

Estimación de Parámetros utilizando Distancias Estocásticas para Datos con Ruido Speckle

J. Casseti¹, J. Gambini^{2,3}, A. Frery⁴

¹Instituto de Desarrollo Humano, Universidad Nacional de General Sarmiento,
Los Polvorines, Pcia. de Buenos Aires, Argentina

²Depto. de Ingeniería en Informática, Instituto Tecnológico de Buenos Aires,
Buenos Aires, Argentina

³Depto. de Ingeniería en Computación, Universidad Nacional de Tres de Febrero,
Caseros, Pcia. de Buenos Aires, Argentina

⁴Laboratório de Computação Científica e Análise Numérica, Universidade Federal de
Alagoas, Maceió, Brasil

{julia.cassetti, juliana.gambini, afrery}@gmail.com

Resumen Las imágenes obtenidas con dispositivos que poseen iluminación coherente como ultrasonido, laser o radar de apertura sintética son afectadas por el ruido speckle, el cual es no gaussiano, no aditivo y difícil de eliminar. Para analizar este tipo de imágenes es necesario utilizar distribuciones estadísticas que modelen los datos en forma adecuada y en este sentido la distribución \mathcal{G}_I^0 es una buena elección porque bajo este modelo pueden caracterizarse regiones con diferente grado de rugosidad a través de un parámetro. Luego, la estimación de parámetros cumple un rol fundamental. En este trabajo se propone y analiza una nueva estrategia para la estimación del parámetro de rugosidad de la distribución \mathcal{G}_I^0 por medio de la minimización de distancias estocásticas entre la distribución empírica de los datos y la función de densidad de la distribución \mathcal{G}_I^0 . En este estudio utilizamos muestras de tamaño pequeño y mediano. Los experimentos se realizan con las distancias de Hellinger, Rényi y Triangular y se comparan con el estimador de máxima verosimilitud.

1. Introducción

Los dispositivos de captura de imágenes que emplean iluminación coherente introducen ruido speckle, como en las imágenes de ultrasonido B, sonar, laser y de radar de apertura sintética – SAR (*Synthetic Aperture Radar*). Las imágenes SAR son muy importantes porque pueden ser utilizadas, para la inspección de lugares de difícil acceso, para detección de talas y cultivos ilegales o inspección de la superficie oceánica, pero tienen la desventaja de que poseen ruido speckle. El ruido speckle no es ni gaussiano ni aditivo, y la interpretación automática de imágenes con este tipo de ruido es un problema de difícil resolución [16].

Para analizar e interpretar imágenes con ruido speckle es necesario utilizar distribuciones estadísticas que describan los datos adecuadamente. Se han

presentado en la literatura diversos modelos, aplicados a clasificación de imágenes [15], detección de bordes [11], reconocimiento y detección de objetos [6] y segmentación [9]. El modelo más exitoso es el multiplicativo, y en los últimos años se ha usado la familia de distribuciones \mathcal{G}^0 para datos de amplitud (\mathcal{G}_A^0) y de intensidad (\mathcal{G}_I^0) [10]. Este modelo permite describir las áreas muy rugosas o extremadamente rugosas mucho mejor que la distribución \mathcal{K} [15]. En el artículo [12] el autor discute una gran variedad de modelos para este tipo de imágenes, con especial énfasis en la distribución \mathcal{G}^0 .

Bajo el modelo \mathcal{G}^0 , regiones con diferente grado de homogeneidad pueden ser caracterizadas por sus parámetros y por la tanto la precisión en la estimación de los mismos es muy importante. Existen varios métodos de estimación de parámetros, Vasconcellos et al. [23] cuantifican el error en la estimación del parámetro de rugosidad para la distribución \mathcal{G}_A^0 y proponen una técnica analítica para mejorar la estimación. Silva et al. [21] proponen otro método analítico para mejorar la estimación del parámetro de rugosidad que reduce el error cuadrático medio. Cribari-Neto et al. [5] proponen el uso de bootstrap para el mismo fin. Allende et al. [1] y Bustos et al. [4] plantean mejoras para la estimación del parámetro de rugosidad, pero con foco en su robustez. Sin embargo, todos estos métodos aumentan el costo computacional de la estimación y no funcionan adecuadamente bajo contaminación. Un problema que surge en la estimación de los parámetros de la distribución \mathcal{G}_I^0 es la falta de convergencia en los métodos de optimización, especialmente para muestras de tamaño pequeño. Frery et al. [8] y Pianto & Cribari-Neto [17] proponen técnicas que tratan este problema pero agregando costo computacional. De ahí la importancia de buscar nuevos métodos de estimación de parámetros para muestras pequeñas.

