

UNIVERSIDAD NACIONAL DE CÓRDOBA  
FACULTAD DE MATEMÁTICA, ASTRONOMÍA Y FÍSICA

---

SERIE "B"

# TRABAJOS DE MATEMÁTICA

Nº 50/06

## INTRODUCCIÓN AL PROCESAMIENTO DE IMÁGENES DIGITALES

*Oscar Bustos, Adriana Mallea, Myriam Herrera*



---

Editores: Jorge R. Lauret - Elvio A. Pilotta

CIUDAD UNIVERSITARIA – 5000 CÓRDOBA

REPÚBLICA ARGENTINA

# Introducción al Procesamiento de Imágenes Digitales

Oscar Bustos

Adriana Mallea

Myriam Herrera

2006

# Índice General

<b>1</b>	<b>Introducción</b>	<b>4</b>
1.1	Representación y modelación de imágenes . . . . .	5
1.2	Mejoramiento de Imágenes . . . . .	5
1.3	Restauración de Imágenes . . . . .	6
1.4	Análisis de Imágenes . . . . .	6
1.5	Reconstrucción de imagen desde las proyecciones . . . . .	6
1.6	Compresión de los Datos de una Imagen . . . . .	7
<b>2</b>	<b>Análisis de Imágenes y Visión</b>	<b>7</b>
<b>3</b>	<b>Extracción de Características Espaciales</b>	<b>8</b>
3.1	Características de los Rangos de Variación . . . . .	8
3.2	Características de Histograma . . . . .	8
3.3	Transformación para extraer características . . . . .	9
3.4	Detección de Bordes . . . . .	10
<b>4</b>	<b>Representación de Imágenes por Modelos Estocásticos</b>	<b>13</b>
<b>5</b>	<b>Modelos de Covarianza</b>	<b>17</b>
<b>6</b>	<b>Modelos Causales Unidimensionales</b>	<b>17</b>
6.1	Modelos Autorregresivos (AR) . . . . .	18
6.2	Representaciones en Promedios Móviles (MA) . . . . .	21
6.3	Representaciones Autorregresivas de Promedios Móviles (ARMA) . . . . .	21
<b>7</b>	<b>Predicción Lineal en dos dimensiones</b>	<b>22</b>
7.1	Predicción Causal . . . . .	22
7.2	Predicción Semicausal . . . . .	23
7.3	Predicción No Causal . . . . .	23
7.4	Predicción de Mínima Varianza . . . . .	24
7.5	Representación estocástica de Campos . . . . .	25
7.6	Representaciones de Mínima Varianza de orden finito . . . . .	26
<b>8</b>	<b>Principios de Información de Imágenes SAR</b>	<b>26</b>
8.1	Geometría de imágenes de radar . . . . .	27
8.2	Parámetros del sistema SAR . . . . .	32
8.3	La naturaleza de la información en imágenes SAR . . . . .	33
<b>9</b>	<b>Tipos de Imagen de un único canal</b>	<b>35</b>

<b>10 Estimación del RCS en datos multilook</b>	<b>39</b>
10.1 El Modelo de Ruido Multiplicativo para el RCS . . . . .	40
<b>11 Modelos para los datos</b>	<b>42</b>
11.1 Introducción . . . . .	42
11.2 El modelo producto . . . . .	42
11.3 El modelo <i>RCS</i> . . . . .	43
11.3.1 La FDP de la Intensidad . . . . .	44
<b>12 Clasificación de RCS y segmentación</b>	<b>44</b>
12.1 Clasificación de RCS . . . . .	45
12.2 El Modelo Cartoon y Segmentación . . . . .	45
12.3 Detección de Bordes . . . . .	46
12.4 Fusión de Regiones . . . . .	47
12.5 Teoría sobre fusión/división . . . . .	48
<b>13 Referencias</b>	<b>50</b>

# Introducción al Procesamiento de Imágenes Digitales

Oscar Bustos<sup>(1)</sup>-Adriana Mallea<sup>(2)</sup>- Myriam Herrera<sup>(2)</sup>

Universidad Nacional de Córdoba<sup>(1)</sup>- Universidad Nacional de San Juan<sup>(2)</sup>

## Resumen

Este trabajo consiste en una introducción a las técnicas de procesamiento de imágenes y de su representación por modelos estocásticos. También se proporciona una introducción a las técnicas de clasificación y segmentación de imágenes.

PALABRAS CLAVES: Procesos- Estocásticos- Clasificación- Segmentación.

## 1 Introducción

El Procesamiento de Imágenes Digitales generalmente se refiere al procesamiento de ciertas matrices hechas por un computador. El mismo tiene un extenso campo de aplicación en sensores remotos vía satélite o naves espaciales, transmisión de imágenes y almacenaje para aplicaciones de negocios o procesamiento médicos.

Las imágenes adquiridas por satélites son útiles para relevar recursos del suelo, en confección de mapas, predicción en cosechas agrícolas, crecimientos urbanos y otras aplicaciones. Podemos considerar los siguientes problemas que se abordan con el procesamiento de imágenes:

- Representación y modelación de imágenes
- Mejoramiento de imágenes
- Restauración de imágenes
- Análisis de imágenes
- Reconstrucción de imágenes
- Compresión de los datos dados por la imagen

## 1.1 Representación y modelación de imágenes

La representación de una imagen tiene que ver con caracterizaciones de la cantidad que cada píxel representa. Una imagen puede representar la luminosidad (radiación de la luz) de objetos en escena, la sección transversal de radar del blanco (imagen de radar), imágenes infrarrojo o el campo gravitacional en imágenes geofísicas.

En general cualquier función de dos dimensiones que lleva información puede ser considerada una imagen. La modelación de una imagen proporciona una descripción lógica o cuantitativa de las propiedades de esta función. Una importante consideración a tener en cuenta en la representación de imágenes es el criterio de fidelidad para medir la calidad de una imagen o la performance de una técnica de procesamiento.

El requerimiento fundamental para el procesamiento de una imagen es que ella debe ser muestreada y cuantificada. La razón de muestreo, número de píxeles por unidad de área, tiene que preservar la utilidad de la información en la imagen.

Un método clásico de representación de señales es la expansión de series ortogonales, tales como la serie de Fourier. Para imágenes, una representación análoga es posible vía funciones ortogonales dos dimensionales llamada imagen base. Luego cualquier imagen puede ser expresada por una suma ponderada de imágenes base. Varias características de las imágenes pueden ser estudiadas por tal expansión.

Modelos estadísticos describen una imagen como un elemento de un ensamble, frecuentemente caracterizado por su media y covarianza. Esto permite el desarrollo de algoritmos que son usados por una clase entera o ensamble de imágenes más que por una única imagen. Frecuentemente el ensamble se supone estacionario, lo que facilita la estimación de su media y covarianza. Para caracterizar propiedades locales de los píxeles, una alternativa puede ser caracterizar un píxel por su relación con píxeles vecinos. Por ejemplo en un modelo causal, un píxel está caracterizado por píxeles del “pasado”, en un modelo no causal por píxeles del pasado y futuro.

## 1.2 Mejoramiento de Imágenes

La idea del mejoramiento de imagen es acentuar ciertas características para un análisis posterior. Se usa para la extracción de características, análisis de imagen e información visual. Este proceso no aumenta la información contenida en los datos, sino que enfatiza ciertas características. Un ejemplo de técnica usada es el método de equalización de histograma, donde los niveles de gris son transformados tal que el nivel de gris de salida tiene distribución uniforme. Este es un método potente para mejorar los contrastes bajos de imagen.

### 1.3 Restauración de Imágenes

La restauración de imágenes se refiere a suprimir o minimizar degradaciones conocidas de una imagen. Incluye la supresión del borrono de imágenes degradadas por las limitaciones del sensor, filtrado del ruido y la distorsión por corrección geométrica y radiométrica. La imagen de un punto original es borronoada y degradada debido a la atmósfera, ruido térmico de componentes del sensor, proceso de formación de imágenes, etc. Si suponemos un modelo lineal, la imagen de un objeto puede ser expresada como

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y; \alpha, \beta) f(\alpha, \beta) d\alpha d\beta + \eta(x, y),$$

donde  $\eta(x, y)$  es el ruido aditivo,  $f(\alpha, \beta)$  es el objeto,  $g(x, y)$  es la imagen y  $h(x, y; \alpha, \beta)$  es llamada la función de propagación (spread) puntual PSF. El problema de restauración es la estimación de la función  $f(\alpha, \beta)$  dada PSF, la imagen “con ruido” y las propiedades estadísticas del ruido.

Un resultado fundamental en la teoría de filtro, usado para la restauración de imágenes, es el llamado filtro de Wiener. Este filtro da la mejor estimación lineal de cuadrados medios de un objeto a partir de las observaciones.

### 1.4 Análisis de Imágenes

El análisis de imágenes considera medidas cuantitativas de una imagen para producir una descripción de ella, produciendo información cuantitativa y se usa para tomar decisiones.

Las técnicas de análisis de imagen requieren la extracción de ciertas características que ayudan a la identificación del objeto. Las técnicas de segmentación son usadas para aislar el objeto deseado de la escena. Las medidas cuantitativas de las características permiten la clasificación y descripción de la imagen.

### 1.5 Reconstrucción de imagen desde las proyecciones

La reconstrucción de imagen desde las proyecciones es una clase especial de los problemas de restauración de imágenes donde un objeto de dimensión dos o mayor es reconstruido con proyecciones unidimensionales. Tales técnicas son importantes en imágenes médicas, astronomía, imágenes de radar y exploraciones geológicas.

Matemáticamente los problemas de reconstrucción de imágenes se basan en la teoría de transformaciones aleatorias.

## 1.6 Compresión de los Datos de una Imagen

La cantidad de datos asociada con la información visual es grande, requiere una enorme capacidad de almacenaje. Las técnicas de compresión de datos de imagen permiten una reducción del número de bits que se requieren para almacenar o transmitir imágenes sin pérdida de información. Entre la transmisión de imágenes tenemos las mostradas por televisión; sensores remotos vía satélite, avión, radar, sonar; teleconferencias. El almacenaje de imágenes es requerido comúnmente en documentos de educación, negocios, imágenes médicas usadas en pacientes con sistemas monitoreados. Por la gran variedad de aplicaciones la compresión de datos es de gran interés en el procesamiento de imágenes digitales.

## 2 Análisis de Imágenes y Visión

### Computacional

El objetivo esencial en un gran número de aplicaciones del procesamiento de imágenes es extraer *características* importantes de los datos de la imagen, para lo cual la computadora puede ayudarnos a hacer una descripción, interpretación, o comprensión de la escena. Por ejemplo, un sistema de visión puede distinguir partes sobre una línea de ensamble y listar sus características, tales como el número y tamaño de agujeros. Sistemas de visión más sofisticados son capaces de interpretar los resultados del análisis y describir varios objetos y sus relaciones en la escena.

El análisis de imágenes involucra básicamente el estudio de *extracción de características, segmentación y técnicas de clasificación*.

En los sistemas de visión computacional, la imagen de entrada es primero preprocesada, lo cual puede involucrar restauración, aumento o representación apropiada de los datos. Entonces se extraen ciertas características para la segmentación de la imagen en sus componentes, por ejemplo separación de diferentes objetos por extracción de sus bordes. La imagen segmentada es llevada a un clasificador o un sistema de compresión de la imagen. La clasificación de imágenes transforma diferentes regiones o segmentos (imagen satelital) en uno de varios objetos, cada uno identificado por una etiqueta (mapa temático).

Los sistemas de comprensión de imágenes determinan la relación entre los diferentes objetos de la escena para dar una descripción adecuada. Por ejemplo un sistema de comprensión de imágenes puede dar el siguiente reporte:

*la vista del campo contiene un camino de barro rodeado de hierba.*

Tales sistemas pueden permitir la clasificación de diferentes texturas tales como arena, hierba o grano, usando conocimientos a priori y reglas predefinidas para generar una descripción.



### 3 Extracción de Características Espaciales

Las características espaciales de un objeto pueden ser caracterizadas por sus niveles de gris (niveles digitales), sus distribuciones de probabilidad conjunta, distribución espacial y otras estadísticas.

#### 3.1 Características de los Rangos de Variación

Las características más simples y quizás las más útiles de un objeto son los rangos de variación de sus propiedades físicas, tal como reflectividad, transmisibilidad, o respuestas multiespectrales. Por ejemplo, en imágenes de rayos X en medicina, la variación de los niveles de gris representan las características de las masas del cuerpo y permiten discriminar tejidos óseos en tejidos sanos o enfermos. En imágenes infrarrojas los rangos de variación de la temperatura, facilitan la segmentación entre nubes y terrenos. En imágenes de radar la variación de la *sección transversal del radar*, o coeficiente de retrodispersión, determina el tamaño del objeto.

#### 3.2 Características de Histograma

Las características de Histograma se basan en el histograma de una región de la imagen. Sea  $U$  la variable aleatoria que representa un nivel de gris en una región dada de la imagen. Definimos

$$p_U(x) = P[U = x] = \frac{\text{nro. de pixeles con nivel gris } x}{\text{nro. total de pixeles en la región}} \quad x = 0, 1, \dots, L - 1. \quad (1)$$

Las características comunes de  $p_U(x)$  son sus momentos, momentos absolutos, la entropía dada por

$$H = E[-\log_2 p_U] = - \sum_{x=0}^{L-1} p_U(x) \log_2 p_U(x).$$

Algunas características comunes de histogramas son la *media*, *varianza*, *valor cuadrático medio o energía promedio* y *coeficientes de asimetría y de curtosis*. Otras características útiles son la mediana y la moda. Un histograma estrecho indica una región de bajo contraste. La varianza puede usarse para medir la actividad local en las amplitudes. Las características del histograma son también útiles para analizar la forma de objetos desde sus proyecciones.

Frecuentemente estas características son medidas sobre una pequeña ventana móvil  $W$ . Algunas de las características del histograma pueden medirse

sin determinar explícitamente el histograma, por ejemplo

$$m_i(k, l) = \frac{1}{N_W} \sum_{(m, n) \in W} [u(m - k, n - l)]^i,$$

$$\mu_i(k, l) = \frac{1}{N_W} \sum_{(m, n) \in W} [u(m - k, n - l) - m_1(k, l)]^i,$$

donde  $i = 1, 2, \dots$ , siendo  $m_1$  el momento no centrado de orden 1 y  $N_W$  el número de píxeles en la ventana  $W$ .

