

## Detección de bordes para imágenes SAR polarimétricas

J. I. Fernández Michelli<sup>a,b</sup>, M. Hurtado<sup>a,b</sup>, J. Areta<sup>b,c</sup>, C. H. Muravchik<sup>a,d</sup>.

<sup>a</sup>Laboratorio de Electrónica Industrial, Control e Instrumentación (LEICI),  
Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP).

<sup>b</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).

<sup>c</sup>Universidad Nacional de Río Negro (UNRN).

<sup>d</sup>Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC-pBA).  
e-mail: jfernandez@ing.unlp.edu.ar

**Resumen**— En este trabajo se estudia la detección de bordes en una imagen SAR polarimétrica comparando el valor medio entre dos regiones. Se considera un modelo de datos complejo Gaussiano de media no nula que conduce a la utilización del test de Hotelling generalizado con distribución F. Se evalúa el desempeño del test para datos sintéticos y se verifica su correspondencia con el modelo teórico, independiente de la cantidad de datos considerados. Por otro lado, se aplica el test a datos reales, obteniéndose resultados insatisfactorios. Finalmente, se analizan las causas de los resultados, y se propone una posible mejora del modelo considerando ruido multiplicativo.

**Palabras Clave**— Detección de Bordes, Synthetic Aperture Radar, Test Hotelling Generalizado.

### 1. Introducción

#### 1.1. Radar de Apertura Sintética

El Radar de Apertura Sintética (SAR) es un sistema que sensa la superficie terrestre y forma imágenes de la misma para su estudio. Opera sobre una plataforma aerotransportada o satelital, emitiendo pulsos de radio frecuencia y recibiendo las ondas reflejadas en la superficie terrestre a medida que se desplaza. Mediante un adecuado procesamiento, el conjunto de estas reflexiones es transformado para formar una imagen del terreno, a partir de la cual es posible analizar diversas características del mismo, como humedad, vegetación, rugosidad, entre otras.

La señal recibida por el receptor está determinada por parámetros de dos diferentes tipos: los parámetros relacionados con el objeto iluminado, como su estructura geométrica y características dieléctricas, y por parámetros relacionados con el propio sensor, como la frecuencia de transmisión, estado de polarización y ángulo de incidencia. Por lo tanto, es necesario que los datos obtenidos a partir de las reflexiones sean representados por modelos que reflejen estas interrelaciones. En este sentido, en un sistema SAR polarimétrico la información proveniente de cada punto del terreno es representada por una matriz compleja  $S$  denominada matriz de scattering, cuyos elementos relacionan el campo eléctrico reflejado y

el incidente<sup>1</sup>. Adicionalmente, los datos son de naturaleza aleatoria, cuyas características estadísticas están determinadas por el área iluminada y por el fenómeno físico de la reflexión de las ondas. Por lo tanto, el procesamiento de la información para la formación de imágenes debe tener un enfoque estadístico que contemple estas características.

#### 1.2. Detección de Bordes

La segmentación de datos SAR consiste básicamente en la transformación de una imagen de la escena en una nueva imagen, donde los nuevos píxeles son organizados en diferentes grupos. Cada grupo se define por una característica compartida por datos que lo conforman, reflejando información estructural de la escena subyacente. La detección de los bordes sobre la imagen para delimitar estos grupos es de gran importancia y constituye un tópico de gran actividad en el procesamiento SAR.

Para este propósito, el modelo de datos ampliamente utilizado es el complejo Gaussiano de media nula. Esto implica la necesidad de evaluar momentos de mayor orden y conduce con frecuencia a tests estadísticos con distribuciones asintóticas, que difieren de los resultados obtenidos al considerar una cantidad finita de datos. Por ejemplo en [2] se aplica un test basado en la matriz de covarianza de los datos, asumiendo un modelo Gaussiano de media nula. La distribución del estadístico que se utiliza no tiene una expresión analítica. Para poder evaluar su desempeño debe considerarse su comportamiento asintótico (infinito número de datos), que resulta en una distribución Wishart para la hipótesis nula. Dado que en la aplicación se considera un número finito de píxeles para determinar un borde, para alcanzar el desempeño teórico debe ajustarse manualmente el umbral asociado al test.