Por otro lado, la teoría de la información ha sido aplicada a los métodos de estadística y probabilidades con éxito [14]. Shannon [20] definió la información $I(X, Y)$ entre las variables aleatorias X e Y como una divergencia calculada entre sus densidades de probabilidad. Estas divergencias fueron ampliamente estudiadas por Kullback y Leibler [13] y por Rényi [19], entre otros. Este tipo de divergencias poseen múltiples aplicaciones en procesamiento de señales e imágenes [2], análisis de imágenes médicas [24], clasificación de texturas [18], restauración de imágenes [3] e incluso en detección automática de regiones con diferente grado de rugosidad en imágenes SAR [7,22].

En este trabajo proponemos una técnica de estimación con potencial de robustez. El estimador será aquel punto del espacio paramétrico que minimice una distancia entre el histograma de los datos y la densidad que caracteriza el modelo. El costo computacional de evaluar cada punto es relativamente pequeño, puesto que el histograma está fijo y la densidad no involucra funciones especiales. La robustez (que no es evaluada en este trabajo) deviene de la relativa poca sensibilidad del histograma ante desvíos del modelo. La idea principal de este trabajo es estimar el parámetro de rugosidad de la distribución \mathcal{G}_I^0 por medio de la minimización de distancias estocásticas entre la distribución empírica de los datos y la función de densidad de la distribución. Utilizamos muestras de tamaño pequeño y mediano. Los experimentos se realizan con las distancias de

Hellinger, Rényi y Triangular, y los resultados se comparan con el estimador de Máxima Verosimilitud.

Este artículo está compuesto de la siguiente manera: en la Sección 2 se presenta la distribución \mathcal{G}_I^0 y sus principales propiedades. Con el objetivo de que el artículo sea autocontenido, en la Sección 3 se muestra el estimador de Máxima Verosimilitud del parámetro de rugosidad. En la Sección 4 se presentan las distancias estocásticas utilizadas en este trabajo y el algoritmo de estimación del parámetro; esta sección constituye el aporte más importante de este artículo. En la Sección 5 se muestran los resultados obtenidos y, finalmente, en la Sección 6, se extraen conclusiones y se presentan trabajos futuros.

2. El Modelo \mathcal{G} y la distribución \mathcal{G}_I^0

En esta Sección presentamos una breve explicación del modelo multiplicativo, el ruido speckle y la distribución \mathcal{G}_I^0 , para mayor información sobre este tema ver [16]. En lo que sigue, tener en cuenta que $\Gamma(\cdot)$ es la función Gamma y que $\Gamma(s, t)$ es la distribución Gamma con parámetros (s, t) .

El modelo multiplicativo indica que una variable aleatoria Z puede modelarse como el producto de dos variables aleatorias independientes $Z = XY$, donde X e Y corresponden el backscatter y el ruido speckle, respectivamente. La variable aleatoria Y se modela con la distribución $\Gamma(L, L)$, donde $L \geq 1$ es el número equivalente de looks, mientras que si X sigue una distribución Gaussiana Inversa Generalizada, denotada por $\mathcal{N}^{-1}(\alpha, \lambda, \gamma)$, tenemos que Z obedece una distribución \mathcal{G}_I [10]. Para valores particulares de los parámetros de la distribución \mathcal{N}^{-1} se obtienen las distribuciones $\Gamma(\alpha, \lambda)$, $\Gamma^{-1}(\alpha, \gamma)$ e $IG(\gamma, \lambda)$ (Gaussiana Inversa), que llevan a las distribuciones \mathcal{K} , \mathcal{G}_I^0 y \mathcal{G}^H para Z , respectivamente.

