.6175	14.13
Adaptativo	3 540.4875	

Para la segunda imagen mw0512.bmp, después de la iteración 20, se actualizan las medias y el sistema de vecindades a uno de 4-conectividad, las medias actualizadas tiene el valor de  $\mu_i' = 18$  y  $\mu_j' = 64$  y varianza de  $\sigma_i'^2 = 529$ . Se vuelve a ejecutar el algoritmo hasta la iteración 40 para tener un total de 60 iteraciones con el resultado mostrado en la Fig. 6.10.

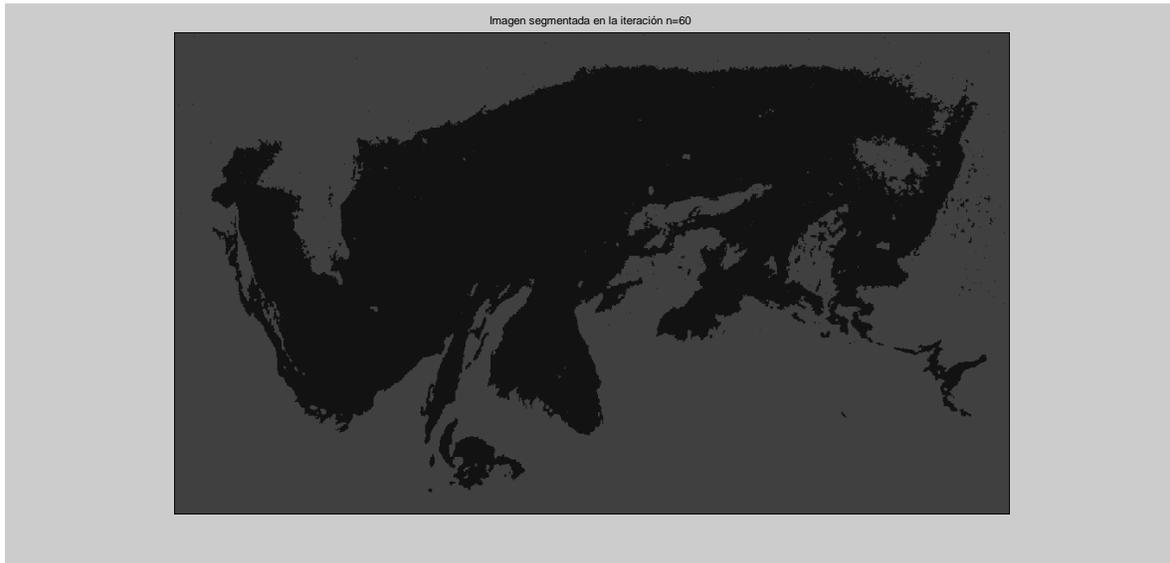


Fig. 6.10. Imagen segmentada en la iteración 60 con el modelo propuesto.

Se tienen resultados parecidos con la Fig. 6.7, sin embargo con el modelo propuesto se tiene una mejor definición de la textura y los contornos sobre las manchas de petróleo.

Finalmente se tiene una comparación sobre las tres imágenes (original, modelo convencional y con el esquema adaptativo) en la Fig. 6.11.

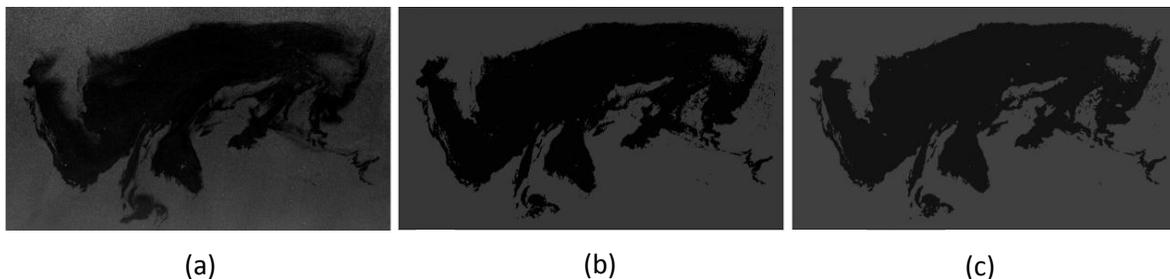


Fig. 6.11 (a) Imagen Original, (b) imagen segmentada con RS convencional e (c) imagen segmentada con esquema adaptativo