Otros enfoques proponen la detección de bordes sin considerar el modelo estadístico subyacente, basándose en la estructura de la imagen directamente, como el cálculo del gradiente o en la transformada Wavelet. Por ejemplo, en [3] se evalúa el gradiente de la imagen en módulo y dirección para cuatro diferentes orientaciones, previo filtrado de los datos con un filtro Gaussiano 2-D para mitigar el ruido. En [4] se utiliza la transformada wavelet

<sup>1</sup>Para una definición rigurosa de la matriz de scattering, consultar [1]

para la detección de bordes. Se aplica el método “soft-thresholding”, que consiste en descartar en cada nivel de descomposición a los coeficientes menores a un umbral y reducir la amplitud al resto. Debido a que este procedimiento tiende a descartar coeficientes con información útil, el trabajo propone aplicarlo a coeficientes preclasificados como no-perteneientes a bordes, preservando así su ubicación y nitidez. Si bien se propone un modelo de ruido speckle para los datos, los ajustes de la preclasificación son completamente ad-hoc, sin tener en cuenta su distribución.

### 1.3. Propuesta

En este artículo se propone un método para la detección de bordes en una imagen SAR polarimétrica, basado en la comparación de las medias entre dos zonas adyacentes de la imagen. Se procesan los datos de la matriz compleja  $\mathbf{S}$  para preservar la información polarimétrica. Se deriva un test de hipótesis de distribución F, cuya expresión es cerrada bajo ambas hipótesis. Esta característica permite predecir su desempeño sin recurrir a expresiones asintóticas y establece un ajuste automático del umbral de detección en función de la probabilidad de falsa alarma deseada, sin la necesidad de ajustes posteriores.

En la primera parte se presenta el modelo de datos utilizado. Se deriva el test de hipótesis junto con la distribución del estadístico resultante, verificando su validez mediante simulación. Luego se muestran los resultados de la aplicación del test a datos SAR simulados y reales, y se concluye con un análisis de la diferencia entre ellos.

## 2. Modelo

En aplicaciones de sensado remoto donde el área de interés (objetivo) del radar es mayor que su celda de resolución y se dispone de diversidad polarimétrica, los datos son representados por la matriz de scattering  $\mathbf{S}$  compleja. Bajo la suposición de que la zona iluminada es homogénea, donde todos los reflectores puntuales sobre la superficie contribuyen en igual proporción a la energía reflejada, es usual adoptar un modelo Gaussiano complejo para cada elemento de la matriz  $\mathbf{S}$ . Esto está justificado bajo tres hipótesis: 1. Existe un gran número de reflectores puntuales en la celda de resolución; 2. La distancia entre el radar y la zona objetivo es mucho mayor que la longitud de onda del radar; 3. La rugosidad de la superficie es menor que la longitud de onda, lo que permite considerar que la fase resultante de cada contribución está uniformemente distribuida en el intervalo  $[-\pi, \pi]$ .

Considerar datos de media nula significa atribuir a momentos de orden superior la diferencia de brillo en la imagen formada. Esto provoca que los estadísticos utilizados en la detección de bordes presenten distribuciones complicadas, debiendo recurrir a expresiones asintóticas para establecer una probabilidad de falsa alarma determinada. Por ejemplo en [5], donde se comparan matrices de covarianza para determinar bordes, se tiene una expresión asintótica para la distribución del estadístico bajo la hipótesis nula, mientras que no se cuenta con una ex-

presión para determinar la probabilidad de detección. El modelo propuesto en este artículo considera datos complejos Gaussianos de media no-nula, atribuyendo a este parámetro el brillo de las diferentes zonas en la imagen.

### 2.1. Datos

Cada dato SAR está representado por la matriz de scattering  $\mathbf{S}$ :

$$\mathbf{S} = \begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix} \quad (1)$$

Cada elemento  $S_{pq}$  es una cantidad compleja que relaciona el campo reflejado con polarización  $q$ , con el campo incidente con polarización  $p$ , para  $p = h, v$ ,  $q = h, v$ , polarizaciones horizontal y vertical respectivamente. Se asume que  $S_{pq}$  tiene distribución compleja normal, de media  $\mu$  y varianza  $\sigma^2$ .