En este trabajo utilizamos la distribución para datos de intensidad \mathcal{G}^0 , denotada por \mathcal{G}_I^0 . Su función de densidad está dada por:

$$f_{\mathcal{G}_I^0}(z) = \frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} \cdot \frac{z^{L-1}}{(\gamma + zL)^{L-\alpha}},$$

donde $-\alpha, \gamma, z > 0$ y $L \geq 1$. A diferencia del modelo \mathcal{K} , esta densidad no involucra funciones especiales, por lo que el costo computacional de evaluarla es comparativamente pequeño. Los momentos de orden r son

$$E(z^r) = \left(\frac{\gamma}{L}\right)^r \frac{\Gamma(-\alpha - r)}{\Gamma(-\alpha)} \cdot \frac{\Gamma(L + r)}{\Gamma(L)}. \quad (1)$$

Una de las características más importantes de la distribución \mathcal{G}^0 , tanto para datos de intensidad o amplitud es que puede interpretarse el parámetro α en términos de la textura de la zona observada. Para valores de α cercanos a cero (típicamente en el intervalo $(-3, 0)$), la zona de la imagen corresponde a una región muy texturada, como es el caso de las zonas urbanas en las imágenes SAR. A medida que el valor del parámetro α disminuye, corresponde a zonas con cada vez menos textura, como son las regiones de forestación (usualmente

4 J. Cassetti, J. Gambini, A. Frery

$(-6, -3]$ y pastura (en $(-\infty, -6)$). Esta es la razón por la cual este parámetro se llama el parámetro de rugosidad y también la razón por la cual su estimación precisa es de suma importancia en el análisis de imágenes con ruido speckle.

3. Estimador de Máxima Verosimilitud

Sea $\mathbf{z} = (z_1, \dots, z_n)$ una muestra de tamaño n , la función de verosimilitud bajo la distribución $\mathcal{G}_I^0(\alpha, \gamma, L)$ está dada por

$$\mathcal{L}(\alpha, \gamma, L, \mathbf{z}) = \left(\frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} \right)^n \prod_{i=1}^n z_i^{L-1} (\gamma + Lz_i)^{\alpha-L}.$$

Con el objetivo de maximizar la función de verosimilitud $\mathcal{L}(\alpha, \gamma, L, \mathbf{z})$ en función de los parámetros α y γ , se le aplica logaritmo a la función \mathcal{L} , con lo que los estimadores se calculan maximizando la función $\mathcal{LL}(\alpha, \gamma, L, \mathbf{z})$ dada por

$$\begin{aligned} \mathcal{LL}(\alpha, \gamma, L, \mathbf{z}) = & n[L \log L + \log \Gamma(L - \alpha) - \alpha \log \gamma + \log \Gamma(-\alpha) + \log \Gamma(L)] \\ & + (L - 1) \sum_{i=1}^n \log z_i + (\alpha - L) \sum_{i=1}^n \log(\gamma + Lz_i). \end{aligned}$$

En este trabajo consideramos el valor del parámetro γ que hace que el momento de primer orden de la distribución \mathcal{G}_I^0 sea 1. Esto se realiza con el doble propósito de comparar conjuntos de datos que posean un parámetro de escala cuya esperanza sea la unidad, y al mismo tiempo simplificar los cálculos. Utilizando la Ec. (1) para $r = 1$, obtenemos $\gamma = 1 - \alpha$. Luego el máximo se encuentra únicamente en función del parámetro α . Entonces, el estimador de Máxima Verosimilitud del parámetro α , llamado $\hat{\alpha}_{MV}$ es la solución de la siguiente ecuación no lineal

$$\begin{aligned} & \psi^0(\hat{\alpha}_{MV}) - \psi^0(L - \hat{\alpha}_{MV}) - \log(1 - \hat{\alpha}_{MV}) + \frac{\hat{\alpha}_{MV}}{1 - \hat{\alpha}_{MV}} + \\ & \frac{1}{n} \sum_{i=1}^n \log(1 - \hat{\alpha}_{MV} + Lz_i) - \frac{\hat{\alpha}_{MV} - L}{n} \sum_{i=1}^n \frac{1}{1 - \hat{\alpha}_{MV} + Lz_i} = 0, \end{aligned}$$

donde $\psi^0(\cdot)$ es la función digamma.