Las probabilidades conjuntas de segundo orden son útiles en aplicaciones tal como extracción de texturas. La probabilidad conjunta de segundo orden se define por

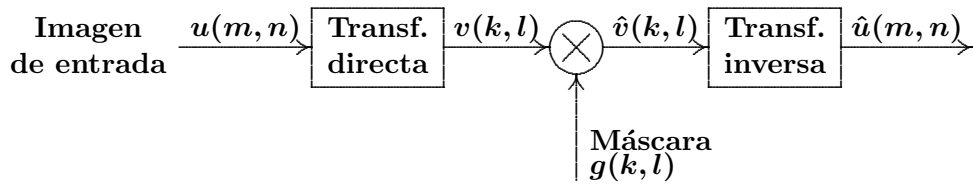
$$\begin{aligned} p_U(x_1, x_2) &= p_{U_1, U_2}(x_1, x_2) \\ &= P[U_1 = x_1, U_2 = x_2] \\ &= \frac{\text{nro. de pares de píxeles } u_1 = x_1, u_2 = x_2}{\text{nro. total de pares de píxeles en la región}}, \end{aligned}$$

$$x_1, x_2 = 0, \dots, L - 1$$

donde  $U_1$  y  $U_2$  son los dos píxeles en la región de la imagen especificados por alguna relación. Por ejemplo  $U_2$  puede especificarse como un píxel ubicado a distancia  $r$  y ángulo  $\theta$  de  $U_1$ . El arreglo  $L \times L \{p_U(x_1, x_2)\}$  se denomina “matriz de concurrencia”.

### 3.3 Transformación para extraer características

Las transformaciones de imagen proveen la información del dominio de frecuencia en los datos. La extracción de características se obtiene por filtrado-zonal de la imagen en el espacio de transformación seleccionado.



El filtro zonal es simplemente una ventana o rendija. Generalmente las características de alta frecuencia pueden usarse para la detección de lados y bordes, y las rendijas angulares pueden usarse para detección de orientación. Por ejemplo, una imagen que contiene varias líneas paralelas con orientación  $\theta$  exhibirá una fuerte energía a lo largo de una línea con ángulo  $\frac{\pi}{2} + \theta$  pasando

por el origen de su transformada de Fourier dos dimensional. Una combinación de una rendija angular con un filtro (en el dominio de las frecuencias) pasa baja o pasa alta, de banda limitada, puede ser usada para discriminación de texturas periódicas o cuasiperiódicas. Otras transformaciones tales como, Haar y Hadamard son potencialmente útiles para la extracción de características.

### 3.4 Detección de Bordes

Un problema de fundamental importancia en el análisis de imágenes es la detección de bordes. Los bordes caracterizan los límites de objetos y por lo tanto son útiles para segmentación, registro e identificación de objetos en escenas. Los puntos de borde pueden pensarse como las localizaciones de píxeles que cambian abruptamente los tonos de gris. Por ejemplo, es razonable definir puntos de borde en imágenes binarias como *píxeles negros con al menos uno blanco en algún entorno del mismo*, esto es, las ubicaciones de píxeles  $(m, n)$  tal que  $u(m, n) = 0$  y  $g(m, n) = 1$ , donde

$$g(m, n) = [u(m, n) \oplus u(m \pm 1, n)].O.[u(m, n) \oplus u(m, n \pm 1)], \quad (2)$$

donde  $\oplus$  denota la operación lógica O-excluyente.

Para una imagen continua  $f(x, y)$  su derivada toma un máximo local en la dirección del borde. Luego, una técnica de la detección del borde es medir el gradiente de  $f$  a lo largo de  $r$  en la dirección de  $\theta$ , esto es

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = f_x \cos \theta + f_y \sin \theta. \quad (3)$$

El máximo valor de  $\frac{\partial f}{\partial r}$  se obtiene cuando  $\frac{\partial}{\partial \theta} \frac{\partial f}{\partial r} = 0$ . Esto da

$$-f_x \sin \theta_g + f_y \cos \theta_g = 0 \Rightarrow \theta_g = \tan^{-1} \left( \frac{f_y}{f_x} \right), \quad (4)$$

$$\left( \frac{\partial f}{\partial r} \right)_{max} = \sqrt{f_x^2 + f_y^2}, \quad (5)$$

donde  $\theta_g$  es la dirección del borde. En base a estos conceptos, se pueden introducir dos tipos de operadores para detección de bordes, *operadores gradientes* y *operadores compass*. Para imágenes digitales estos operadores, también llamados *máscaras*, representan aproximaciones en diferencias finitas de o bien sus gradientes ortogonales  $f_x, f_y$  o el gradiente direccional  $\frac{\partial f}{\partial r}$ . Sea  $\mathbf{H}$  que denota una máscara  $p \times p$  y definimos, para una imagen arbitraria  $\mathbf{U}$ , su producto interior en la ubicación  $(m, n)$  como la correlación

$$\langle \mathbf{U}, \mathbf{H} \rangle_{m,n} = \sum_i \sum_j h(i,j)u(i+m, j+n). \quad (6)$$

### Operadores Gradiente

Estos son representados por un par de máscaras  $\mathbf{H}_1$  y  $\mathbf{H}_2$ , que miden el gradiente de la imagen  $u(m, n)$  en dos direcciones ortogonales.

Si definimos los gradientes direccionales  $g_1(m, n) = \langle \mathbf{U}, \mathbf{H}_1 \rangle_{(m,n)}$  y  $g_2(m, n) = \langle \mathbf{U}, \mathbf{H}_2 \rangle_{(m,n)}$ , la magnitud y dirección del vector gradiente se dan por

$$g(m, n) = \sqrt{g_1^2(m, n) + g_2^2(m, n)},$$

$$\theta_g(m, n) = \tan^{-1} \left( \frac{g_2(m, n)}{g_1(m, n)} \right).$$

Frecuentemente la magnitud del vector gradiente se mide con la norma  $L_1$ , dada por

$$g(m, n) = |g_1(m, n)| + |g_2(m, n)|,$$

en lugar de la norma  $L_2$ . Este cálculo es más fácil de implementar y se prefiere, especialmente cuando se implementa en hardware digital.

La ubicación  $(m, n)$  de un píxel es declarada ubicación de borde si  $g(m, n)$  excede algún umbral  $t$ . Las ubicaciones de los puntos de bordes constituyen una *función de borde*  $\epsilon(m, n)$  la cual está definida como

$$\epsilon(m, n) = \begin{cases} 1 & (m, n) \in I_g \\ 0 & \text{en otro caso} \end{cases}, \quad (7)$$

donde

$$I_g = \{(m, n) : g(m, n) > t\}.$$

La función de borde brinda los datos necesarios para trazar o marcar los bordes en una imagen. Generalmente  $t$  puede seleccionarse usando el histograma acumulativo de  $g(m, n)$  de modo que del 5 al 10% de píxeles con gradientes mayores son declarados como borde.

### Operadores de Laplace y cruces de ceros

Los métodos anteriores de estimación de gradientes funcionan bien cuando la transición de los niveles de gris es muy abrupta, semejante a una función escalonada. Cuando la región de transición se vuelve más amplia, es más ventajoso aplicar derivadas de segundo orden. Un operador frecuentemente usado es el operador Laplaciano, definido como

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

Por trabajar con derivadas de segundo orden, este operador es más sensitivo al ruido blanco que los previamente definidos. También, la magnitud de  $\nabla^2 f$  produce bordes dobles. Por esas razones, juntamente con su incapacidad para detectar la dirección del borde, el Laplaciano como tal no es un buen operador para la detección de bordes. Una mejor utilización del Laplaciano es usar sus cruces de ceros para detectar la ubicación de los mismos.

Una generalización del operador Laplaciano, que aproxima al Laplaciano de las funciones Gaussianas, es un potente detector de cruces de ceros. Se define como

$$h(m, n) = c \left[ 1 - \frac{(m^2 + n^2)}{\sigma^2} \right] \exp \left( -\frac{m^2 + n^2}{2\sigma^2} \right),$$

donde  $\sigma$  controla el ancho del núcleo Gaussiano y  $c$  normaliza a la unidad la suma de los elementos de una máscara de tamaño dado. Los cruces de ceros de una imagen convolucionados con  $h(m, n)$  dan sus ubicaciones de borde. Sobre una grilla bi-dimensional, se dice que ocurre un cruce de ceros cuando ocurre cruces de ceros en al menos una dirección.

$h(m, n)$  es la respuesta impulso muestreada de un filtro analógico de pasa-banda cuya respuesta de frecuencia es proporcional a  $(\xi_1^2 + \xi_2^2) \exp[-2\sigma^2(\xi_1^2 + \xi_2^2)]$ . Por lo tanto, el detector de cruces de ceros es equivalente a un filtro pasa-baja que tiene una respuesta de impulso Gaussiano, seguido por un operador de Laplace. El filtro pasa-baja sirve para atenuar la sensibilidad al ruido del Laplaciano. El parámetro  $\sigma$  controla la amplitud de la respuesta de salida del filtro pero no afecta la ubicación de los cruces de ceros.

La información direccional de los bordes puede obtenerse investigando los cruces de ceros de las derivadas de segundo orden a lo largo de  $r$  para cada dirección  $\theta$ . De (3), obtenemos

$$\frac{\partial^2 f}{\partial r^2} = \frac{\partial f_x}{\partial r} \cos \theta + \frac{\partial f_y}{\partial r} \sin \theta = \frac{\partial^2 f}{\partial x^2} \cos^2 \theta + \frac{\partial^2 f}{\partial y^2} \sin^2 \theta, \quad (8)$$

los cruces de ceros se buscan variando  $\theta$ .

## 4 Representación de Imágenes por Modelos Estocásticos

En las representaciones estocásticas una imagen se considera como una función muestral de un arreglo de variables aleatorias llamado un *campo aleatorio*. Esta caracterización de un ensamble de imágenes es útil para el desarrollo de técnicas de procesamiento de imágenes que son válidas para la clase completa y no sólo para una imagen en particular.

Los modelos estocásticos que se han usado en el procesamiento de imágenes pueden resumirse en:

- Modelos de covarianza
- Modelos uni-dimensionales (1-D)
- Modelos bi-dimensionales (2-D).

Aplicaciones de estos modelos se dan en compresión de datos de imagen, restauración de una imagen, síntesis y análisis de textura, estimación del espectro de potencia bi-dimensional, extracción de bordes de imágenes distorsionadas y otras varias situaciones.

Antes de displayarnos en estos modelos, daremos algunos conceptos previos.

### Campos Aleatorios Discretos

En la representación estadística de imágenes, cada píxel se considera como una variable aleatoria. Luego podemos pensar una imagen dada como una función muestral de un ensamble de imágenes. Tal ensamble puede ser adecuadamente definido por una densidad de probabilidad conjunta del arreglo de variables aleatorias. En la práctica, al analizar una imagen, el número de variables aleatorias es muy grande (262144 para imágenes 512x512). Luego es difícil especificar una función de densidad conjunta real. Una posibilidad es especificar el ensamble por sus momentos de primer y segundo orden solamente. Existen varios enfoques para modelar estocásticamente una imagen.

#### Definiciones

Cuando cada muestra de una secuencia bidimensional es una variable aleatoria, lo llamamos *Campo Aleatorio Discreto*. Cuando el campo aleatorio representa un campo de imágenes, tal como imágenes satelitales, lo llamaremos una *imagen aleatoria*.

Las funciones de media y covarianza de un campo aleatorio complejo se definen como:

$$\begin{aligned} E[u(m, n)] &= \mu(m, n) \\ Cov[u(m, n), u(m', n')] &= E[(u(m, n) - \mu(m, n))(u(m', n') - \mu^*(m', n'))] \quad (9) \\ &= r_u(m, n; m', n') = r(m, n; m', n'). \end{aligned}$$

Con frecuencia consideraremos el caso estacionario donde

$$\begin{aligned} \mu(m, n) &= \mu = \text{constante} \\ r(m, n; m', n') &= r(m - m', n - n'). \end{aligned} \quad (10)$$

Un campo aleatorio que satisface (10) también se denomina estacionario en sentido amplio, salvo que indiquemos lo contrario la estacionariedad será en este sentido. De lo anterior se deduce que

$$\begin{aligned} r(m, n) &= Cov[u(m, n), u(0, 0)] = Cov[u(m + m', n + n'), u(m', n')] \\ &\text{para todo } (m', n'). \end{aligned}$$

Un campo aleatorio  $\epsilon(m, n)$  se dice un *ruido blanco* cuando su media es constante (generalmente igual a cero) y cuando dos elementos  $\epsilon(m, n)$  y  $\epsilon(m', n')$  son no correlacionados, es decir su función de covarianza tiene la forma

$$r_\epsilon(m, n; m', n') = \sigma_\epsilon^2(m, n)\delta(m - m', n - n').$$

Un campo aleatorio es llamado *Gaussiano* si todos sus segmentos definidos sobre una grilla arbitraria finita es Gaussiana. Esto significa que todo segmento finito de  $u(m, n)$  puede transformarse en un vector que tenga densidad conjunta normal.