Vectorizando la matriz  $\mathbf{S}$ , se obtiene para cada dato un vector columna complejo  $\mathbf{x} = \{x_i\} = [S_{hh}, S_{hv}, S_{vh}, S_{vv}]^t$ ,  $i = 1, \dots, 4$ . En el caso de radares monoestáticos se cumple que  $S_{hv} = S_{vh}$ , en ese caso puede definirse  $\mathbf{x}$  con 3 componentes. Se considerará aquí el caso general de 4 componentes.

Los datos se asumen independientes entre sí, sin embargo se trata vectorialmente a cada uno para conservar la información que posee la correlación entre sus elementos. Finalmente, el modelo de datos considerado es el siguiente:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad \mathbf{x} \sim CN(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (2)$$

donde  $CN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  significa distribución compleja normal (Gaussiana), de vector media  $\boldsymbol{\mu}$  y matriz de covarianza  $\boldsymbol{\Sigma}$ , cuya expresión es:

$$f dp_{\mathbf{x}}(\mathbf{x}) = \frac{1}{\pi^p |\boldsymbol{\Sigma}|} \exp \left[ (\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right] \quad (3)$$

donde  $(\cdot)^H$  indica conjugado transpuesto,  $|\cdot|$  indica determinante y  $p$  es el número de componentes de  $\mathbf{x}$ . En este caso  $p = 4$ .

### 2.2. Test de Hipótesis

En esta sección se desarrollará un test de hipótesis para el modelo de datos explicado. En [6] se detalla la derivación completa para el caso de datos reales. Dado que su extensión para el caso complejo es directa, se incluirán aquí los puntos más relevantes por completitud.

El problema consiste en comparar la media de dos zonas de la imagen y decidir si son iguales o no. Formalmente, el problema de decisión consiste en elegir una de dos hipótesis posibles: la hipótesis nula  $\mathcal{H}_0$  (la medias comparadas son iguales) o la hipótesis alternativa  $\mathcal{H}_1$  (las medias comparadas son distintas):

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)} \\ \mathcal{H}_1 : \boldsymbol{\mu}^{(1)} \neq \boldsymbol{\mu}^{(2)} \end{cases} \quad (4)$$

En (4)  $\boldsymbol{\mu}^{(i)}$ ,  $i = 1, 2$  indica la media del  $i$ -ésimo grupo de datos a comparar. Si  $\mathcal{H}_0$  se rechaza, entonces el pixel en medio de las dos zonas corresponde a un borde.

El test se desarrolla en base al *cociente de verosimilitud*. Dado que los parámetros  $\boldsymbol{\mu}$  y  $\boldsymbol{\Sigma}$  son desconocidos, se utilizarán sus estimadores de máxima verosimilitud (MLE) en las expresiones. Este método se denomina test de *Cociente de Verosimilitud Generalizado* (GLR) [7].

Sean dos muestras aleatorias de  $N$  observaciones  $\{\mathbf{x}_\alpha^{(i)}\}$  del tipo de datos definido en (2);  $i = 1, 2$ ;  $\alpha = 0 \dots N$  con  $N > p$ . Es decir cada vector  $\mathbf{x}_\alpha^{(i)}$  tiene distribución  $CN(\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma}^{(i)})$ . Se define el vector aleatorio  $\mathbf{y}_\alpha$ :

$$\mathbf{y}_\alpha = \mathbf{x}_\alpha^{(1)} - \mathbf{x}_\alpha^{(2)}, \quad \mathbf{y}_\alpha \sim CN(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y) \quad (5)$$

donde  $\boldsymbol{\mu}_y = \boldsymbol{\mu}^{(1)} - \boldsymbol{\mu}^{(2)}$  y  $\boldsymbol{\Sigma}_y = \boldsymbol{\Sigma}^{(1)} + \boldsymbol{\Sigma}^{(2)}$ . El objetivo del test es de decidir si  $\boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^{(2)}$  o, equivalentemente,  $\boldsymbol{\mu}_y = 0$ . La función de verosimilitud para  $N$  observaciones es