4. Estimador por Minimización de Distancias Estocásticas

En esta sección presentamos las fórmulas de las distancias estocásticas que utilizamos para hallar estimadores $\hat{\alpha}$ de la distribución \mathcal{G}_I^0 . Observamos el comportamiento de las curvas de distancias para diferentes valores del parámetro y minimizamos las distancias entre la distribución empírica y la densidad del modelo en función de α .

Las distancias estocásticas permiten comparar dos distribuciones de probabilidad. En este trabajo las utilizamos para comparar la función de densidad de

la distribución \mathcal{G}_I^0 con la distribución empírica. De esta forma, es posible hallar el mínimo de la distancia en función del parámetro α , lo cual provee una forma de estimación del parámetro.

Sean V y W dos variables aleatorias definidas sobre el mismo espacio de probabilidad cuyas funciones de densidad son $f_V(x; \theta_1)$ y $f_W(x; \theta_2)$. Se definen las siguientes distancias estocásticas:

1. Distancia de Hellinger $d_H(V, W) = 1 - \int_{-\infty}^{\infty} \sqrt{f_V f_W}$.
2. Distancia de Bhattacharyya $d_B(V, W) = -\log \int_{-\infty}^{\infty} \sqrt{f_V f_W}$.
3. Distancia Triangular $d_T(V, W) = \int_{-\infty}^{\infty} \frac{(f_V - f_W)^2}{f_V + f_W}$.
4. Distancia de Rényi con parámetro $\beta \in (0, 1)$

$$d_R^\beta(V, W) = \frac{1}{2(\beta - 1)} \log \int_{-\infty}^{\infty} (f_V^\beta f_W^{1-\beta}) + \log \int_{-\infty}^{\infty} (f_V^{1-\beta} f_W^\beta).$$

Observamos que las distancias de Hellinger y Bhattacharyya obedecen $d_B = -\log(1 - d_H)$. Como además la función logaritmo es creciente resulta que

$$\arg \min_{\alpha} d_B(\alpha) = \arg \min_{\alpha} d_H(\alpha)$$

por lo que las distancias de Battacharya y Hellinger poseen el mismo mínimo y por eso utilizamos solamente la distancia de Hellinger que tiene menor costo computacional.

Calculamos las curvas de las distancias entre las funciones de densidad de la distribución $\mathcal{G}_I^0(\alpha_0, 1 - \alpha_0, L)$ y $\mathcal{G}_I^0(\alpha, 1 - \alpha, L)$, $\alpha < -1$, para distintos valores de α_0 y de L . La Figura 1 muestra los gráficos de las distancias de Hellinger, Rényi y Triangular entre dos distribuciones \mathcal{G}_I^0 , para distintos valores del número looks $L = \{1, 8\}$ y para $\alpha_0 = \{-2, -3, -4, -5, -6, -7\}$, graficadas en colores magenta, amarillo, cian, azul, verde y rojo, respectivamente. En estos gráficos puede observarse la variación de las curvas de distancias para los distintos valores de α_0 y del número de looks. En el caso de $L = 1$ cuanto menor es el valor de α_0 la curva se hace más plana en un entorno de α_0 y por lo tanto en el momento de hallar el mínimo, los métodos de optimización son más inestables y el mínimo más difícil de encontrar en forma precisa. A medida que aumenta el valor de L esta situación se revierte y es posible encontrar el mínimo de una forma eficiente.

Para estimar el parámetro α se toma el conjunto de datos $\{z_1, \dots, z_n\}$ de tamaño n . Se calcula la distribución empírica de los datos $p(z)$ y finalmente, se calcula

$$\hat{\alpha} = \arg \min_{\alpha} dK(f_{\mathcal{G}_I^0}(z, \alpha, 1 - \alpha, L), p(\mathbf{z}))$$

donde dK es alguna de las distancias consideradas.

La distribución empírica $p(z)$ se calcula utilizando el histograma de los datos y utilizando el método de Freedman-Diaconis.