Las funciones de covarianza y de autocorrelación de campos aleatorios bi-dimensionales tienen propiedades de simetría y no negatividad similares a los procesos aleatorios de una dimensión:

- *Simetría:*  $r(m, n; m', n') = r^*(m', n'; m, n)$
- *No negatividad:*  $\sum_m \sum_n \sum_{m'} \sum_{n'} u(m, n)r(m, n; m', n')u^*(m', n') \geq 0,$   
 $u(m, n) \neq 0, \quad \text{para todo } (m, n).$

## La Función de Densidad Espectral

Sea  $u(n)$  una sucesión estacionaria. Su *función generadora de covarianzas* (FGC), denotada por  $S_u(z)$ , se define por la  $Z$ -transformada de su función de covarianza  $r_u(n)$ , es decir,

$$S_u(z) = \sum_{n=-\infty}^{\infty} r_u(n)z^{-n}. \quad (11)$$

La *función de densidad espectral* (FDE) se define como la transformada de Fourier de  $r_u(n)$ , que es la FGC evaluada en  $z = \exp(j\omega)$  y la denotaremos por  $S_u(\omega)$ , es decir

$$S_u(\omega) = \sum_{n=-\infty}^{\infty} r_u(n) \exp(-j\omega n). \quad (12)$$

La función de covarianza  $r_u(n)$  es simplemente la transformada de Fourier inversa

$$r_u(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_u(\omega) \exp(j\omega n) d\omega. \quad (13)$$

Para dos dimensiones la FGC y FDE tienen definiciones análogas:

$$S_u(z_1, z_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} r_u(m, n) z_1^{-m} z_2^{-n},$$

$$S_u(\omega_1, \omega_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} r_u(m, n) \exp[-j(\omega_1 m + \omega_2 n)],$$

$$r_u(m, n) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} S_u(\omega_1, \omega_2) \exp[j(\omega_1 m + \omega_2 n)] d\omega_1 d\omega_2.$$

Esto prueba que

$$\sigma_u^2 = r_u(0, 0) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} S_u(\omega_1, \omega_2) d\omega_1 d\omega_2,$$

es decir, el volumen bajo  $S_u(\omega_1, \omega_2)$  es igual a la potencia promedio en el campo aleatorio  $u(m, n)$ . Por lo tanto, físicamente,  $S_u(\omega_1, \omega_2)$  representa la densidad de potencia en la imagen en las frecuencias espaciales  $(\omega_1, \omega_2)$ , de allí que la FDE también se conozca como la función densidad del espectro de potencia o simplemente como el espectro de potencia del campo aleatorio. Frecuentemente el espectro de potencia se define como la transformada de Fourier de la autocorrelación en lugar de la covarianza. Salvo que se indique lo contrario continuaremos usando las definiciones en base a covarianzas.

Las propiedades que verifica la FDE son

1. La FDE es real:

$$S(\omega_1, \omega_2) = S^*(\omega_1, \omega_2) \quad \text{para todo } \omega_1, \omega_2.$$



2. La FDE es no negativa:

$$S(\omega_1, \omega_2) \geq 0, \quad \text{para todo } \omega_1, \omega_2.$$

Para campos aleatorios continuos las definiciones y propiedades para la FDE son similares.

## Funciones de Covarianza de Imágenes Separables e Isotrópicas

La función de covarianza de un campo aleatorio se llama *separable* cuando puede expresarse como el producto de funciones de covarianza de sucesiones de una dimensión, es decir, si

$$\begin{aligned} r(m, n; m', n') &= r_1(m, m')r_2(n, n') && \text{caso no estacionario} \\ r(m, n) &= r_1(m)r_2(n) && \text{caso estacionario.} \end{aligned} \quad (14)$$

Una función de covarianza estacionaria separable que se usa frecuentemente en procesamiento de imágenes es

$$r(m, n) = \sigma^2 \rho_1^{|m|} \rho_2^{|n|}, \quad |\rho_1| < 1, |\rho_2| < 1, \quad (15)$$

donde  $\sigma^2$  representa la varianza del campo aleatorio y  $\rho_1 = r(1, 0)/\sigma^2$  y  $\rho_2 = r(0, 1)/\sigma^2$  son las correlaciones (marginales) en las direcciones  $m$  y  $n$ , respectivamente.

Otra función de covarianza usada con frecuencia para muchas imágenes es la función *exponencial no separable*

$$r(m, n) = \sigma^2 \exp\left(-\sqrt{\alpha_1 m^2 + \alpha_2 n^2}\right). \quad (16)$$

Cuando  $\alpha_1 = \alpha_2 = \alpha$ ,  $r(m, n)$  resulta una función de la distancia euclídeana  $d = \sqrt{m^2 + n^2}$ , es decir

$$r(m, n) = \sigma^2 \rho^d,$$

donde  $\rho = \exp(-|\alpha|)$ . Una tal función se denomina *isotrópica* o *simétrica circularmente*. Los parámetros de una función de covarianza exponencial no separable se relacionan con las correlaciones marginales a través de  $\alpha_i = -\ln \rho_i$  para  $i = 1, 2$ . Luego los modelos de covarianza (15) y (16) pueden identificarse midiendo la varianza y las correlaciones marginales de sus campos aleatorios con media cero. En la práctica, esas cantidades se estiman a partir de los datos de una imagen dada usando los promedios muestrales; por ejemplo, para una imagen  $u(m, n)$ ,  $M \times N$ ,

$$\hat{\mu} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N u(m, n), \quad (17)$$

y

$$\hat{r}(m, n) = \frac{1}{MN} \sum_{m'=1}^{M-m} \sum_{n'=1}^{N-n} [u(m', n') - \hat{\mu}][u(m + m', n + n') - \hat{\mu}]. \quad (18)$$

Para muchas clases de imágenes,  $\rho_1$  y  $\rho_2$  son valores próximos a 0.95.

## 5 Modelos de Covarianza

En muchas aplicaciones tal como restauración de imágenes y compresión de datos es suficiente caracterizar un ensamble de imágenes por su función de media y covarianza. Generalmente uno comienza estableciendo una representación estacionaria del campo aleatorio donde la media es constante y la función de covarianza se representa por los modelos separables o exponencial no separable, definidos anteriormente.

Los modelos de covarianza son útiles para transformar codificación de imágenes, donde la función de covarianza se usa para determinar las varianzas de los coeficientes transformados. Los modelos de correlación con media y varianza que varían espacialmente, pero correlación espacialmente invariante han sido usados en técnicas de procesamiento block por block, en codificación de imágenes y problemas de restauración.

Una alternativa para representar campos aleatorios por sus funciones de media y covarianza, es caracterizarlos como las salidas de sistemas lineales cuyas entradas son campos aleatorios con propiedades estadísticas conocidas o deseadas (por ejemplo, con entradas de ruido blanco). Tales sistemas lineales se representan por ecuaciones en diferencia y son generalmente útiles para desarrollar algoritmos de procesamiento de imágenes computacionalmente eficientes. Además, los algoritmos adaptativos basados sobre la adaptación de los parámetros de las ecuaciones en diferencias son más fáciles de implementar que aquéllos basados en adaptar funciones de covarianza. El problema de encontrar un modelo en diferencia lineal estocástico que realice las covarianzas de un ensamble se conoce como el *problema de factorización espectral*.

## 6 Modelos Causales Unidimensionales

Una manera sencilla de caracterizar una imagen es considerarla como una señal 1-D que aparece como la salida de un scanner, esto es una sucesión de filas o columnas. Si se ignoran las dependencias entre filas o columnas, entonces los sistemas lineales 1-D son útiles para modelar tales señales.

Sea  $u(n)$  una sucesión aleatoria real y estacionaria con media cero y función de covarianza  $r(n)$ . Si  $u(n)$  se considera como la salida de un sistema lineal

estable  $H(z)$ , invariante por traslación, cuya entrada es una sucesión aleatoria estacionaria  $\epsilon(n)$  con media cero, esto es

$$u(n) = H(\epsilon(n)),$$

siendo  $H$  una función que es

- Lineal:  $H(a_1u_1(n) + a_2u_2(n)) = a_1H(u_1(n)) + a_2H(u_2(n))$
- Estable: la salida del sistema es uniformemente acotado para cualquier entrada acotada
- Invariante por traslación: la traslación de la entrada produce una traslación en la salida

La FDE de  $u(n)$  se da por

$$S(z) = H(z)S_\epsilon(z)H(z^{-1}), \quad z = e^{j\omega}, \quad -\pi < \omega \leq \pi$$

donde  $S_\epsilon(z)$  es la FDE de  $\epsilon(n)$ . Si  $H(z)$  es causal, entonces tiene serie de Laurent de una cola

$$H(z) = \sum_{n=0}^{\infty} h(n)z^{-n}$$

y todos sus polos deben estar dentro del círculo unitario.

## 6.1 Modelos Autorregresivos (AR)

Una sucesión aleatoria con media cero  $u(n)$  se denomina un proceso autorregresivo (AR) de orden  $p$  cuando puede ser generado como la salida del sistema

$$u(n) = \sum_{k=1}^p a(k)u(n-k) + \epsilon(n), \quad \text{para todo } n \quad (19)$$

$$E[\epsilon(n)] = 0, \quad E[(\epsilon(n))^2] = \beta^2, \quad E[\epsilon(n)\epsilon(m)] = 0, \quad m < n \quad (20)$$

donde  $\epsilon(n)$  es es una sucesión de entrada estacionaria con media cero que es independiente de salidas pasadas. Este sistema usa las  $p$  salidas más recientes y la entrada actual para generar recursivamente la próxima salida. Estos modelos adquieren especial significancia en procesamiento de señales e imágenes debido a que poseen varias propiedades importantes.

## Propiedades de Modelos AR

La cantidad

$$\bar{u}(n) = \sum_{k=1}^p a(k)u(n-k),$$

es el mejor predictor lineal de mínimos cuadrados de  $u(n)$  basado en todo su pasado pero depende sólo de las  $p$  muestras previas. Para sucesiones Gaussianas esto significa que una sucesión AR de orden  $p$  es un proceso de Markov. Luego (19) puede escribirse como

$$u(n) = \bar{u}(n) + \epsilon(n),$$

lo que implica que la muestra en  $n$  es la suma de su *estimación de predicción causal de mínima varianza* más el *error de predicción*  $\epsilon(n)$ , que también se denomina la *sucesión de innovaciones*.

El filtro causal definido por

$$A_p(z) = 1 - \sum_{n=1}^p a(n)z^{-n}, \quad (21)$$

se denomina el *filtro de error de predicción*, debido a que genera la sucesión de errores de predicción  $\epsilon(n)$ .

Excepto para posibles ceros en  $z = 0$ , la función de transferencia y la FDE de un proceso AR verifican

$$H(z) = \frac{1}{A_p(z)},$$

y

$$S(z) = \frac{\beta^2}{A_p(z)A_p(z^{-1})}, \quad z = e^{j\omega}, -\pi < \omega \leq \pi.$$

Para sucesiones con media  $\mu$ , el modelo AR puede modificarse como

$$\begin{aligned} x(n) &= \sum_{k=1}^p a(k)x(n-k) + \epsilon(n) \\ u(n) &= x(n) + \mu, \end{aligned} \quad (22)$$

donde las propiedades de  $\epsilon(n)$  son como antes.

## Identificación de Modelos AR

Multiplicando a ambos lados de (19) por  $\epsilon(n)$ , tomando esperanza y teniendo en cuenta (20), resulta que

$$E[u(n)\epsilon(n)] = \beta^2\delta(n - m), \quad m \geq n. \quad (23)$$

Ahora, multiplicando ambos miembros de (19) por  $u(0)$  y tomando esperanza, encontramos que el modelo AR satisface la relación

$$r(n) - \sum_{k=1}^p a(k)r(n - k) = \beta^2\delta(n), \quad \text{para todo } n \geq 0, \quad (24)$$

donde  $r(n)$  es la función de covarianza de  $u(n)$ . Este resultado es importante para la identificación de los parámetros del modelo AR  $a(k)$  y  $\beta^2$  para un conjunto de covarianzas  $\{r(n) : -p \leq n \leq p\}$ . De hecho, el modelo autorregresivo de orden  $p$  puede determinarse unívocamente resolviendo (24) para  $n = 0, \dots, p$ . En notación matricial, esto es equivalente a resolver las ecuaciones normales

$$Ra = r, \quad (25)$$

$$r(0) - a^T r = \beta^2, \quad (26)$$

donde  $R$  es la matriz Toeplitz  $p \times p$

$$R = \begin{bmatrix} r(0) & r(1) & \cdots & r(p-1) \\ r(1) & r(0) & \cdots & r(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(p-1) & \cdots & r(1) & r(0) \end{bmatrix} \quad (27)$$

y  $a = (a(1), a(2), \dots, a(p))^t, = (r(1), r(2), \dots, r(p))^t$ .

Mostramos con un ejemplo aplicaciones del modelo AR para representar líneas escaneadas de imágenes.

**Ejemplo 6.1** *La función de covarianza de una línea escaneada de una imagen puede obtenerse considerando la covarianza entre dos píxeles de la misma fila. Los modelos 2-D dados en (15) y (16) se reducen al modelo 1-D de la forma  $r(n) = \sigma^2\rho^{|n|}$ . Para ajustar a un modelo AR de orden 2, por ejemplo, resolvemos*

$$\sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \end{bmatrix} = \sigma^2 \begin{bmatrix} \rho \\ \rho^2 \end{bmatrix}$$

lo que da  $a(1) = \rho, a(2) = 0$  y  $\beta^2 = \sigma^2(1 - \rho^2)$ . La correspondiente representación para una línea escaneada de la imagen es un modelo AR de primer orden

$$\begin{aligned} x(n) &= \rho x(n-1) + \epsilon(n), & r_\epsilon(n) &= \sigma^2(1 - \rho^2)\delta(n) \\ u(n) &= x(n) + \mu, \end{aligned} \quad (28)$$

con  $A(z) = 1 - \rho z^{-1}$ ,  $S_\epsilon = \sigma^2(1 - \rho^2)$ , y  $S(z) = \sigma^2(1 - \rho^2)/[(1 - \rho z^{-1})(1 - \rho z)]$ , siendo  $\mu$  la media correspondiente.

## 6.2 Representaciones en Promedios Móviles (MA)

Una sucesión  $u(n)$  se denomina un proceso de promedios móviles (MA) de orden  $q$  cuando puede escribirse como un promedio pesado de variables aleatorias no correlacionadas

$$u(n) = \sum_{k=0}^q b(k)\epsilon(n-k), \quad (29)$$

donde  $\epsilon(n)$  es un proceso de ruido blanco con varianza  $\beta^2$ . La FDE de un MA se da por

$$\begin{aligned} S(z) &= \beta^2 B_q(z)B_q(z^{-1}) \\ B_q(z) &= \sum_{k=0}^q b(k)z^{-k}. \end{aligned} \quad (30)$$

Las relaciones anteriores permiten deducir que la covarianza de un proceso MA de orden  $q$  es cero fuera del intervalo  $[-q, q]$ .

## 6.3 Representaciones Autorregresivas de Promedios Móviles (ARMA)

Un modelo AR cuya entrada es una sucesión MA lleva a la representación

$$\sum_{k=0}^p a(k)u(n-k) = \sum_{l=0}^q b(l)\epsilon(n-l), \quad (31)$$

donde  $\epsilon(n)$  es un proceso de ruido blanco con varianza  $\beta^2$ . Esta se denomina representación ARMA de orden  $(p, q)$ . Su función de transferencia y FDE se dan por

$$H(z) = \frac{B_q(z)}{A_p(z)}, \quad (32)$$

$$S(z) = \frac{\beta^2 B_q(z) B_q(z^{-1})}{A_p(z) A_p(z^{-1})}. \quad (33)$$

Claramente para  $q = 0$  el modelo es un AR(p) y para  $p = 0$  es un MA(q).