$$L(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y) = \frac{1}{\pi^{pN} |\boldsymbol{\Sigma}_y|^N} \exp \left[ - \sum_{\alpha=1}^N (\mathbf{y}_\alpha - \boldsymbol{\mu}_y)^H \boldsymbol{\Sigma}_y (\mathbf{y}_\alpha - \boldsymbol{\mu}_y) \right] \quad (6)$$

y el cociente de verosimilitud resulta

$$\lambda = \frac{\max_{\boldsymbol{\Sigma}_y} L(\mathbf{0}, \boldsymbol{\Sigma}_y)}{\max_{\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y} L(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y)} \quad (7)$$

El numerador de (7) resulta de maximizar la función de verosimilitud con respecto a  $\boldsymbol{\Sigma}_y$  bajo la hipótesis nula, es decir, cuando las medias son iguales, mientras que el denominador resulta de maximizarla respecto de  $\boldsymbol{\mu}_y$  y  $\boldsymbol{\Sigma}_y$ , bajo la hipótesis alternativa. Los máximos ocurren cuando los parámetros son iguales a los estimadores de máxima verosimilitud de  $\boldsymbol{\mu}_y$  y  $\boldsymbol{\Sigma}_y$ . Debido a que las distribuciones involucradas son Gaussianas, los estimadores resultan ser las conocidas medias y matrices de covarianzas muestrales:

$$\begin{aligned} \mathcal{H}_1 : \quad \hat{\boldsymbol{\mu}}_y^{\mathcal{H}_1} &= \bar{\mathbf{y}} = \frac{1}{N} \sum_{\alpha=1}^N \mathbf{y}_\alpha \\ \hat{\boldsymbol{\Sigma}}_y^{\mathcal{H}_1} &= \frac{1}{N} \sum_{\alpha=1}^N (\mathbf{y}_\alpha - \bar{\mathbf{y}})(\mathbf{y}_\alpha - \bar{\mathbf{y}})^H \end{aligned} \quad (8)$$

$$\mathcal{H}_0 : \quad \hat{\boldsymbol{\Sigma}}_y^{\mathcal{H}_0} = \frac{1}{N} \sum_{\alpha=1}^N \mathbf{y}_\alpha \mathbf{y}_\alpha^H$$

Reemplazando los resultados en (6) se obtiene:

$$\begin{aligned} \max_{\boldsymbol{\Sigma}_y} L(\mathbf{0}, \boldsymbol{\Sigma}_y) &= \frac{1}{\pi^{pN} |\hat{\boldsymbol{\Sigma}}_y^{\mathcal{H}_0}|^N} e^{-pN} \\ \max_{\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y} L(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y) &= \frac{1}{\pi^{pN} |\hat{\boldsymbol{\Sigma}}_y^{\mathcal{H}_1}|^N} e^{-pN} \end{aligned} \quad (9)$$

Entonces el cociente de verosimilitud resulta:

$$\begin{aligned} \lambda &= \frac{|\hat{\boldsymbol{\Sigma}}_y^{\mathcal{H}_1}|^N}{|\hat{\boldsymbol{\Sigma}}_y^{\mathcal{H}_0}|^N} = \\ &= \frac{|\sum (\mathbf{y}_\alpha - \bar{\mathbf{y}})(\mathbf{y}_\alpha - \bar{\mathbf{y}})^H|^N}{|\sum \mathbf{y}_\alpha \mathbf{y}_\alpha^H|^N} = \\ &= \frac{|\mathbf{A}|^N}{|\mathbf{A} + N\bar{\mathbf{y}}\bar{\mathbf{y}}^H|^N} \end{aligned} \quad (10)$$

donde  $\mathbf{A} = \sum_{\alpha=1}^N (\mathbf{y}_\alpha - \bar{\mathbf{y}})(\mathbf{y}_\alpha - \bar{\mathbf{y}})^H$ . Aplicando sobre (10) el corolario A.3.1, pág. 638 de [6], se obtiene

$$\lambda^{-N} = \frac{1}{1 + N\bar{\mathbf{y}}^H \mathbf{A}^{-1} \bar{\mathbf{y}}} = \frac{1}{1 + \frac{T^2}{N-1}} \quad (11)$$