5. Resultados Experimentales

Para evaluar el desempeño de los estimadores propuestos se realizan mil ensayos Monte Carlo independientes para varios valores de α , y se obtienen

6 J. Casseti, J. Gambini, A. Frery

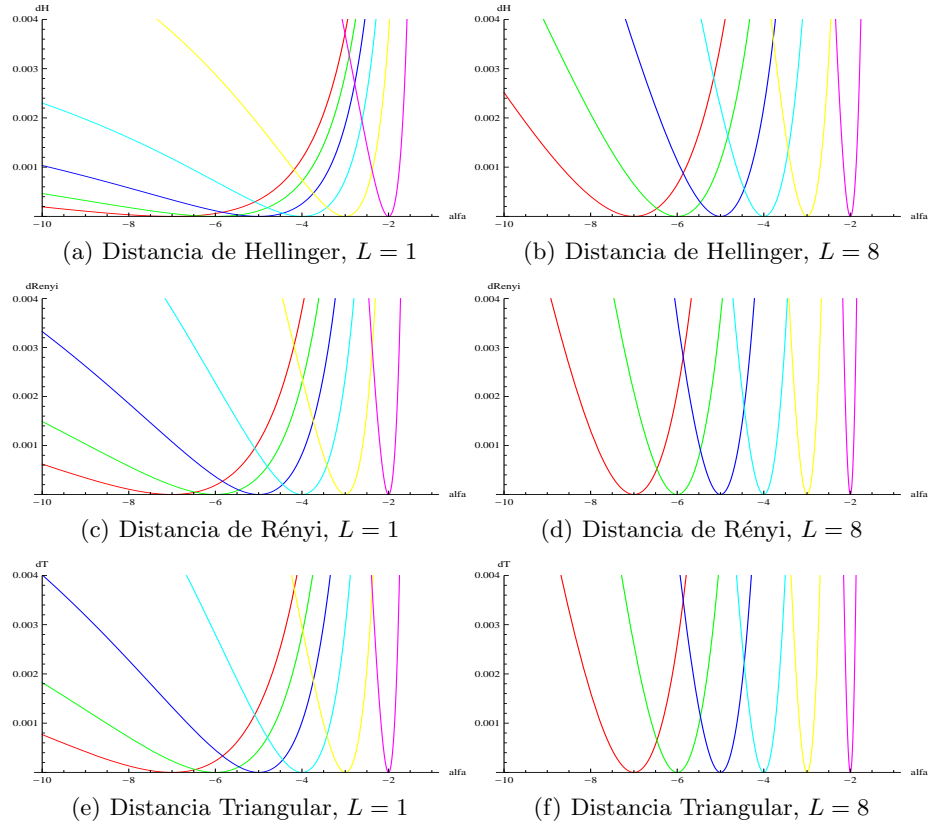


Figura 1. Comparación entre los gráficos de las curvas de distancia, $d(f_{\mathcal{G}_L^0}(z, \alpha_0, L), f_{\mathcal{G}_L^0}(z, \alpha, L))$, $\alpha < -1$, $\alpha = \{-2, -3, -4, -5, -6, -7\}$, $L = \{1, 8\}$ en magenta, amarillo, cian, azul, verde y rojo.

$\{\hat{\alpha}_1, \dots, \hat{\alpha}_{1000}\}$. Se calcula la media de las estimativas $\bar{\hat{\alpha}} = (1000)^{-1} \sum_{i=1}^{1000} \hat{\alpha}_i$, se estima el sesgo $B(\hat{\alpha}) = E(\hat{\alpha}) - \alpha$ con $\hat{B}(\hat{\alpha}) = \bar{\hat{\alpha}} - \alpha$ y el error cuadrático medio con $\widehat{\text{ecm}} = (1000)^{-1} \sum_{i=1}^{1000} (\hat{\alpha}_i - \alpha)^2$.

En este trabajo las pruebas se realizan utilizando los siguientes valores de parámetros:

- Textura: $\alpha = \{-1, 5, -3, -5, -8\}$. Con estos valores se describen regiones extramadamente rugosas, rugosas y homogéneas.
- Looks: $L = \{1, 3, 8\}$, para modelar varios niveles de procesamiento.
- Tamaño de las muestras: $n = \{9, 25, 49, 81, 121, 1000\}$. La elección de estos tamaños se basa en que muchos de los métodos de filtrado de imágenes o detección de bordes utilizan máscaras deslizantes para estimar los parámetros, las cuales suelen ser de tamaño 3×3 , 5×5 , 7×7 , 9×9 u 11×11 . El tamaño de la muestra $n = 1000$ se elige para poder hacer una comparación con la estimación tomando una muestra grande.