Al igual que en la imagen “mw0509.bmp” se hace el comparativo del valor del petróleo derramado por metro cuadrado, tomando en cuenta que el valor de pixel es de 22 500 m<sup>2</sup>. En la tabla 6.2 se muestran los resultados de la imagen “mw0512.bmp”.

Tabla 6.2. Resultados de la segmentación de la imagen mw0512.bmp del 12 de mayo de 2010.

Imagen mw0512.bmp		
Esquema	Cantidad de petróleo [km <sup>2</sup> ]	Diferencia [km <sup>2</sup> ]
Convencional	8 639.595	52.38
Esquema adaptativo	8 587.215	

En la tabla 6.3 se hace el comparativo de las dos imágenes con base en los datos del esquema adaptativo, ya que pertenecen a la misma zona pero con diferente fecha; además se aprecia que la diferencia del área del derrame fue grande con tan solo dos días de diferencia.

Tabla 6.3 Comparación del derrame de las dos imágenes con el esquema adaptativo.

Imagen	Cantidad de petróleo derramado por [km <sup>2</sup> ]	Diferencia [km <sup>2</sup> ]
mw0509.bmp (9 de mayo de 2010)	3 540.4875	5 046.7275
mw0512.bmp (12 de mayo de 2010)	8 587.215	

# 7. Conclusiones

Se describen las conclusiones sobre la segmentación de imágenes, obtenidas del satélite ambiental ENVISAT, utilizando el concepto de los Campos Aleatorios de Markov con una actualización adaptativa sobre la optimización estocástica del Recocido Simulado.

El número de iteraciones para el recocido simulado, para la optimización de los MRF, fueron de 60, aunque con 50 se tuvo el resultado esperado (separación del mar y petróleo con tonos de grises) aún se contaba con muy poco ruido, es decir, que en la textura de la imagen había zonas con ruido mostrado como *partes punteadas*; se tenía el tono más claro (tono de gris asignado al mar) y, sobre este, puntos (píxeles) con el tono más oscuro (tono de gris asignado al petróleo), sin que se tuviera una distinción del mar y la mancha de petróleo. Pero aún con este aumento en el número de iteraciones los cambios fueron muy pocos significativos, ya que permanecieron algunos puntos; además de que la solución siempre varía debido a que se trata de una asignación de variables aleatorias. Al ir aumentando el número de iteraciones el resultado fue muy parecido, aunque en algunas ocasiones hubo zonas que desaparecían y volvían aparecer después de cierto número de iteraciones debido código de optimización, comparando píxeles vecinos del analizado. Sin embargo, al aumentar el número de iteraciones se tuvo un mayor tiempo de procesamiento.

En el proceso de optimización por el algoritmo de recocido simulado, se detecta una dependencia sobre los valores de las medias iniciales (tonos de gris). Con diferentes valores iniciales de las medias se realizaron varias pruebas, definiendo medias individuales para cada imagen, ya que las dos presentaban características diferentes como los tonos de grises, el tamaño, el brillo y las zonas del derrame. Se utilizó el programa ImageJ para tener un valor promedio de las zonas que corresponden al mar y a petróleo; porque al realizar pruebas con medias muy separadas, es decir, tomar valores de gris cercanos a los límites del rango tanto inferior como superior (1 a 255), la segmentación después de algunas iteraciones producía errores al comparar los vecinos del píxel analizado, ya que la mancha que representaba al petróleo abarcaba un área mucho más grande de la que se muestra en la imagen original.