donde

$$T^2 = (N-1)N\bar{\mathbf{y}}^H \mathbf{A}^{-1} \bar{\mathbf{y}} = N\bar{\mathbf{y}}^H \mathbf{S}^{-1} \bar{\mathbf{y}} \quad (12)$$

siendo  $\mathbf{S} = \frac{\mathbf{A}}{N-1}$ . La potencia 2 en (12) está sólo para enfatizar que el estadístico no puede tomar valores negativos. Aplicando el corolario 5.21 de [6], se verifica que el estadístico  $T^2$  tiene la siguiente distribución<sup>2</sup>:

$$\frac{T^2}{N-1} \frac{N-p}{p} \sim \begin{cases} \mathcal{F}_{2p, 2(N-p)} & \text{bajo } \mathcal{H}_0 \\ \mathcal{F}'_{2p, 2(N-p)}(2\gamma) & \text{bajo } \mathcal{H}_1 \end{cases} \quad (13)$$

donde se denota con  $\mathcal{F}_{\nu_1, \nu_2}$  a la distribución  $F$  con  $\nu_1$  y  $\nu_2$  grados de libertad, y  $\mathcal{F}'_{\nu_1, \nu_2}(\nu_3)$  a la distribución  $F$  no central, con  $\nu_1$  y  $\nu_2$  grados de libertad y parámetro de no centralidad  $\nu_3$ . En (13)  $\gamma$  es igual a:

$$\gamma = 2\boldsymbol{\mu}_y^H \boldsymbol{\Sigma}_y \boldsymbol{\mu}_y \quad (14)$$

El test se define sobre la región crítica (zona de rechazo) dada por  $T^2 \geq T_0^2$ , donde  $T_0^2$  es elegido de manera que la probabilidad de (13) cuando la hipótesis nula se cumple sea igual a la probabilidad de falsa alarma deseada. Por lo tanto, el desempeño de la detección está determinado por la siguiente expresión:

$$\begin{cases} P_{FA} = Q_{\mathcal{F}_{2p, 2(N-p)}}(\delta) \\ P_D = Q_{\mathcal{F}'_{2p, 2(N-p)}(2\gamma)}(\delta) \end{cases} \quad (15)$$

donde  $P_{FA}$  denota la probabilidad de falsa alarma,  $P_D$  es la probabilidad de detección,  $Q_{fdp}$  es la distribución acumulada complementaria de la distribución  $fdp$  y  $\delta$  es el umbral de detección.

### 3. Validación

En esta sección se realizará la validación numérica de los resultados de la sección anterior mediante simulaciones en computadora. Se comparan las distribuciones teóricas con el histograma que se obtiene evaluando  $T^2$  con los datos simulados. El procedimiento es el siguiente:

<sup>2</sup>El corolario citado aplica a datos reales, pero puede extenderse fácilmente al caso complejo. El resultado es la duplicación de los parámetros en la distribución.

1. Se generan  $N > p$  datos del tipo descrito en (2), con media y varianza elegidas aleatoriamente.
2. Se evalúa (12). Esto constituye una realización del estadístico  $T^2$ .
3. Se repite los pasos 1 y 2  $L$  veces, para obtener  $L$  realizaciones de  $T^2$  y se realiza el histograma.
4. Finalmente se compara el histograma con las distribuciones teóricas obtenidas en (13).

El procedimiento anterior se repitió para varios valores de  $N$  y  $L$ , para ambas hipótesis, obteniendo en todos ellos coincidencia entre el histograma y la distribución teórica, independientemente de la cantidad  $N$  de datos considerados. La Fig.1 muestra los resultados obtenidos para  $L = 10000$  realizaciones del estadístico, y  $N = 10$ , bajo ambas hipótesis. Se observa que el histograma de  $T^2$  coincide con la distribución teórica en ambos casos.