La Figura 2 muestra las medias de las estimativas del parámetro de rugosidad $\hat{\alpha}$ con las distancias de Hellinger, Rényi, Triangular y con Máxima Verosimilitud, para distintos tamaños de muestras n y números equivalentes de looks L . La línea recta azul es el valor verdadero con que se genera la muestra. Los resultados se muestran en escala semilogarítmica para su mejor comprensión visual. En el gráfico de $L = 1$ observamos que la estimación se aleja del valor verdadero a medida que el valor de α se hace más pequeño. Esto se justifica con la observación de la Figura 1 que muestra lo planas que resultan las curvas de distancia en este caso. Observamos también en estos gráficos que los métodos de minimización de distancias poseen menor error para valores grandes del número de looks y para valores cercanos a cero del parámetro α_0 , que corresponden a las regiones muy rugosas de la imagen. Sin embargo no se observan diferencias considerables entre los resultados obtenidos con las diferentes distancias.

Para evaluar la calidad del estimador consideramos el error cuadrático medio. La Figura 3 muestra una comparación entre los errores cuadráticos medios para $L = 1$ y para los cuatro métodos analizados con las variaciones de parámetros. Este caso es el más difícil de estimar y podemos observar en el gráfico que el aumento del tamaño de la muestra no necesariamente implica un estimador mejor.

La Figura 4 muestra la comparación entre los errores cuadráticos medios para los cuatro métodos analizados con las variaciones de parámetros, para $L = 3$ y $L = 8$. Nuevamente observamos que el aumento en el número de looks aumenta la calidad del estimador, y que en este caso el tamaño de muestra sí cumple un rol preponderante, en el sentido de que a medida que aumenta el tamaño de muestra, el error disminuye considerablemente.

La Figura 5 muestra el resultado de aplicar el algoritmo a una imagen de 80×80 pixels simulada con $L = 8$ looks. Las 40 primeras columnas fueron generadas con $\alpha = -8, \gamma = 9$, y las segundas con $\alpha = -1,5, \gamma = -2,5$, ésto es, tienen la misma media. Para cada pixel se estima α utilizando la distancia de Hellinger y un entorno de 3×3 elementos alrededor de cada pixel, es decir que la muestra en cada caso es de 9 elementos.

6. Conclusiones y Trabajos Futuros

En este trabajo se propone utilizar distancias estocásticas para la estimación del parámetro de rugosidad de la distribución \mathcal{G}_I^0 . Se realizan las estimaciones con las distancias de Hellinger, Rényi y Triangular. En los resultados puede observarse que la estimación está dentro de un rango adecuado para la rugosidad correspondiente a los datos generados, sin embargo, en algunos casos el estimador de Máxima Verosimilitud resulta más preciso. Esto ocurre especialmente para valores de $L = 1$, en los cuales la curva de distancia se hace más plana y los métodos de minimización son más inestables. Sin embargo, el estimador de máxima verosimilitud puede ser muy difícil de encontrar cuando se aplica a datos reales debido a la correlación que los mismos poseen, en las muestras simuladas los datos se generan en forma independiente y por eso el método es muy

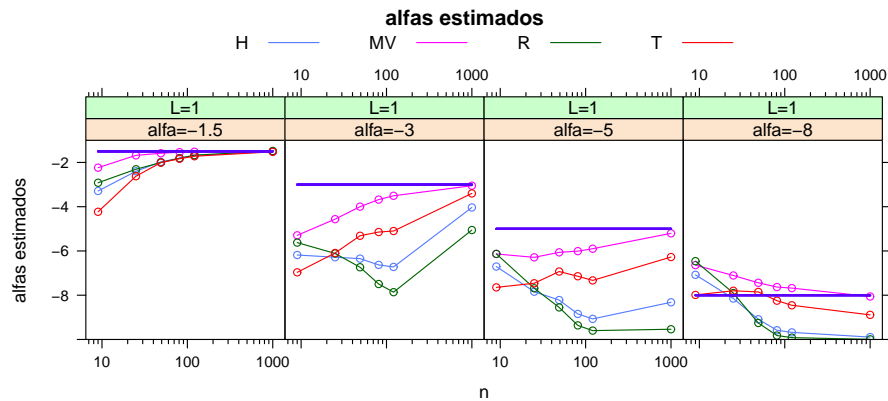
eficiente. Como trabajos futuros experimentaremos con otras distancias, con datos contaminados para medir la robustez de los estimadores propuestos y con datos provenientes de imágenes reales, analizando el resultado de la estimación y comparándolo con otros estimadores, como por ejemplo el de los momentos.