Para lograr una mejor aproximación sobre el área del derrame petrolero, se hace que los contornos sobre las texturas de la imagen sean más definidos y se tenga una mejor segmentación de los tonos de grises. Para lograr lo anterior se propone la actualización tanto de las medias (ventanas de formación de alguna zona con un tono de gris predominante) como del sistema de vecindades de segundo orden a uno de primer orden. Se debe tener en cuenta que por la propia naturaleza de la imagen no se puede tener una medición exacta, además de otros factores como el momento de tomar la imagen, el estado del clima en la zona (estado del mar), la cantidad de ruido de la imagen, definición, textura y la cantidad de datos (dimensiones de la imagen).

Aunque con la segmentación de la imagen se tenga una mejor aproximación del área del derrame, se debe tener en cuenta el factor de la profundidad que llegó a tener el petróleo en las tomas por el satélite ENVISAT, ya que como se puede apreciar en las imágenes segmentadas, la mancha de petróleo fue expandiéndose.

Con la segmentación mediante el esquema adaptativo sobre los MRF se tiene una mejor aproximación detallada sobre el derrame, ya que de acuerdo a los valores de la tabla 6.3 se muestra el impacto ambiental que causó el derrame, puesto que con sólo dos días de diferencia el área cubierta por el petróleo aumentó hasta un poco más de 5 000 km<sup>2</sup>.

Durante el procesamiento de las imágenes con los MRF y su optimización con el algoritmo de Recocido Simulado, se aplican conocimientos básicos de probabilidad hasta los específicos para la carrera de Ingeniería en Telecomunicaciones, como procesos estocásticos, programación y satélites. Resumiéndose en un sistema de transmisión, medio y recepción para el estudio de los datos recabados de este sistema.

Todo aprendizaje durante la carrera es aplicado directa o indirectamente; siendo así que, la base del conocimiento, en el ámbito ingenieril, es el aprendizaje de cada una de las materias cursadas, sabiduría transmitida del profesor al alumno, de los compañeros, de las experiencias y de uno mismo. Todo conocimiento adquirido de una u otra forma es el complemento de las demás. Por lo tanto, este trabajo es un conjunto de conocimientos adquiridos para su aplicación directa y simple a un problema en específico.

# Referencias

AATSR (s. f.). What is AATSR? Recuperado el 21 de junio de 2015 de <http://www.leos.le.ac.uk/aatsr/whatis/index.html>

Cambio Climático en México, Glosario de términos en cambio climático (2001). Recuperado el 21 de junio de 2015 de [http://cambio\\_climatico.ine.gob.mx/glosario.html](http://cambio_climatico.ine.gob.mx/glosario.html)

Canadian Space Agency (s. f.). RADARSAT-2 A Vital New Source of Global Geospatial Intelligence. Recuperado el 21 de junio de 2015 de [http://www.imagingnotes.com/go/article\\_free.php?mp\\_id=147](http://www.imagingnotes.com/go/article_free.php?mp_id=147)

Chellappa, Rama (1993). *Markov Random Fields: Theory and application*. Boston: Academic Press.

CLS, Collecte Localisation Satellites (s. f.). Radar Altimetry Tutorial, Laser RetroReflector. Recuperado el 21 de junio de 2015 de [http://www.altimetry.info/html/missions/envisat/instruments/lrr\\_en.html](http://www.altimetry.info/html/missions/envisat/instruments/lrr_en.html)

CRISP, Center for Remote Imaging, Sensing & Processing (s. f.). Medium Resolution Imaging Spectrometer (on ENVISAT Satellite, European Space Agency). Recuperado el 21 de junio de 2015 de <http://www.crisp.nus.edu.sg/~research/tutorial/meris.htm>

Curlander, John C. (1991). *Synthetic aperture radar: Systems and signal processing*. New York: J. E. Wiley.