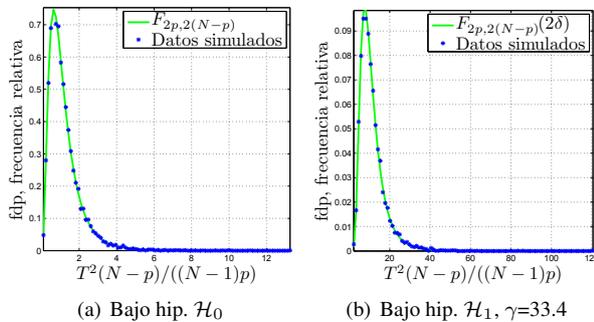


Figura 1: Histograma de  $T^2$ .

#### 4. Aplicación

En esta sección se aplicará la detección de bordes basada en el test propuesto. En primer lugar se explicará el algoritmo utilizado para la detección de bordes. Luego se mostrarán los resultados de su aplicación a una imagen formada a partir de datos simulados por computadora y posteriormente sobre una imagen SAR real de la misión SARAT, provista por la Comisión Nacional de Actividades Espaciales (CONAE).

##### 4.1. Detector de Bordes

El algoritmo de detección de bordes está basado en el propuesto en [5], y está implementado para detectar bordes en dos orientaciones perpendiculares (horizontal y vertical). Naturalmente puede extenderse a más orientaciones, pero no se han implementado en el presente artículo para mantener el mínimo costo computacional que permita obtener conclusiones válidas. En [5] se detalla cómo aumentar el número de orientaciones si fuera necesario.

El detector de bordes se aplica sobre una imagen SAR polarimétrica de  $N_f$  por  $N_c$  píxeles en escala de grises y devuelve una imagen binaria de  $N_f$  por  $N_c$  píxeles, con valor 1 donde ha detectado un borde y 0 en caso contrario. Su esquema de funcionamiento es el siguiente:

1. Se define una ventana de filtrado de  $L$  por  $W$  píxeles, donde  $L \cdot W = N > p$ , se establece la  $P_{FA}$  deseada y se elige una orientación.
2. Se calcula el umbral  $\delta$  que corresponde a  $P_{FA}$  mediante la primera sentencia de (15).
3. Para cada píxel de la imagen de entrada, se consideran los  $N$  píxeles definidos por la ventana a cada lado del píxel considerado. Se evalúa el estadístico  $T^2$  dado en (12).
4. Se repite el paso anterior para cada orientación.
5. Se crea una nueva imagen, asignando el máximo de los estadísticos calculados a la posición del píxel considerado.
6. Se compara el valor cada píxel de la nueva imagen con  $\delta$ . Si es mayor o igual, significa que pertenece a la región crítica del test, por lo tanto se rechaza la hipótesis nula y se detecta un borde asignando el valor 1. En caso, contrario, se asigna 0.

#### 4.2. Datos Sintetizados

El algoritmo descrito en la sección anterior se aplicó en primer lugar a imágenes de datos sintetizados de 100 por 100 píxeles. El valor de cada uno se genera utilizando el modelo indicado en (2). Las imágenes consisten en 4 zonas diferentes y se analizan dos casos: en el primero cada zona es caracterizada con una media diferente, elegida aleatoriamente, y todas las zonas tienen la misma covarianza. En el segundo cada zona posee medias y covarianzas diferentes, elegidas aleatoriamente (Fig. 2). La

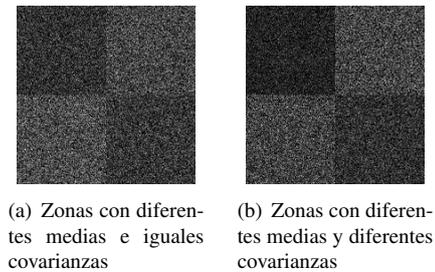


Figura 2: Imágenes de datos sintetizados.

disposición de las zonas se eligió de manera que la fronteras entre ellas tengan la misma orientación que las consideradas en el detector de bordes (horizontal y vertical). Se espera que el detector sea capaz de identificar las fronteras entre zonas con diferente media. Dado que el test estima las matrices de covarianza de los datos, se espera que pueda diferenciar las medias de forma independiente a la covarianza.