## 7 Predicción Lineal en dos dimensiones

La noción de causalidad no se extiende naturalmente a dos o más dimensiones. Puesto que la causalidad no tiene importancia intrínseca en dos dimensiones, es natural considerar otras estructuras de los datos para caracterizar los modelos 2-D. Existen tres formas canónicas: causal, semi-causal y no causal, que consideraremos para la predicción lineal.

Estos tres tipos de modelos estocásticos tienen aplicaciones en muchos problemas de procesamiento de imágenes.

Los modelos de predicción lineal considerados aquí pueden también usarse efectivamente para factorización espectral 2-D y estimación espectral.

Sea  $u(m, n)$  un campo aleatorio estacionario con media cero y covarianza  $r(k, l)$ . Sea  $\bar{u}(m, n)$  que denota una predicción lineal de  $u(m, n)$ , definida por

$$\bar{u}(m, n) = \sum_{(k,l) \in \hat{S}_x} a(k, l) u(m - k, n - l), \quad (34)$$

donde  $a(k, l)$  son llamados coeficientes del predictor y  $\hat{S}_x$ , es un subconjunto del reticulado 2-D, llamada la región de predicción.

Las muestras incluidas en  $\hat{S}_x$  dependen del tipo de predicción considerada: causal ( $x = 1$ ), semicausal ( $x = 2$ ) y no causal ( $x = 3$ ). Con un mecanismo de escaneo hipotético que *escanea secuencialmente desde la parte superior a la inferior y de izquierda a derecha*, las tres regiones de predicción pueden definirse como sigue.

### 7.1 Predicción Causal

Un *predictor causal* es aquel cuya región de predicción (causal) es

$$\hat{S}_1 = \{(k, l) : l \geq 1\} \cup \{(k, l) : l = 0, k \geq 1\}$$

La definición de causalidad incluye el caso especial de predictores causales de un único cuadrante

$$\bar{u}(m, n) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a(k, l) u(m - k, n - l), \quad (35)$$

llamado *predictor fuertemente causal*.

## 7.2 Predicción Semicausal

Un *predictor semicausal* es causal en una de las coordenadas y no causal en la otra. Por ejemplo, la región de *predicción semicausal* que es causal en  $n$  y no causal en  $m$  es

$$\hat{S}_2 = \{(k, l) : l \geq 1\} \cup \{(k, l) : l = 0, k = 0\}.$$

## 7.3 Predicción No Causal

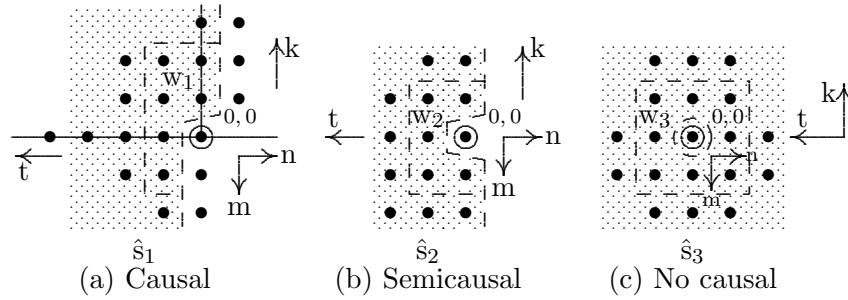
Un *predictor no causal*  $\bar{u}(m, n)$  es una función de todas las posibles variables en el campo aleatorio excepto  $u(m, n)$ . La región de predicción no causal es

$$\hat{S}_3 = \{(k, l) : (k, l) \neq (0, 0)\}.$$

En la práctica sólo una vecindad finita  $\hat{W}_x \subseteq \hat{S}_x$ , llamada ventana de predicción, puede usarse en el proceso de predicción, de modo que

$$\bar{u}(m, n) = \sum_{(k, l) \in \hat{W}_x} a(k, l) u(m - k, n - l), \quad (m, n) \in \hat{S}_x, \quad x = 1, 2, 3. \quad (36)$$

Algunas ventanas comúnmente usadas son



Causal:  $\hat{W}_1 = \{(k, l) : -p \leq k \leq p, 1 \leq l \leq q\} \cup \{(k, l) : 1 \leq k \leq p, l = 0\}$ .

Semicausal:  $\hat{W}_2 = \{(k, l) : -p \leq k \leq p, 0 \leq l \leq q, (k, l) \neq (0, 0)\}$ .

No causal:  $\hat{W}_3 = \{(k, l) : -p \leq k \leq p, -q \leq l \leq q, (k, l) \neq (0, 0)\}$ .

También definimos

$$W_x = \hat{W}_x \cup (0, 0), \quad x = 1, 2, 3.$$



**Ejemplo 7.1** *Los siguientes son tres ejemplos de predictores causales, semicausales y no causales, respectivamente.*

*Causal:*

$$\bar{u}(m, n) = a_1 u(m-1, n) + a_2 u(m, n-1) + a_3 u(m-1, n-1).$$

*Semicausal:*

$$\bar{u}(m, n) = a_1 u(m-1, n) + a_2 u(m+1, n) + a_3 u(m, n-1).$$

*No causal:*

$$\bar{u}(m, n) = a_1 u(m-1, n) + a_2 u(m+1, n) + a_3 u(m, n-1) + a_4 u(m, n+1).$$

## 7.4 Predicción de Mínima Varianza

Dada una región de predicción para formar una estimación de  $\bar{u}(m, n)$ , los coeficientes  $a(m, n)$  se pueden determinar usando el criterio de *predicción de mínima varianza*. Este criterio requiere que la varianza del error de predicción se minimice, esto es,

$$\beta^2 = \min E[\epsilon^2(m, n)], \quad \epsilon(m, n) = u(m, n) - \bar{u}(m, n).$$

La condición de ortogonalidad asociada con esta predicción de mínima varianza es

$$E[\epsilon(m, n)u(m-k, n-l)] = 0, \quad (k, l) \in \hat{S}_x, \text{ para todo } (m, n). \quad (37)$$

Usando las definiciones de  $\epsilon(m, n)$  y  $\bar{u}(m, n)$ , esto lleva a

$$E \left\{ \left[ u(m, n) - \sum_{(i,j) \in \hat{W}_x} a(i, j)u(m-i, n-j) \right] u(m-k, n-l) \right\} = \beta^2 \delta(k, l),$$

$$(k, l) \in \hat{S}_x, \text{ para todo } (m, n), \quad (38)$$

donde

$$S_x = \hat{S}_x \cup (0, 0), \quad x = 1, 2, 3,$$

por la cual obtenemos

$$r(k, l) - \sum_{(i,j) \in \hat{W}_x} a(i, j)r(k-i, l-j) = \beta^2 \delta(k, l),$$

$$(k, l) \in \hat{S}_x. \quad (39)$$

La solución de las ecuaciones simultáneas anteriores dan los coeficientes del predictor  $a(i, j)$  y la varianza del error de predicción  $\beta^2$ . Usando la propiedad de simetría de la covarianza, puede deducirse de la ecuación anterior que

$$\begin{aligned} a(-i, 0) &= a(i, 0) && \text{para predictores causales} \\ a(-i, -j) &= a(i, j) && \text{para predictores no causales} \end{aligned} \quad (40)$$

## 7.5 Representación estocástica de Campos Aleatorios

En general, el campo aleatorio  $u(m, n)$  puede caracterizarse como

$$u(m, n) = \bar{u}(m, n) + \epsilon(m, n), \quad (41)$$

donde  $\bar{u}(m, n)$  es una predicción arbitraria de  $u(m, n)$  y  $\epsilon(m, n)$  es otro campo aleatorio tal que (41) realiza las propiedades de covarianza de  $u(m, n)$ . Existen tres tipos de representaciones que son de interés en el procesamiento de imágenes:

1. Representaciones de Mínima Varianza (MVR)
2. Representaciones conducidas por Ruido Blanco (WNDR)
3. Representaciones Autorregresivas en Promedios Móviles (ARMA)

Para las representaciones de mínima varianza,  $\bar{u}(m, n)$  se elige como el predictor de mínima varianza y  $\epsilon(m, n)$  es el error de predicción. Para WNDR  $\epsilon(m, n)$  se elige como un ruido blanco. En representaciones ARMA,  $\epsilon(m, n)$  es un promedio móvil en dos dimensiones, esto es, un campo aleatorio con una función de covarianza truncada:

$$E[\epsilon(i, j)\epsilon(m, n)] = 0, \quad \forall(i, j) : |i - m| > K, |j - l| > L,$$

para enteros fijos  $K > 0$  y  $L > 0$ .

**Ejemplo 7.2** *La función de covarianza*

$$r(k, l) = \delta(k, l) - \alpha[\delta(k - 1, l) + \delta(k + 1, l) + \delta(k, l - 1) + \delta(k, l + 1)]$$

*representa un campo aleatorio MA. Su FDE es un polinomio de orden finito, bi-dimensional dado por*

$$S(z_1, z_2) = 1 - \alpha(z_1 + z_1^{-1} + z_2 + z_2^{-1}), \quad |\alpha| > 1/4.$$

## 7.6 Representaciones de Mínima Varianza de orden finito

Los predictores finitos de (36) llevan a MVR de la forma

$$u(m, n) = \sum_{((k,l) \in \hat{W}_x)} a(k, l) u(m - k, n - l) + \epsilon(m, n). \quad (42)$$

Un campo aleatorio caracterizado por esta MVR debe satisfacer (37) y (38). Multiplicando ambos lados de (42) por  $\epsilon(m, n)$ , tomando esperanza y usando (37), obtenemos

$$E[u(m, n)\epsilon(m, n)] = \beta^2.$$

Usando los resultados anteriores encontramos

$$\begin{aligned} r_\epsilon(k, l) &= E[\epsilon(m, n)\epsilon(m - k, n - l)] \\ &= E[\epsilon(m, n) (u(m - k, n - l) - \bar{u}(m - k, n - l))] \\ &= \beta^2 \left[ \delta(k, l) - \sum_{((i,j) \in \hat{W}_x)} a(i, j) \delta(k - i, l + j) \right], \end{aligned} \quad (43)$$

con  $a(0, 0) = -1$  y usando (40), la función de covarianza del error de predicción  $\epsilon(m, n)$  es

$$r_\epsilon(k, l) = \begin{cases} \beta^2 \delta(k, l), & \text{p.t. (k,l) para MVR causal} \\ -\beta^2 \delta(l) \sum_{i=-p}^p a(i, 0) \delta(k + i), & \text{p.t. (k,l) para MVR semicausal} \\ -\beta^2 \sum_{i=-p}^p \sum_{j=-q}^q a(i, j) \delta(k - i) \delta(l - j), & \text{p.t. (k,l) para MVR no causal} \end{cases}$$

## 8 Principios de Información de Imágenes SAR

Cuando intentamos extraer información de imágenes de *radar de apertura sintética* (SAR), debemos distinguir dos tipos de propiedades de las imágenes. La más importante es dónde las propiedades de la escena ejemplo, su geometría, su movimiento, etc. producen efectos en la imagen; por lo tanto un examen de la imagen puede proveer información acerca de la escena. El segundo tipo se genera exclusivamente por el sistema y procesamiento de señales.

Nuestro objetivo se centrará en establecer un modelo ideal para una imagen SAR y mostrar que muchas propiedades intrínsecas se determinan por un pequeño número de elecciones en el diseño del sistema.

El SAR es esencialmente un artificio complicado cuyo propósito es simplemente medir la interacción local entre la superficie de la Tierra y una onda incidente. Una mayor calidad de imágenes SAR de esta interacción puede ser producida por sistemas modernos.

Los métodos para extraer información de datos SAR deben tener en cuenta el conocimiento de cómo es el mundo real. Para esto es fundamental establecer modelos viables para los datos, en los cuales la información pueda ser relacionada a parámetros medibles y su disposición espacial. Una dificultad primaria al tratar de obtener información de imágenes SAR es que todas las mediciones o inferencias a partir de los datos deben tener en cuenta el ruido (speckle). Estudiaremos propiedades básicas del speckle y cómo se manifiestan en diferentes tipos de imágenes que pueden formarse cuando se obtienen datos SAR complejos. La teoría para el speckle de un único canal permite un modelo para los datos en el cual la información es transportada por la intensidad promedio o RCS (coeficiente de retrodispersión) en cada píxel afectado por el ruido. No obstante, la información contenida por píxel es muy baja dentro de los blancos extensos, puesto que en este caso un gran número de píxeles deben ser caracterizados por un pequeño número de parámetros. Aumentar el número de píxeles sirve para mejorar la estimación del parámetro. Con este objetivo analizaremos el proceso llamado “multilooking” que reduce la resolución espacial a fin de mejorar la estimación del RCS en cada píxel individual.

También introduciremos el modelo de ruido multiplicativo en el cual datos SAR de un único canal son tratados como el producto de un RCS subyacente y un proceso de ruido estacionario, correspondiente al speckle.