De Samaniego Steta, Fernando (2007). Recuperado el 21 de junio de 2015 de [http://www.fisica.unam.mx/personales/naumis/index\\_archivos/Tesis/tesfin.pdf](http://www.fisica.unam.mx/personales/naumis/index_archivos/Tesis/tesfin.pdf)

El satélite Envisat, un “coloso” en el espacio (2002). Recuperado el 21 de junio de 2015 de <http://www.gobiernodecanarias.org/educacion/3/usrn/lentiscal/2-CD-Fiisca-TIC/2-1Gravitacion/Envisat.htm>

Encinas Castro Marco Antonio (2005). *Esquema de Segmentación de Imágenes de Percepción Remota: Aproximación Contextual y de Regularización* (Tesis para Maestría). Universidad Nacional Autónoma de México, México.

Estrada Medina, Juan Manuel (2009). *Procesos estocásticos: apuntes*. México, D.F.: UNAM.

European Space Agency (s. f.). Observing The Earth ERS Overview. Recuperado el 21 de junio de 2015 de [http://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/ERS\\_overview](http://www.esa.int/Our_Activities/Observing_the_Earth/ERS_overview)

European Space Agency Earth Online (s. f.). Recuperado el 21 de junio de 2015, de <https://earth.esa.int/web/guest/home>

Herrera Domínguez, Pável (s. f.). Construcción de imágenes de súper-resolución usando un método probabilista. Recuperado el 21 de junio de 2015 de <http://ccc.inaoep.mx/~esucar/Clases-mgp/Proyectos/finalPavelHerrera.pdf>

Ibe, Oliver C. (2009). *Markov processes for stochastic modeling*. Amsterdam: Elsevier Academic Press.

Kato, Zoltan (1994). *Modélisation markoviennes multirésolution en vision par ordinateur. Application à la segmentation d'images SPOT* (Tesis Doctoral). L'Université de Nice Sophia Antipolis.

Larson, Harold J. (1978). *Introducción a la teoría de probabilidades e interferencia estadística*. México: Limusa.

Li, S. Z. (1995). *Markov Random Field Modeling in Computer Vision*. London, UK: Springer-Verlag.

López López, Ludwin Ventura (2008). *Modelo Jerárquico Multitemperatura para Segmentación Markoviana* (Tesis Doctoral). Universidad Nacional Autónoma de México, México.

Moctezuma, Miguel & Parmiggiani Flavio (2014). *Adaptive stochastic minimization for measuring marine oil spill extent in synthetic apertura radar images*. Recuperado de <http://remotesensing.spiedigitallibrary.org>

Monti, Andrea; Prati Claudio (s. f.). SCANSAR FOCUSING AND INTERFEROMETRY. Recuperado el 3 de agosto de 2015 de <http://home.deib.polimi.it/monti/papers/monti96.pdf>

Obregon Sanin, Ivan (1975). *Teoría de la probabilidad*. México: Limusa.

Ratier, G. Levrini, A. Popescu y T. Paulsen (1999). GOMOS: Envisat's Contribution to Measuring Long-Term Trends in Ozone and Other Trace Gases. Recuperado el 21 de junio de 2015 de <http://www.esa.int/esapub/bulletin/bullet97/ratier.pdf>

Sharing Earth Observation Resource (s. f.). EnviSAT. Recupertado el 21 de junio de 2015 de <https://directory.eoportal.org/web/eoportal/satellite-missions/e/envisat>

Skolnik, Merrill I. (1988). *Radar applications*. New York: IEEE PRESS.

The ERS-1 and ERS-2 Satellite (s. f.). Recuperado el 21 de junio de 2015 de <http://southport.jpl.nasa.gov/polar/ers1.html>

The World Data Center for Remote Sensing of the Atmosphere (s. f.).  
SCIAMACHY, Scanning Imaging Absorption Spectrometer for Atmospheric  
Cartography. Recuperado el 21 de junio de 2015 de  
<http://wdc.dlr.de/sensors/sciamachy/>

Winkler, Gerhard (1995). *Image analysis, random fields, and dynamic Monte Carlo methods: a mathematical introduction*. New York: Springer.