La Fig. 3 muestra el resultado del paso 5 del detector de bordes. Son imágenes en escala de grises con los valores que toma el estadístico  $T^2$  en cada píxel. Se observa claramente cómo los valores altos corresponden a los píxeles de frontera entre cada zona, mientras que en el resto el valor es más bajo, indicando ausencia de bordes.

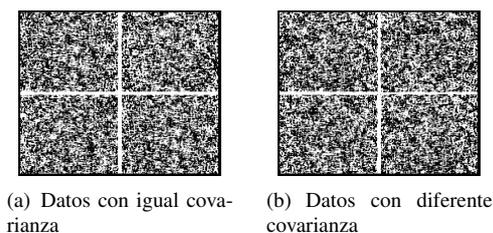


Figura 3: Valor del estadístico luego de aplicar el detector.

La Fig. 4 muestran las imágenes binarias finales para cada caso, para  $P_{FA} = 10^{-3}$ . Se observa la detección correcta de los bordes, confirmando el resultado esperado en ambos casos: para igual covarianza y para diferente covarianza para cada zona.

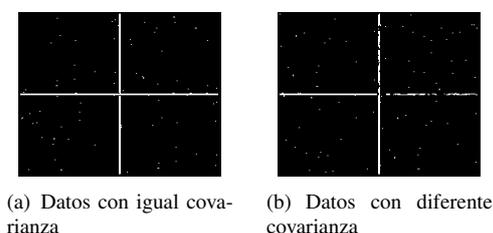


Figura 4: Resultados del detector para  $P_{FA} = 10^{-3}$ . El color blanco corresponde a  $\mathcal{H}_1$  y el negro a  $\mathcal{H}_0$ .

### 4.3. Datos Reales

El mismo algoritmo se aplicó sobre una imagen SAR polarimétrica real. Es una imagen parcial *Single Look Complex* que se obtuvo a partir del enfoque de datos crudos (*Raw Data*) de la misión SARAT, provistos por CONAE [8]. La Fig. 5 muestra los datos  $S_{hh}$  en magnitud.

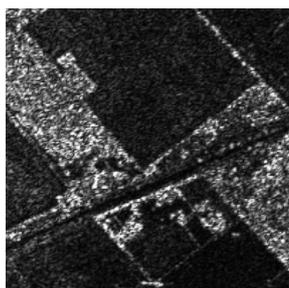


Figura 5: Imagen SAR. Magnitud de  $S_{hh}$ .

El resultado se muestra en la Fig. 6. Se observa que los resultados no son los esperados. El estadístico toma valores altos aún en zonas homogéneas, detectando bordes donde no los hay. Una posible explicación a este resultado se expone en la siguiente sección.

## 5. Discusión

Se ha desarrollado un detector de bordes basado en la comparación de medias. Bajo la hipótesis de gaussianidad de los datos complejos, la distribución del estadístico

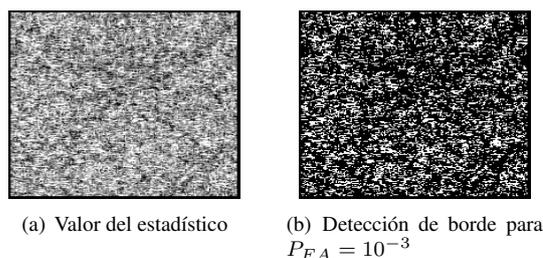


Figura 6: Resultados del detector para datos reales.

que se obtiene es bien conocida y de expresión cerrada, para ambas hipótesis, permitiendo caracterizar el desempeño sin recurrir a expresiones asintóticas. El detector ha funcionado exitosamente en imágenes generadas por computadora, como era de esperarse. Sin embargo, ese no fue el caso para imágenes SAR reales. El motivo más evidente es que la hipótesis de gaussianidad con media diferente de cero es muy fuerte, de difícil cumplimiento en los datos SAR, aún en las zonas más homogéneas donde se ha probado. Si bien zonas con diferentes niveles de brillo presentan diferentes medias, la varianza varía en forma proporcional, lo que indica un modelo multiplicativo subyacente que difiere del considerado. Dicho modelo es ampliamente utilizado para explicar el fenómeno de ruido speckle que se produce debido a la naturaleza coherente de la señal reflejada ([9],[10],[11]).