El coeficiente de *backscatter* o de retrodispersión, llamado *Radar Cross Section* (RCS), está definido como la razón entre la energía que hubiera emitido el sensor si la misma hubiera sido totalmente dispersada en forma isotrópica y la energía recibida por el sensor. Llamaremos a la primera  $E_1$  y a la segunda  $E_2$ , entonces

$$\sigma = \frac{E_1}{E_2}.$$

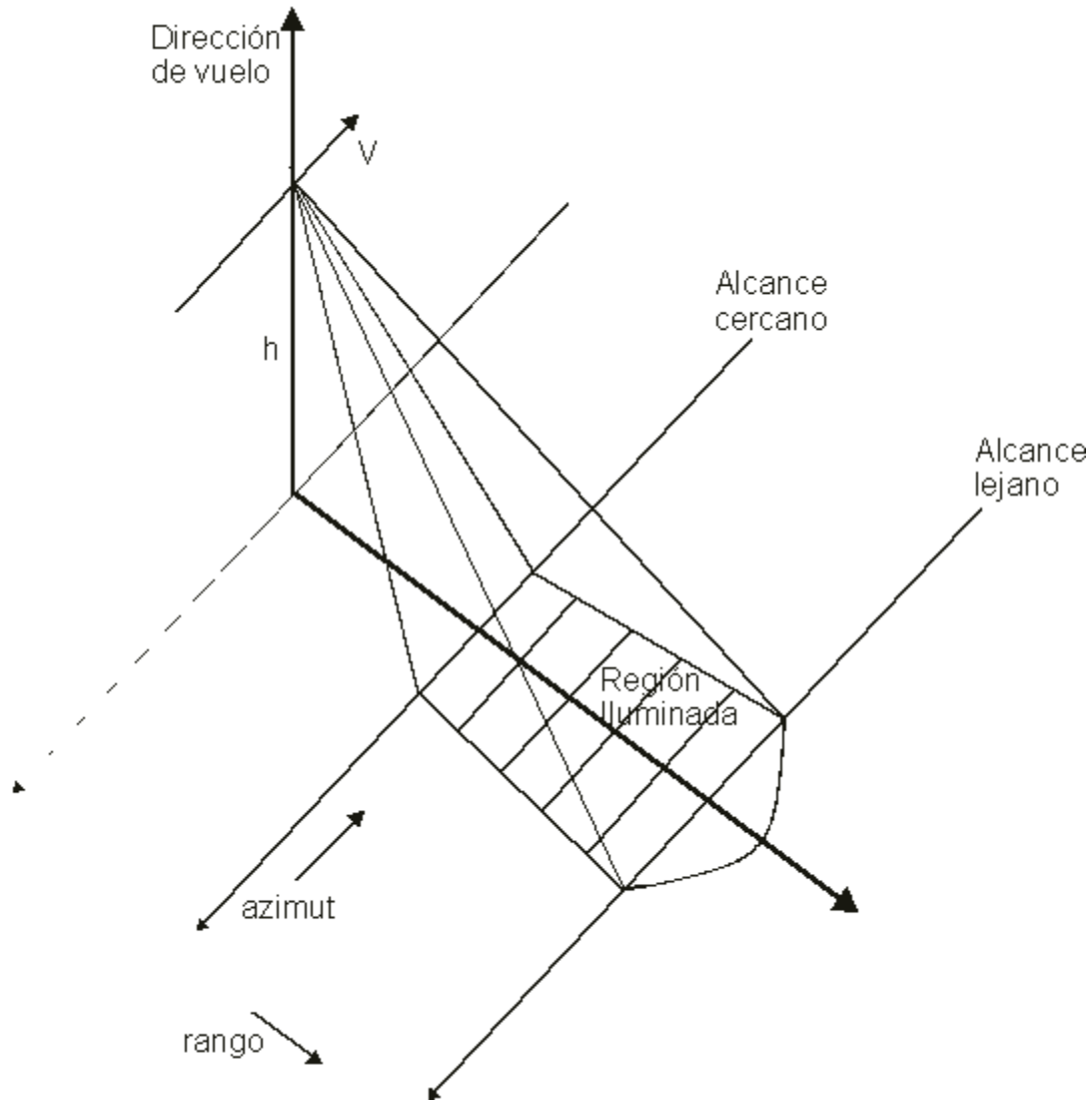
Usualmente se expresa el coeficiente anterior en decibeles, por lo que finalmente queda

$$\sigma^0 = 10 \log \frac{E_1}{E_2}.$$

## 8.1 Geometría de imágenes de radar

La teoría básica de la geometría SAR consiste en una plataforma que se mueve con velocidad  $V$  y altitud  $h$  que transporta una antena de radar que ilumina la superficie de la Tierra con pulsos de radiación electromagnética. La dirección de

la trayectoria de la plataforma (dirección de vuelo) se conoce como la dirección *azimut* ; mientras que la dirección perpendicular a esta última se denomina dirección del *rango*. Esta geometría se muestra en la figura



A continuación definimos sus elementos más importantes.

### 1 Visión lateral.

Una plataforma desplazándose a una velocidad  $V$  y altura  $h$  lleva una antena de radar de observación lateral que ilumina la superficie de la tierra con pulsos de radiación electromagnética.

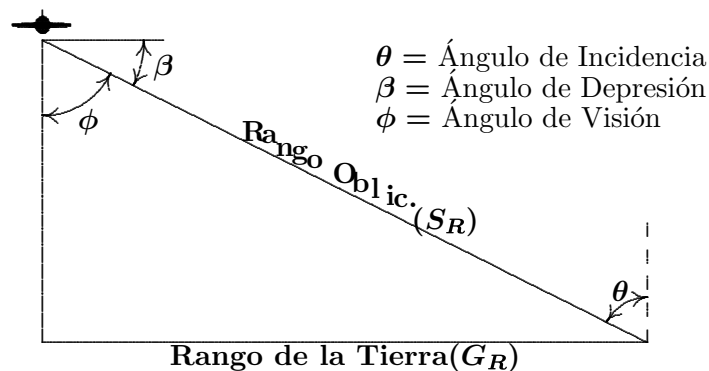
La zona rayada es la iluminada por la antena y la zona barrida por el radar al desplazarse es la franja cuyo ancho va desde el alcance cercano

hasta el alcance lejano. El radar emite una sucesión de pulsos de radiación electromagnética y al captar la energía reflejada por la tierra obtiene una sucesión de ecos (retornos) con los que se generará la imagen.

Los retornos provenientes de los distintos puntos en el terreno en el caso de visión vertical arribarán al mismo tiempo, en cambio, en la visión lateral lo harán a distintos tiempos. Este hecho permite distinguirlos entre sí y calcular su distancia al radar en la dirección de vuelo o rango. Es por ello que es necesaria la visión lateral.

## 2 Resolución en la dirección del Rango.

La dirección del rango es la perpendicular a la dirección de vuelo. El rango de un punto sobre la tierra puede medirse como la distancia del mismo a la trayectoria de vuelo (Rango Oblicuo: *Slant Range*) o como la proyección de esta distancia sobre la tierra (Rango de Tierra: *Ground Range*). La resolución de una imagen de radar en la dirección del rango se determina por la longitud del pulso, que es la longitud física de la señal de microonda. La longitud del pulso es la longitud de la luz multiplicada por la duración de la transmisión. Longitudes de pulso más cortas dan una mejor resolución, aunque menos iluminación sobre la superficie, y por lo tanto una señal de retorno más débil. Luego para asegurarse que una señal de retorno sea suficientemente fuerte, a fin de dar una imagen definida, los sistemas de radar son limitados en la resolución del rango que pueden alcanzar.



Si el radar emite pulsos de duración  $\tau$ , la resolución del mismo  $\Delta\tau_{min}$  queda determinada por la fórmula

$$\Delta\tau_{min} = \frac{c\tau}{2},$$

donde  $c$  es la velocidad de la luz. Ésta corresponde a la resolución en el rango oblicuo.

La resolución en el rango de la tierra es

$$\frac{\Delta\tau_{min}}{\sin\theta} = \frac{c\tau}{2\sin\theta},$$

donde  $\theta$  es el ángulo de incidencia lejano.

### 3 Resolución en la dirección Azimutal.

Radar de Apertura Sintética vs. Radar de Apertura Real.

El Radar de Apertura Sintética posee la ventaja de que la resolución azimutal es independiente de la distancia, y puede ser tan pequeña como:

$$\frac{L}{2},$$

donde  $L$  es la longitud de la antena.

Este hecho resulta paradójico a primera vista, ya que mejoramos la resolución reduciendo el largo de la antena y, por lo tanto, ensanchando el haz de radiación de la misma.

Esto ocurre porque, en el procesamiento que se realiza a los pulsos captados por el radar, la posición de un punto sobre la tierra se determina en base a la historia de fase del mismo.

La historia de fase permite una mejor ubicación de la posición azimutal de un punto cuanto mayor sea su duración, y esta duración es directamente proporcional al ancho del haz.

Durante el procesamiento de los pulsos captados por el radar, se sintetiza una antena de muy alta resolución. Además, durante este procesamiento, pueden sintetizarse varias pasadas sobre la misma región con un solo juego de datos.

Las imágenes generadas por dichas pasadas corresponden a antenas sintéticas que observan la región desde ángulos ligeramente diferentes entre sí. A cada una de estas imágenes se las denomina un *look*. Al promediarlas entre sí se genera una sola imagen *multilook* la cual es de más fácil interpretación visual gracias a la reducción del fenómeno de speckle respecto de una imagen ordinaria (de un solo look).

En el SAR la resolución azimutal está limitada por

$$\frac{h\lambda}{\cos\theta},$$

donde  $h$  es la altura de vuelo,  $\lambda$  la longitud de onda y  $\theta$  el ángulo de incidencia.

Vemos que para tener mejor resolución a una longitud de onda  $\lambda$  fija, debemos reducir  $h$  y aumentar  $L$  (es decir, vuelo bajo y antena grande).

Varios ángulos son importantes de definir en la geometría de un sistema de radar como son:

- ángulo de incidencia ( $\theta$ ):

Es el ángulo entre la dirección de propagación de onda y la normal a la superficie; el ángulo con el que incide la radiación sobre la tierra. Crece con la distancia en la dirección del rango, decrece a medida que la pendiente del terreno aumenta. La señal de retorno disminuye cuando aumenta el ángulo de incidencia local.

Dependiendo de la altura del radar sobre la superficie de la tierra, el ángulo de incidencia variará desde el alcance cercano al alcance lejano, lo que a su vez influye en la geometría de la observación. Debido a esta variación, la iluminación del terreno resulta más oblicua. Por lo tanto, el sombreado se hace más pronunciado hacia el alcance lejano. Puede obtenerse información sobre la altura del objeto sensado a partir de las sombras.

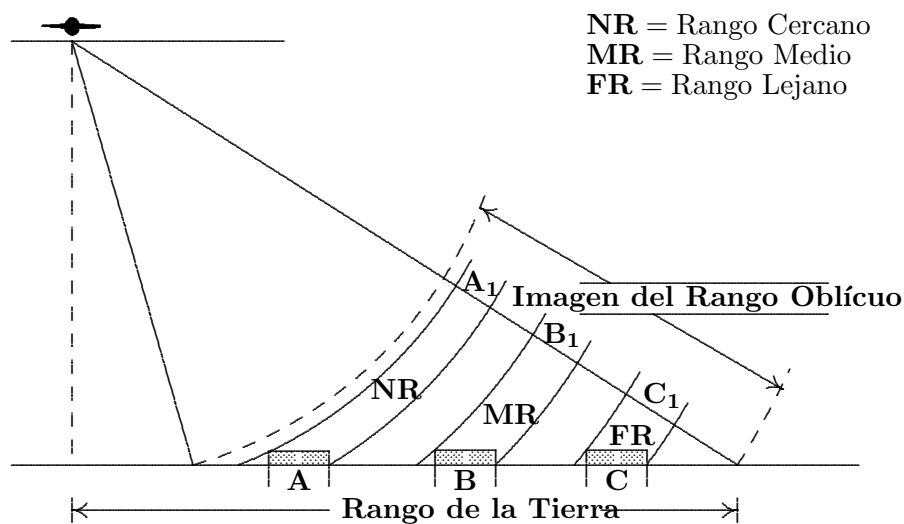
- ángulo de depresión ( $\beta$ ):

Es el ángulo entre la horizontal y la dirección de propagación de onda, luego es el complemento del ángulo de visión.

- ángulo de Azimut y Squint:

El ángulo Azimut es el ángulo entre la dirección azimut y la dirección rango, usualmente es un ángulo de  $90^{\circ}$ . El ángulo Squint es el complemento del ángulo Azimut, luego es usualmente de  $0^{\circ}$ .

El sistema de radar registra la ubicación de objetos usando la distancia del sensor al objeto a lo largo de la línea de visión, más que a lo largo de la superficie. Esta geometría es referida como geometría de visión lateral.





Todas las distorsiones siguientes se deben al hecho que el radar colecciona datos originalmente a lo largo del rango de inclinación, más que del rango del suelo:

- **Foreshortening:** es una distorsión topográfica, en la cual la pendiente del terreno que mira hacia el sensor del radar aparecerá mucho más corta que la pendiente que mira lejos del sensor, aún cuando ambas pendientes sean de igual longitud. El efecto es más pronunciado para pendientes más empinadas, y mayores ángulos de depresión.
- **Layover:** es una distorsión topográfica, en la cual la parte más alta de un objeto topográfico aparece desplazada horizontalmente hacia el rango cercano en relación a su base. El layover hace que elementos montañosos en imágenes de radar aparezcan plegados hacia el sensor. El efecto es más pronunciado para pendientes más empinadas, y mayores ángulos de depresión.
- **Sombra del Radar:** la retroinclinación (lejos de la pendiente del rango) de un elemento topográfico es oscurecida por el rayo del radar, resultando en un no retorno de tal área. La longitud de una sombra del radar puede usarse para determinar el ángulo de pendiente del terreno y un realce relativo del elemento.

## 8.2 Parámetros del sistema SAR

La frecuencia de la onda electromagnética emitida por la antena del radar se denomina frecuencia de emisión. Otros parámetros en el sistema SAR son:

**Polarización.** La polarización emitida por el radar es polarizada, esto significa que el vector campo eléctrico tiene una dirección determinada, que puede ser vertical, horizontal (respecto del plano de la tierra) o tener ambas componentes. Al ser reflejada, la radiación experimenta, en general, una rotación del vector campo eléctrico. Entonces se debe decidir si se capta la componente horizontal del mismo, la vertical o ambas. Por lo tanto, existen cuatro posibilidades:  $HH$ ,  $HV$ ,  $VH$ ,  $VV$ , donde la primera letra se refiere a la señal emitida y la segunda a la recibida.

**Velocidad y pulsos/segundo.** De la velocidad de desplazamiento del radar y de la cantidad de pulsos por segundo que éste emita, depende el espaciamiento entre las líneas de la imagen generada.

**Swath.** Es el ancho de la zona barrida por el radar y corresponde al ancho de las imágenes.

### 8.3 La naturaleza de la información en imágenes SAR

La cantidad básica medida por un SAR de una única frecuencia y única polarización en cada píxel es un par de voltajes en los canales fase y cuadratura. Los valores medidos representan los efectos de la escena sobre la onda transmitida, pero no en forma directa. En primer lugar, el ancho de banda finito del sistema tanto en rango como en azimuth implica que los valores medidos sean promedios ponderados de una región determinada por la función de dispersión puntual del SAR. En segundo lugar, se necesitan aplicar numerosas funciones de peso en el procesamiento para convertir los voltajes medidos a unidades “geofísicas” correspondientes a la reflectividad compleja, RCS o coeficiente de retrodispersión de la escena. Con esas condiciones, las mediciones hechas por el SAR son fundamentalmente determinadas por procesos de dispersión electromagnética.

Desde el punto de vista de la teoría electromagnética, son importantes dos enfoques. El primero de ellos es el *Problema Directo*, en el cual las propiedades del medio dispersante y la onda entrante se especifican y se usan para predecir el campo donde se dispersó la señal. Soluciones a este problema se obtienen frecuentemente sólo imponiendo restricciones considerables sobre el scattering medio, por ejemplo, suponiendo que todas las longitudes de escala en el medio son grandes (o pequeñas) relativas a la longitud de onda, o que solamente la superficie (no volumen) es observada (o viceversa), o que los efectos de sombra se pueden ignorar.

Otra debilidad de la mayoría de los enfoques de modelación incluyen:

- La suposición de ondas planas infinitas y media infinita, en lugar de la imagen donde tenemos una escena estructurada.
- Su habilidad para producir solamente momentos de orden bajo del campo de dispersión (frecuentemente sólo el coeficiente de retrodispersión medio), mientras son importantes los momentos de orden superior como fuentes de información, especialmente los momentos de cuarto orden son necesarios para describir la varianza de la intensidad y correlación, que son críticos en el análisis de textura.

Ellos proveen sólo una parte del conocimiento necesario para discutir información en imágenes SAR y pueden frecuentemente ser irrelevantes (por ejemplo, cuando las estructuras dentro de la escena son muy heterogéneas). Aún dentro del contexto de la teoría electromagnética, el Problema Directo no es lo que realmente interesa cuando tratamos con información en imágenes SAR. Lo más importante es el *Problema Inverso*: ¿qué propiedades de una escena desconocida pueden inferirse a partir de la señal retrodispersada observada?

Los métodos para resolver el Problema Inverso en forma matemática son naturalmente presentados en términos de los mismos modelos que para el Problema Directo, con las mismas limitaciones, y son normalmente dificultosos.

No obstante, el origen real de la dificultad en el Problema Inverso no es la matemática sino que existe escasa información para proveer una única solución. En otras palabras el número de parámetros necesarios para caracterizar la escena observada en el modelo excede el número de mediciones independientes obtenibles por el sensor. En este caso, la dimensionalidad del espacio paramétrico puede reducirse asumiendo que algunos parámetros se conocen o que la solución no es sensible a ellos (esos parámetros entonces no son medibles a partir de las observaciones, pero el acceso a los otros parámetros es posible). Aún cuando una inversión apropiada sea posible, las mediciones frecuentemente fallan en confirmar la teoría, reflejando de alguna forma la falla de la teoría para captar la relevancia física del proceso de dispersión.

Esos enfoques para describir información en imágenes SAR a través de la teoría electromagnética tienen la ventaja aparente de que ellos aplican la señal observada en propiedades físicas de la escena, tal como constante dieléctrica, factores geométricos (por ejemplo: tamaño, peso y rugosidad), y condiciones de polarización. Al hacerlo ellos proveen una relación entre las observaciones y ciertas cantidades medibles sobre el suelo. Generalmente, ambos Problemas Directo e Inverso requieren algún conocimiento a priori antes de ser llevados a cabo. El primer requerimiento es que la región en las que ellos se aplican sean homogéneas en algún sentido, por ejemplo se sabe que la región está formada por bosque.

La teoría física sin una teoría de medición adecuada asociada no es del todo física: capturar la percepción visual en mediciones, si es posible, abre el camino al entendimiento cuantitativo de la imagen. Los modelos para los datos son en la mayoría de los casos dependientes de un pequeño número de parámetros que proveen ventajas, siempre y cuando ellos sean consistentes con los datos:

- Toda la información obtenida es representada por algunos pocos parámetros.
- Es posible desarrollar procedimientos de estimación sólidos y cuantificar la adecuación con la que los parámetros pueden obtenerse.
- Existe una base rigurosa sobre la cual desarrollar técnicas de análisis de imágenes, tal como filtrado, clasificación y detección de blancos.

Nuestro enfoque para modelar datos es sistemático y progresivo. Comenzando por las propiedades de la imagen derivamos descripciones físicas y/o fenomenológicas de los datos para proveer nuestro modelo. Examinamos los tipos de características de la imagen sobre las cuales el modelo es aplicable; éstas varían generalmente con las características del sensor, siendo la resolución particularmente importante. En combinación con un “modelo real” van incorporando tipos de conocimiento general acerca de la escena (por ejemplo, un área agrícola probablemente contiene campos de un único cereal). Se usa el modelo para los datos para formular técnicas de interpretación de la imagen que son óptimas en algún sentido. Estos proveen una imagen modificada en la cual

la información es más nítida que en la imagen original. Cuando las técnicas de interpretación son locales, nos interesa la reconstrucción de alguna propiedad subyacente de la imagen; esto es esencialmente un filtrado de la misma. Cuando el modelo impone una descripción a gran escala, estaremos interesados en la estructura de la imagen. Se elige la escala de acuerdo al contexto de una aplicación: ¿qué propiedades de la realidad resultan detectables o medibles como resultado de la técnica de interpretación?

Aunque hemos descrito las mediciones como un tema central, éstas necesitan alguna interpretación. Para detección de blancos en datos de un sólo canal, se necesita frecuentemente la calibración de la imagen debido a que no estamos interesados en mediciones precisas del RCS o coeficiente de retrodispersión, solamente el contraste entre el objetivo y su ambiente. Consideraciones similares se aplican si nuestro interés está en cartografía estructural. Para detecciones de blancos, especialmente si se combinan canales múltiples, la calibración se vuelve de mayor interés, aunque la calibración relativa puede ser más importante que la absoluta.

La noción de información y su relación a mediciones también necesita alguna calificación. Particularmente en escalas grandes, encontramos dificultades de consistencia entre un modelo explotable para los datos y el conocimiento humano del mundo real.

La distinción entre información descrita por mediciones y la derivada del conocimiento y las aplicaciones necesitan entonces ser más exactas.

## 9 Tipos de Imagen de un único canal

### Speckle

Cuando observamos una imagen, inmediatamente reconocemos formas, objetos, líneas, y regiones, algunas de las cuales se distinguen por su brillo, otras por sus cualidades de textura. También, sobre todo en regiones grandes, con diversos campos, generalmente observamos que los píxeles muestran una variabilidad considerable. Por lo tanto, es pertinente preguntarnos cuál es el origen de esta variabilidad y cuando un píxel individual en un blanco extenso puede decirse que transporta información.

Antes de responder a estas cuestiones necesitamos reconocer que una imagen involucra una elección de cómo se muestran las mediciones por el SAR que se ha usado. Hemos visto que la cantidad geofísica que es determinante en los datos SAR es la reflectividad compleja del radar de la escena y que el SAR provee una medición de esa reflectividad. Cualitativamente la reflectividad compleja expresa el hecho que cuando una onda electromagnética se dispersa desde la posición  $(x, y)$  sobre la superficie de la Tierra, las propiedades físicas

del terreno causan cambios tanto en la fase,  $\phi(x, y)$ , como en la amplitud,  $A(x, y)$  de la onda. El SAR, de hecho mide el par  $(A \cos \phi, A \sin \phi)$  en los canales fase y cuadratura del receptor, ponderado por la PSF (función de dispersión puntual) del SAR. Esta estimación de la reflectividad local en cada píxel puede también representarse por el número complejo  $Ae^{i\phi}$ , de esta forma los datos SAR se conocen como *imagen compleja*.

Desde la imagen compleja puede obtenerse una gran variedad de otras imágenes, por ejemplo, imágenes de la parte real  $A \cos \phi$ , de la parte imaginaria  $A \sin \phi$ , la amplitud  $A$ , la fase  $\phi$ , la intensidad  $I = A^2$ , o la log intensidad  $\log I$ .

Generalmente las imágenes de la parte real e imaginaria muestran alguna estructura que puede aparecer extremadamente ruidosa o difusa, la imagen de fase es similar y no muestra estructura, mientras que las imágenes de amplitud, intensidad y log intensidad, aunque con ruido, son claramente más fáciles de interpretar.

Las características de calidad difusa de estos tipos de imagen se conoce como *speckle*.

Cuando la onda interactúa con el blanco, cada fuente de dispersión contribuye a la onda retrodispersada con un cambio en fase y amplitud, de modo que la modulación total retornada de la onda incidente es

$$Ae^{i\phi} = \sum_{k=1}^N A_k e^{i\phi_k}. \quad (44)$$

Esta suma es sobre el número total de fuentes de dispersión “iluminados” por el rayo incidente; incluye los pesos impuestos por la PSF del SAR y cualquier efecto de atenuación causado por el proceso de dispersión y propagación. Las amplitudes  $A_k$  y las fases  $\phi_k$  no son observables debido a que los fuentes de dispersión individuales son de escalas mucho más pequeñas que la resolución del SAR, y existen normalmente varias de tales fuentes por celda de resolución.

Una conclusión inmediata de (44) es que la señal observada será afectada por los efectos de interferencia como una consecuencia de las diferencias de fases entre fuentes de dispersión. En realidad, el speckle puede entenderse como un fenómeno de interferencia en el cual el origen principal de la calidad ruidosa de los datos observados es la distribución de los términos  $\phi_k$ . Como un resultado, podemos pensar en la práctica que las fases  $\phi_k$  se distribuyen uniformemente en  $[-\pi, \pi]$  e independiente de la amplitud  $A_k$ . La suma (44) entonces se mira como un camino aleatorio en el plano complejo, donde en cada paso la longitud  $A_k$  está en una dirección completamente aleatoria. Para un gran número de fuentes de dispersión estadísticamente idénticos, el análisis revela que:

1. Las componentes observadas en fase y cuadratura,  $z_1 = A \cos \phi$  y  $z_2 = A \sin \phi$ , serán variables aleatorias Gaussianas independientes e idénticamente distribuidas con media cero y cuya varianza  $\sigma/2$  se determina por

las amplitudes  $A_k$ , esto es, tienen función densidad de probabilidad conjunta (FDP) dada por

$$P_{z_1, z_2} = \frac{1}{\pi\sigma} \exp - \left( \frac{z_1^2 + z_2^2}{\sigma} \right). \quad (45)$$

2. La fase observada  $\phi$  será distribuída uniformemente sobre  $[-\pi, \pi]$ .
3. La amplitud  $A$  tendrá una distribución Rayleigh

$$P_A(A) = \frac{2A}{\sigma} \exp \left( -\frac{A^2}{\sigma} \right) \quad A \geq 0, \quad (46)$$

con media  $\sqrt{\pi\sigma}/2$  y desviación estándar  $\sqrt{(1 - \pi/4)\sigma}$ .

4. La intensidad observada es  $I = A^2$  tendrá una distribución exponencial negativa

$$P_I(I) = \frac{1}{\sigma} \exp \left( -\frac{I}{\sigma} \right) \quad I \geq 0, \quad (47)$$

con media y desviación estándar  $\sigma$ .

5. La log intensidad  $D = \ln I$  tiene una distribución Fischer-Tippett

$$P_D(D) = \frac{e^D}{\sigma} \exp \left( -\frac{e^D}{\sigma} \right), \quad (48)$$

cuyo valor medio y varianza son  $\ln \sigma - \gamma_E$  y  $\pi^2/6$  respectivamente. El símbolo  $\gamma_E$  denota la constante de Euler con valor aproximado 0.57722.

Las distribuciones 1 a 5 (para imágenes muy homogéneas) son de importancia fundamental en la manipulación de datos SAR. Notemos que, con excepción de la distribución para la fase, ellas se caracterizan por un único parámetro  $\sigma$ , que transporta toda la información acerca del blanco. A partir de (47) vemos que  $\sigma$  corresponde a la intensidad promedio.

Observamos además que las diferentes representaciones de los datos llevan diferente cantidad de información acerca de blancos extensos en la escena. La fase no lleva información debido a que su distribución es independiente del blanco. Los canales real e imaginario proveen información pero requieren el ojo para estimar la desviación estándar local para detectar diferencias entre regiones. Los datos de amplitud llevan la información en su valor medio, aunque difusa (con ruido) debido a la forma de la distribución de Rayleigh. Los datos de intensidad tienen un rango dinámico más amplio que los datos de amplitud debido a que el coeficiente de variación es mayor para la distribución exponencial que para la de Rayleigh, para los datos de amplitud es 0.52, mientras que para los datos de intensidad es 1.

La observación que muchos blancos extensos obedecen al modelo simple del speckle, permite responder parcialmente a nuestra pregunta acerca de si un píxel de la imagen transporta información. Considerado aisladamente, cada píxel provee una medición real de la potencia o amplitud retrodispersada. Cuando un píxel se forma por varias fuentes de dispersión elementales, la potencia observada es una estimación de un RCS subyacente cuyo valor se oculta por efectos de interferencia. Si el píxel es parte de blancos uniformemente repartidos, la información por píxel es baja puesto que la colección de píxeles en el blanco puede proveer solamente un número útil,  $\sigma$ . Mayor número de píxeles reducen el error de estimación de  $\sigma$  cuando promediamos píxeles.

El valor de la intensidad en cada píxel de un blanco repartido puede mirarse como una estimación escalada de  $\sigma^0$  (coeficiente de retrodispersión promedio para blancos extensos) o, si la corrección de área no se ha hecho, el RCS medio de cada píxel. Con esta observación frecuentemente usaremos el término RCS o “sección transversal” para referirnos a la intensidad media, aún cuando los datos no han sido propiamente calibrados.