El modelo multiplicativo considera a cada dato como el producto de dos variables aleatorias  $Z = XY$ , donde  $X$  modela la *retrodispersión* del terreno (BackScatter) e  $Y$  modela el ruido *Speckle* ([9]). La distribución asignada a  $X$  depende del grado de homogeneidad del terreno y la de  $Y$  depende de qué tipo de dato se está considerando (amplitud, intensidad o datos complejos). En [11] se extiende el modelo de speckle a la matriz de covarianza muestral  $C$ .

Considerando el trabajo futuro, una de las opciones para obtener mejores resultados con las imágenes reales sin modificar radicalmente el test, consiste en aprovechar su robustez frente a la distribución de datos subyacente (Teorema 5.2.3. de [6]). Esta propiedad indica que para muestras de gran tamaño, puede prescindirse de la hipótesis Gaussiana. En ese caso el estadístico se distribuye asintóticamente como  $\chi^2$  (Chi-cuadrado) con  $p$  grados de libertad, bajo la hipótesis  $\mathcal{H}_0$ . Esta propiedad podría permitir comparar medias de diferentes zonas, considerando el modelo multiplicativo adecuado. Sin embargo, se estaría tratando con una distribución asintótica, lo que obliga a trabajar con grandes muestras ( $N$  grande). Esto ocasiona que pixeles más alejados del pixel considerado influyan en la estimación. Si existe algún borde en la zona de la muestra, el estadístico incrementará su valor en toda esa zona, provocando que el borde detectado sea de mayor ancho, con la consecuente pérdida de resolución en la ubicación del borde.

### Agradecimientos

Este trabajo fue financiado a través de ANPCyT PICT 2011-11-0909, UNLP 11-I-166, CONICET PIP 2010-2012 y CICpBA.

### Referencias

- [1] J. S. Lee and E. Pottier, *Polarimetric Radar Imaging: from Basics to Applications*. CRC Press, Taylor and Francis Group, 2009, ch. 3, pp. 55–61.
- [2] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, “A test statistic in the complex wishart distribution and its application to change detection in polarimetric SAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 1, pp. 4–19, January 2003.
- [3] Z. Wenge, L. Fang, J. Licheng, and G. Xinbo, “SAR image despeckling based on edge detection and nonsubsampling second generation bandelets,” *Journal of Systems Engineering and Electronics*, vol. 20, no. 3, pp. 519–526, 2009.
- [4] M. Rosa-Zurera, A. M. Cóbreces-Álvarez, J. C. Nieto-Borge, M. P. Jarabo-Amores, and D. Mata-Moya, “Wavelet denoising with edge detection for speckle reduction in SAR images,” in *Eusipco*, Poznan, Poland, 2007.
- [5] J. Schou, H. Skriver, A. A. Nielsen, and K. Conradsen, “CFAR edge detector for polarimetric SAR images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 1, no. 1, pp. 20–31, January 2003.
- [6] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed., ser. Wiley Series in Probability and Statistics, W. Interscience, Ed. Wiley Interscience, 2003.
- [7] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, ser. Signal Processing Series, A. V. Oppenheim, Ed. Upper Saddle River, NJ 07458: Prentice Hall, 1998, vol. 2.
- [8] “Comisión Nacional de Actividades Espaciales, proyecto SARAT,” <http://www.conae.gov.ar/satelites/sacom/sarat.html>.
- [9] A. C. Frery, S. J. S. Sant’Anna, N. D. A. Mascarenhas, and O. H. Bustos, “Robust inference techniques for speckle noise reduction in 1-look amplitude SAR images,” *Applied Signal Processing*, vol. 4, pp. 61–76, 1997.
- [10] A. C. Frery, C. da Costa Freitas Yanasse, and S. Sant’Anna, “Alternative distributions for the multiplicative model in SAR images,” in *Geoscience and Remote Sensing Symposium, 1995. IGARSS ’95. ’Quantitative Remote Sensing for Science and Applications’*, *International*, vol. 1, 1995, pp. 169–171.
- [11] C. Lopez-Martinez and X. Fabregas, “Polarimetric SAR speckle noise model,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 10, pp. 2232–2242, 2003.