Un punto a destacar es la necesidad de distinguir el valor medido en un píxel y el valor del parámetro  $\sigma$ . De acuerdo a (44) el valor observado en cada píxel es el resultante de ondas Huygens interferentes, salvo que un único scatterer domine completamente el retorno. Luego el valor de  $\sigma$  es específico para cada píxel; el valor medido es justamente una muestra de la distribución parametrizada por  $\sigma$ . Para hacer esto más claro, notemos que la intensidad media en un píxel se da, de acuerdo a (44), por

$$E[I] = E[A^2] = \sum_{j=1}^N \sum_{k=1}^N E[A_j A_k e^{i(\phi_j - \phi_k)}].$$

Suponiendo que la amplitud y la fase en cada scatterer son independientes de los otros y que las fases de scatterers diferentes son independientes y distribuidos uniformemente tenemos

$$E[I] = \sum_{k=1}^N E[A_k^2]. \quad (49)$$

Luego, la intensidad media se determina por la suma incoherente de la intensidad media de todos los scatterers y es independiente de su configuración geométrica (suponiendo que las orientaciones no se cambian). No obstante, cambios en la posición de las fuentes de dispersión cambiarán todas las relaciones de fase en el proceso, alterando la intensidad observada, pero sin alterar el parámetro  $\sigma$ . En este sentido, la intensidad *observada* en un píxel es determinística y repetible pero para el propósito de medición es mejor considerarla como aleatoria y como una estimación de  $\sigma$ .

## 10 Estimación del RCS en datos multilook

Puesto que  $\sigma$  es el parámetro de interés, es pertinente preguntarnos cómo mejoramos su estimación para un píxel dado. Usando (45), se demuestra que la *estimación de máxima verosimilitud* (EMV) de  $\sigma$ , dadas las componentes en fase y cuadratura  $(z_1, z_2)$ , se dá por

$$\hat{\sigma} = z_1^2 + z_2^2 = I, \quad (50)$$

esto es, la intensidad observada. Puesto que las distribuciones (46) a (48) son todas transformaciones del par de números  $(z_1, z_2)$ , todas llevan a la misma estimación de  $\sigma$ . Luego, si no se tiene ninguna suposición o conocimiento anterior, la mejor estimación de  $\sigma$  en un píxel dado es formar la imagen intensidad.

Se puede mejorar la estimación de  $\sigma$ , dadas  $L$  mediciones independientes, se promedien en *intensidad* (frecuentemente este promedio se denomina incoherente debido a que se descarta toda la información de fase). Es fácil probar que se obtiene de esta forma un estimador de máxima verosimilitud de  $\sigma$  dadas  $L$  mediciones independientes e idénticamente distribuídas en un píxel combinando mediciones en esa posición. Puesto que  $\sigma$  es la potencia media, ésto sugiere forma exponencial con media  $\sigma$ . Además, de acuerdo a (47), se preserva el valor medio  $\sigma$  mientras reduce la varianza de las mediciones a  $\sigma/L$ .

Se pueden obtener mediciones independientes separando la apertura sintética del ancho de banda del Doppler. Las imágenes separadas se denominan looks, de modo que este proceso de promediar la intensidad se conoce como multilooking, y la imagen resultante se conoce como  $L - look$ . La premisa fundamental del multilooking es que cada look deberá estimar el mismo RCS, por lo tanto es importante que los looks deban estar en la misma posición, casi simultánea, y sin distorsión radiométrica relativa. También necesitan estar en la misma frecuencia y configuración de polarización. No obstante, pueden ocurrir diferencias entre looks debido a ángulos de visión algo diferentes usados en cada look.

Un segundo enfoque para mejorar la estimación de  $\sigma$  en una posición dada depende de hacer suposiciones acerca de las propiedades espaciales de  $\sigma$ . Quizás la más simple de esas suposiciones es que el RCS es constante sobre algún entorno alrededor del píxel de interés (escena homogénea). En este caso, dados  $L$  valores de píxeles independientes, la situación es estadísticamente idéntica a la descrita previamente, donde looks independientes en el mismo punto se forman en el procesador. Luego, nuevamente las mediciones deberán combinarse promediando en intensidad. Ambos casos son conocidos como multilooking, aunque ellos contienen diferentes tipos de información a fin de mejorar la precisión radiométrica. En el primer caso, la variación angular en el RCS es ignorada, mientras que en el segundo, se pierde la variación espacial. Ambos proveen estimaciones adecuadas de  $\sigma$  sobre un blanco uniforme, pero en



regiones heterogéneas los dos métodos pueden proveer estimaciones muy diferentes del RCS medio en una posición dada, debido a los diferentes anchos de la PSF.

Notemos que en ambos casos la resolución espacial se habrá degradado. En el primer caso esto se debe a que cada look usa solamente una fracción  $1/L$  del ancho de banda disponible, luego la resolución se degrada por el factor  $L$ . El segundo caso es más complicado, puesto que el proceso de promediar actúa como un filtro de pasa-bajo; el valor asignado a la resolución después del filtrado depende de la definición precisa de la resolución que se ha usado. Ambos casos resultan ser complicados cuando los looks no son independientes, tal como ocurre si los looks se forman con anchos de banda del Doppler que presenten solapes en el dominio de frecuencias o si los píxeles son correlacionados en el dominio espacial.

La intensidad promedio  $L$ -look resultante es

$$I = \frac{1}{L} \sum_{k=1}^L I_k, \quad (51)$$

donde los  $I_k$  son variables independientes e idénticamente distribuidas con media  $\sigma$ , por lo tanto  $I$  (que en realidad deberíamos denotar con  $\bar{I}$ ) tiene una distribución gamma con parámetros  $L$  y  $1/\sigma$

$$P_I(I) = \frac{1}{\Gamma(L)} \left(\frac{L}{\sigma}\right)^L I^{L-1} e^{-LI/\sigma}. \quad (52)$$

La intensidad promedio tiene momentos

$$E[I^m] = \frac{\Gamma(m+L)}{\Gamma(L)} \left(\frac{\sigma}{L}\right)^m \quad I \geq 0. \quad (53)$$

en particular  $E[I] = \sigma$  y  $var(I) = \sigma^2/L$ . La relación anterior motiva la definición del *número equivalente de looks* (NEL) como

$$NEL = \frac{(media)^2}{varianza}, \quad (54)$$

donde los promedios son llevados a cabo en intensidad sobre blancos uniformemente distribuidas. El NEL es equivalente al número de valores de intensidad independientes promediados por píxel.

## 10.1 El Modelo de Ruido Multiplicativo para el RCS

El modelo Scatterer discreto para blancos extensos implica que toda la información en cada píxel en datos de un único canal, es transportada por la potencia

media incoherente,  $\sigma$ , apropiada para su conjunto particular de scatterers. La intensidad observada en cada píxel entonces tiene la probabilidad condicional

$$P_I(I) = \frac{1}{\sigma} e^{-I/\sigma} \quad I \geq 0. \quad (55)$$

Haciendo el cambio de variable

$$I = \sigma n, \quad (56)$$

resulta la FDP

$$P_n(n) = e^{-n} \quad n \geq 0. \quad (57)$$

Luego, la intensidad observada en cada punto puede mirarse como un valor determinístico *RCS* multiplicado por la unidad media speckle distribuído exponencialmente. Equivalentemente, la reflectividad compleja en cada píxel puede escribirse como

$$S = S_1 + iS_2 = \sqrt{\sigma}(m_1 + im_2), \quad (58)$$

donde la contribución al speckle (voltaje)  $m_1 + im_2$  tiene *FDP*

$$P_{m_1, m_2}(m_1, m_2) = \frac{1}{\pi} \exp(-m_1^2 - m_2^2), \quad (59)$$

esto es, las componentes del speckle en fase y cuadratura son variables aleatorias independientes e idénticamente distribuídas en forma normal con media cero y varianza  $1/2$ .

Análogamente, para datos *L*-look, la descomposición (56) es posible, donde  $n$  es ahora una unidad media con distribución gamma con parámetro de orden  $L$ .

Esta formulación de las imágenes SAR como un *RCS* determinístico modulando a un proceso speckle estacionario permite una descripción del speckle como un ruido multiplicativo, aunque es un fenómeno de interferencia repetible.

Aunque el *RCS* subyacente a los valores observados es esencialmente determinístico, su magnitud es desconocida y normalmente es más útil describirla por un modelo probabilístico. Dado un conjunto de  $N$  píxeles, necesitamos considerar la distribución conjunta del *RCS*  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$  y los valores del speckle  $= (n_1, n_2, \dots, n_N)$ . Si suponemos que los valores del speckle son independientes entre sí y del *RCS*, podemos escribir

$$P_\sigma(\sigma) = P_\sigma(\sigma) \prod_{k=1}^N P_n(n_k). \quad (60)$$

La correspondiente intensidad  $= (I_1, I_2, \dots, I_N)$  se dan por  $I_k = \sigma_k n_k$ . Un cambio de variables lleva a la distribución conjunta de intensidad y speckle

$$P_{\mathbf{I}}(\mathbf{I}) = P_{\sigma}(I_1/n_1, \dots, I_N/n_N) \prod_{k=1}^N \frac{P_n(n_k)}{n_k}. \quad (61)$$

Integrando con respecto al speckle obtenemos la *FDP* conjunta de la intensidad como

$$P_{\mathbf{I}}(\mathbf{I}) = \int P_{\sigma}(I_1/n_1, \dots, I_N/n_N) \prod_{k=1}^N \frac{P_n(n_k)}{n_k} d. \quad (62)$$

Luego podemos, en principio, encontrar la distribución conjunta de la intensidad correspondiente a cualquier distribución conjunta del *RCS*.

## 11 Modelos para los datos

### 11.1 Introducción

Derivar información de imágenes SAR es un proceso en el cual los datos de la imagen son interpretados a la luz del conocimiento a priori que es encapsulado en modelos para los datos y las propiedades de la escena. Mostraremos cómo diferentes modelos permiten extraer información óptima para la interpretación de una imagen.

### 11.2 El modelo producto

Un progreso significativo en el entendimiento de las propiedades de imágenes de radar fue la apreciación que muchos tipos de fondos de la escena derivan de los dos procesos no relacionados que pueden encapsularse en el modelo producto. En su forma más simple, éste combina un *RCS* subyacente  $\sigma$ , con un speckle multiplicativo no correlacionado  $n$ ; de modo que la intensidad  $I$ , en una imagen SAR puede expresarse como el producto  $I = \sigma n$ . Para datos de intensidad *L*-look,  $I$  puede ser representada como el producto de  $\sigma$  con una contribución speckle  $n$ , cuya *FDP* se dá por

$$P(n) = \frac{L^L n^{L-1}}{\Gamma(L)} \exp(-Ln). \quad (63)$$

Este modelo forma la base de filtros de reconstrucción y métodos de segmentación que se diseñan para valores específicos de  $\sigma$  asociados con un valor de intensidad observada  $I$ . Notemos que para explotar el modelo, es esencial

que el speckle y las fluctuaciones del *RCS* tengan escalas espaciales muy diferentes. Normalmente remuestreamos la imagen SAR de modo que el speckle contribuya con poca correlación entre las muestras. Cualquier fluctuación del *RCS* sobre una escala espacial mayor puede entonces separarse para nuevos estudios.

La próxima componente del modelo producto es incorporar un *modelo RCS* que describa las fluctuaciones subyacentes. Para una distribución del *RCS* dada  $P(\sigma)$ , el modelo producto asegura que la FDP de la intensidad observada se dá por

$$P(I) = \int_0^{\infty} P(I/\sigma)P(\sigma)d\sigma = \frac{L^L I^{L-1}}{\Gamma(L)} \int_0^{\infty} \frac{d\sigma}{\sigma^L} \exp\left[-\frac{LI}{\sigma}\right]P(\sigma). \quad (64)$$

El speckle es puramente una consecuencia de la distribución aleatoria de los elementos scattering, mientras el *RCS* depende de las propiedades físicas (tal como constante dieléctrica) de aquellos elementos.

### 11.3 El modelo *RCS*

Habiendo definido la contribución del speckle al modelo producto, la etapa siguiente es discutir el modelo *RCS* para el caso de una imagen heterogénea. Algunos estudios del fondo de la escena del mar hechos por Ward, mostraron que la amplitud detectada tenía una distribución Rayleigh en regiones de oleaje despreciable, correspondiente al speckle puro. Los oleajes tienen una escala espacial mucho mayor que el speckle, permitiendo que la contribución de éste sea promediada sin degradar apreciablemente la componente del oleaje. El *RCS* del oleaje era usualmente consistente con una FDP gamma, dada por

$$P(\sigma) = \left[\frac{\nu}{\mu_\sigma}\right]^\nu \frac{\sigma^{\nu-1}}{\Gamma(\nu)} \exp\left[-\frac{\nu\sigma}{\mu_\sigma}\right], \quad (65)$$

donde  $\nu$  es un parámetro de orden, con momentos dados por

$$E[\sigma^n] = \left(\frac{\mu_\sigma}{\nu}\right)^n \frac{\Gamma(n + \nu)}{\Gamma(\nu)}. \quad (66)$$

Esta FDP se caracteriza completamente por el *RCS* medio  $\mu_\sigma$  y el parámetro de orden  $\nu$ . El cuadrado del contraste, esto es, la varianza normalizada, es una cantidad útil para describir tales datos y se dá por  $var[\sigma]/\mu_\sigma^2 = 1/\nu$ .

### 11.3.1 La FDP de la Intensidad

Si ahora combinamos la FDP del speckle con una FDP gamma para el *RCS*, la FDP de la intensidad observada será

$$P(I) = \int_0^{\infty} P(I/\sigma)P(\sigma)d\sigma \quad (67)$$

$$= \frac{2}{\Gamma(L)\Gamma(\nu)} \left(\frac{L\nu}{\mu_I}\right)^{(L+\nu)/2} I^{(L+\nu-2)/2} K_{\nu-L} \left[ 2 \left(\frac{\nu LI}{\mu_I}\right)^{1/2} \right], \quad (68)$$

para SAR L-look, donde  $K_{\nu-L}[\cdot]$  es la función de Bessel modificada de orden  $\nu - L$ . Notemos que ésta tiende a la distribución gamma para speckle L-look cuando  $\nu \rightarrow \infty$ . Los momentos de la intensidad se dan por

$$E[I^n] = \mu_I^n \frac{\Gamma(L+n)}{L^n \Gamma(L)} \frac{\Gamma(\nu+n)}{\nu^n \Gamma(\nu)}. \quad (69)$$

La FDP de la amplitud  $A = \sqrt{I}$  es también *K*-distribuida

$$P(A) = \int_0^{\infty} P(A/\sigma)P(\sigma)d\sigma \quad (70)$$

$$= \frac{4}{\Gamma(L)\Gamma(\nu)} \left(\frac{L\nu}{\mu_I}\right)^{(L+\nu)/2} A^{L+\nu-2} K_{\nu-L} \left[ 2A \left(\frac{\nu L}{\mu_I}\right)^{1/2} \right], \quad (71)$$

con momentos

$$E[A^n] = \mu_I^{n/2} \frac{\Gamma(L+n/2)}{L^{n/2} \Gamma(L)} \frac{\Gamma(\nu+n/2)}{\nu^{n/2} \Gamma(\nu)}. \quad (72)$$

Estas derivaciones tanto para la intensidad como para la amplitud suponen que las fluctuaciones del *RCS* son sobre una escala espacial mucho mayor que el speckle de modo que los promedios multilooking del speckle no afecten las fluctuaciones del *RCS*. Si las contribuciones no fueran separables en esta forma el resultado no sería *K*-distribuida.

## 12 Clasificación de RCS y segmentación

Anteriormente reconstruimos el RCS en cada píxel. Ahora asignaremos a cada píxel una clase particular, tomando un conjunto restringido de valores representativos de RCS. Mostraremos la necesidad de promediar sobre varios píxeles antes de hacer tal asignación para lograr efectividad, la cual es sólo sensible si los bordes de los promedios de las ventanas están dentro de una región con RCS constante en la imagen. Esto nos lleva a el modelo cartoon y su aplicación en segmentación.

## 12.1 Clasificación de RCS

Más que estimar valores de RCS ahora clasificaremos cada píxel en términos de su RCS. Supongamos que se selecciona un conjunto restringido de posibles valores de RCS y cada píxel se asigna a la clase más cercana.

El speckle causa incertidumbre en la asignación. Una clasificación óptima para un modelo speckle conocido puede difícilmente definirse, basado sobre la probabilidad condicional de cada clase RCS, dada por

$$P(I/\sigma) = \left(\frac{L}{\sigma}\right)^L \frac{I^{L-1}}{\Gamma(L)} \exp\left(-\frac{LI}{\sigma}\right).$$

El logaritmo de la verosimilitud para L-look SAR es de la forma

$$\lambda_k = L \ln L - L \ln \sigma_k + (L - l) \ln I - \ln \Gamma(L) - \frac{LI}{\sigma_k}. \quad (73)$$

Puesto que L es una constante conocida para la clasificación de máxima verosimilitud necesitamos identificar la clase  $k$  que maximiza la cantidad  $-\ln \sigma_k - \frac{I}{\sigma_k}$  para cada píxel. Las distintas clases cross-section deberán estar suficientemente separadas para que el ruido no cause malas clasificaciones. Para que las tasas de error sean razonables el espaciamiento o separabilidad de las clases deberían relacionarse a la DE de la intensidad de la imagen la cual es  $\frac{\sigma_k}{\sqrt{L}}$ .

## 12.2 El Modelo Cartoon y Segmentación

La segmentación depende de la asignación que hace el Modelo Cartoon a imágenes en regiones separadas por los bordes en las cuales algún parámetro tal como el RCS es constante.

El número y posición de estos segmentos y valores medios son desconocidos y deben ser determinados a partir de los datos. En lugar de reconstruir una estimación de RCS para cada píxel, la segmentación busca vencer al ruido identificando regiones de RCS constante. En posteriores secciones mostraremos que provee una herramienta muy potente en la interpretación de imágenes por dos razones:

- el Modelo Cartoon permite que el RCS sea reconstruido sin distorsión radiométrica.
- la estructura de la escena estará estrechamente relacionada con los bordes encontrados por la segmentación que delimitan regiones con diferentes propiedades. Esto es importante para un alto nivel de interpretación.

Existen dos condiciones para que un algoritmo de segmentación sea exitoso.

1. Todo procesamiento debe ser automático, sin intervención de operador. Sólo de esta forma habrá rapidez computacional adecuada y los resultados serán reproducibles.
2. El algoritmo deberá eludir parámetros específicos a cada imagen particular. Debe permitir modificar los parámetros globales tal como razón de falsa alarma para regiones fusionadas o detección de bordes entre ellas para optimizar la performance en una aplicación dada. En otras palabras el algoritmo debe ser general, capaz de ser aplicado a cualquier conjunto de datos SAR.

Hay tres clases básicas de algoritmos de segmentación identificado por el Modelo Cartoon: detección de bordes, regiones combinadas y filtro.

### 12.3 Detección de Bordes

Si la escena está hecha de regiones uniformes una técnica de segmentación puede basarse en identificar la presencia y posición de los bordes entre las regiones. Los algoritmos de detección de bordes se pueden subdividir en técnicas de procesamiento generales u otras que explotan las propiedades de SAR y los Modelos Cartoon.

Un algoritmo de detección de bordes fue propuesto por White (RGW) el cual explota el Modelo Cartoon e incluye tamaño de filtro adaptable.

Inicialmente son detectados los bordes dentro de la imagen. Las regiones están entre estos bordes. La segunda etapa consiste en aumentar, eventualmente, el tamaño de las regiones usando un algoritmo de fusión entre regiones.

Las etapas del algoritmo son:

1. se construyen una serie de ventanas rectangulares, variando de  $3 \times 1$  a  $57 \times 13$ , alrededor de cada píxel.
2. se calcula la intensidad promedio en cada mitad  $\bar{A}$  y  $\bar{B}$ .
3. si  $|\bar{A} - \bar{B}|$  es mayor que el umbral multiplicado por el error estándar de la diferencia de medias, entonces un borde es registrado en el centro de la ventana.

Un importante resultado de este algoritmo es que el criterio de detección 3) no se basa en el error estándar dentro de la ventana, sino sobre regiones que la ventana solapa.

Inicialmente la escena entera se mira como una única región. A medida que la segmentación progresa y los segmentos comienzan a corresponder a regiones de  $RCS$  constante, la detección de bordes se determina sólo por estadísticos del ruido.

Notemos que los bordes de brillo se detectan por máscaras más pequeñas que son usadas primero, esto es, de alta resolución. Estos bordes son retenidos

y los píxeles excluidos de las sucesivas etapas para detectar bordes. Los bordes débiles se detectan subsiguientemente con ventanas más grandes. De modo que el algoritmo adapta su resolución a señales fuertes.

La etapa creciente del algoritmo procede como sigue:

1. se coloca una serie de discos sobre la imagen, de modo que no contengan ningún borde detectado.
2. discos adyacentes se combinan para formar regiones más grandes.

Inicialmente los discos son grandes, después el tamaño de los discos que se van agregando se reduce y se repite el mismo proceso. Eventualmente, cada píxel se asigna a una región particular y se construye un mapa de la región completa. El ciclo total se repite llevando a la detección de nuevos bordes. La iteración continúa hasta que el coeficiente de variación alcance algún umbral relacionado a estadísticos del ruido.

Este algoritmo opera con más éxito sobre imágenes de amplitud que sobre imágenes de intensidad.

## 12.4 Fusión de Regiones

Los algoritmos de fusión de regiones explotan el Modelo Cartoon más directamente e intentan aumentar grandes regiones uniformes fusionando o combinando pequeñas regiones vecinas. Lo crucial es tomar la decisión de fusionar o no las regiones, lo cual depende del criterio estadístico adoptado. De hecho criterios idénticos pueden ser usados para fusionar o dividir regiones. Una forma de fusionar regiones (merge using moment MUM) se basa en el test t de Student. Los pasos de este algoritmo se pueden resumir como sigue:

1. se inicia subdividiendo la imagen en pequeñas regiones arbitrarias, por ejemplo de tamaño 2 por 2 o grillas con un sólo píxel.
2. para cada par de regiones adyacentes A y B se calcula la varianza combinada definida por

$$var(A + B) = \frac{N_A var A + N_B var B}{N_A + N_B - 2} \left( \frac{1}{N_A} + \frac{1}{N_B} \right),$$

donde  $N_A$  y  $N_B$  es el número de píxeles de las regiones A y B respectivamente.

3. se calcula  $t = |\bar{A} - \bar{B}| / \sqrt{var(A + B)}$  y se aplica el test t de Student, esto evalúa la probabilidad que A y B puedan fusionarse.
4. se compila una lista en orden de prioridad para fusionar todo par de subregiones adyacentes.



5. se fusionan las regiones de mayor prioridad y se repiten las operaciones para nuevas subregiones.

El proceso es iterado hasta que la probabilidad de falsa alarma de no fusionar una región constante alcanza algún valor prefijado.

El test t es aplicable sólo cuando distinguimos entre regiones con FDP gaussiana, la cual es inapropiada para el speckle. Se puede mejorar la performance significativamente si usamos una derivación de máxima verosimilitud usando estadísticos correctos, dados en la siguiente sección.

## 12.5 Teoría sobre fusión/división

La fusión y división de regiones depende de estimar si las dos regiones pertenecen o no a una misma población. En esta sección derivaremos un test de hipótesis de máxima verosimilitud, para determinar esto, usando las distribuciones SAR vistas anteriormente. De hecho existe más de un criterio para el cual el test puede optimizarse. O la probabilidad de detección de borde en RGW o la probabilidad de fusionar en MUM es maximizada o la adecuada posición del borde es optimizada. Encontraremos estimadores de máxima verosimilitud para ambos casos y demostraremos la mejora si los introducimos en los algoritmos existentes.

Asumimos que se tiene un tamaño adaptable de la ventana. El Modelo Cartoon implica que una ventana contiene o una región uniforme o dos regiones con un borde en alguna posición desconocida dentro de la ventana (cualquier otra posibilidad ha sido eliminada por la detección de bordes en una ventana menor). La densidad conjunta de estos dos estados puede ser expresada en términos de la condicional  $P(I/\sigma)$ . Suponemos que los píxeles 1 a  $k$  están en la región  $A$  con RCS ( $\sigma_A$ ), mientras que los píxeles  $k+1$  a  $M$  están en la región  $B$  con RCS ( $\sigma_B$ ). La probabilidad conjunta para la ventana completa dada la intensidad single-look  $I_j$  en el píxel  $j$  es dado por

$$\begin{aligned} P_{A,B}(\sigma_A, \sigma_B/I_j, k) &\propto \prod_{j=1}^k P_A(I_j/\sigma_A) \prod_{j=k+1}^M P_B(I_j/\sigma_B) \\ &= \prod_{j=1}^k \frac{1}{\sigma_A} \exp\left(-\frac{I_j}{\sigma_A}\right) \prod_{j=k+1}^M \frac{1}{\sigma_B} \exp\left(-\frac{I_j}{\sigma_B}\right). \end{aligned}$$

La probabilidad que los datos sean consistentes con una región fusionada de RCS medio  $\sigma_0$  tiene la misma forma con  $\sigma_A = \sigma_B = \sigma_0$ . La log-verosimilitud para estados divididos o fusionados  $\lambda_{split}$  y  $\lambda_{merge}$  pueden encontrarse tomando logaritmos de esas probabilidades conjuntas y se define un estimador log-

diferencia como

$$\begin{aligned}
\lambda_D(k) &= \lambda_{split}(k) - \lambda_{merge} \\
&= -k \left( \ln A + \frac{\bar{I}_A}{\sigma_A} \right) - (M - k) \left( \ln B + \frac{\bar{I}_B}{\sigma_B} \right) \\
&\quad + M \left( \ln \sigma_0 + \frac{\bar{I}_0}{\sigma_0} \right),
\end{aligned} \tag{74}$$

donde las intensidades medias se estiman sobre ventanas apropiadas. Las estimaciones de máxima verosimilitud de RSC en cada región son:  $\bar{I}_A$  de  $\sigma_A$ ,  $\bar{I}_B$  de  $\sigma_B$  y  $\bar{I}_0$  de  $\sigma_0$ . Notemos que  $\bar{I}_A$  y  $\bar{I}_B$  son funciones de  $k$  luego el estimador de máxima verosimilitud de  $k$ , posición del borde, se obtiene maximizando

$$\lambda_D(k) = -k \ln \bar{I}_A - (M - k) \ln \bar{I}_B + M \ln \bar{I}_0. \tag{75}$$

Un análisis de la performance de detección de bordes basada en esta diferencia, expresada en términos de probabilidades de detección y de falsa alarma, requiere el conocimiento de su función PDF, que no se obtiene directamente. No obstante, brindaremos otra forma de detectar bordes, basada en la razón de intensidades  $r = \frac{\bar{I}_A}{\bar{I}_B}$ . Puede demostrarse que la PDF de  $r$  es

$$P_r(r/R, k, M - k) = \frac{\Gamma(M)}{r \Gamma(k) \Gamma(M - k)} \left\{ \frac{\left( \frac{rk}{R(M-k)} \right)^k}{\left( 1 + \frac{rk}{R(M-k)} \right)^M} \right\}, \tag{76}$$

donde  $R$  denota la verdadera razón de valores medios. Si un borde es “detectado” cuando  $r$  está fuera del rango  $[t_1, t_2]$ , la probabilidad de detección en la posición  $k$ , dado que el borde verdadero está en esa posición, se da por

$$P_d = 1 - \int_{t_1}^{t_2} P(r) dr.$$

La probabilidad de falsa alarma, probabilidad de detectar un borde cuando en realidad no existe,  $P_{fa}$ , corresponde a reemplazar  $R = 1$  en la ecuación (76). Los valores umbrales  $t_1$  y  $t_2$  están por debajo y arriba de 1 y pueden relacionarse al valor equivalente de  $\lambda_D$ , que puede derivarse a partir de (75). Luego

$$\lambda_D(k) = -k \ln t_1 - M \ln M + m \ln[M - k + kt_1] \quad \text{para } t_1 < 1$$

y

$$\lambda_D(k) = -k \ln t_2 - M \ln M + m \ln[M - k + k/t_2] \quad \text{para } t_2 > 1.$$

## 13 Referencias

1. Chuvieco Emilio (1998), Fundamentos de Teledetección Espacial. Ediciones Rialp SA. Madrid.
2. Jain Anil K.(1989), Fundamentals of Digital Image Processing. Prentice Hall.
3. Mejail Marta E.(1999), La Distribución  $g_A^0$  en el Modelado y Análisis de Imágenes de Radar y de Apertura Sintética. Tesis de Doctorado en Ciencias de la Computación Facultad de Ciencias Exactas y Naturales. UBA.
4. Oliver Chris and Quegan Shaun (1998), Understanding Synthetic Aperture Radar Images. Artech House. Boston. London.