Referencias

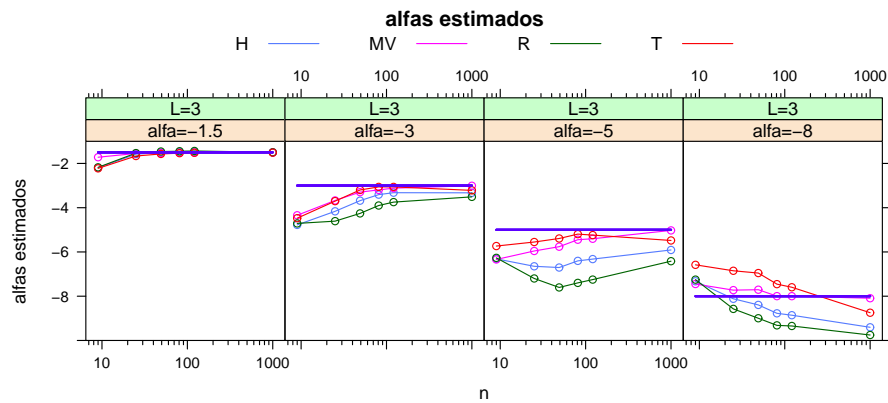
1. H. Allende, A. C. Frery, J. Galbiati, and L. Pizarro. M-estimators with asymmetric influence functions: the GA0 distribution case. *Journal of Statistical Computation and Simulation*, 76(11):941–956, 2006.
2. S. Aviyente, F. Ahmad, and M. G. Amin. Information theoretic measures for change detection in urban sensing applications. In *IEEE Workshop on Signal Processing Applications for Public Security and Forensics*, pages 1–6, 2007.
3. E. Bratsolis and M. Sigelle. Kullback-Leibler divergence and Markov random fields for speckled image restoration. In *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings*, volume 1, pages 425–428, 2003.
4. O. H. Bustos, M. M. Lucini, and A. C. Frery. M-estimators of roughness and scale for GA0-modelled SAR imagery. *EURASIP Journal on Applied Signal Processing*, 1(1):105–114, 2002.
5. F. Cribari-Neto, A. C. Frery, and M. F. Silva. Improved estimation of clutter properties in speckled imagery. *Computational Statistics and Data Analysis*, 40(4):801–824, 2002.
6. A. Farrouki and M. Barkat. Automatic censoring CFAR detector based on ordered data variability for nonhomogeneous environments. *IEE Proceedings in Radar, Sonar and Navigation*, 152(1):43–51, 2005.
7. A. C. Frery, R. J. Cintra, and A. D. C. Nascimento. Entropy-based statistical analysis of PolSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6), in press.
8. A. C. Frery, F. Cribari-Neto, and M. O. Souza. Analysis of minute features in speckled imagery with maximum likelihood estimation. *EURASIP Journal on Applied Signal Processing*, 1(16):2476–2491, 2004.
9. A. C. Frery, E. D. V. Lucca, C. C. Freitas, and S. J. S. Sant’Anna. SAR segmentation algorithms: A quantitative assessment. In *International Geoscience and Remote Sensing Symposium: Remote Sensing of the System Earth – A Challenge for the 21st Century*, pages 1–3, Hamburg, Germany, 1999. IEEE, IEEE. CD-ROM.
10. A. C. Frery, H.-J. Müller, C. C. F. Yanasse, and S. J. S. Sant’Anna. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):648–659, 1997.
11. J. Gambini, M. Mejail, J. Jacobo-Berlles, and A. C. Frery. Accuracy of edge detection methods with local information in speckled imagery. *Statistics and Computing*, 18(1):15–26, 2008.
12. G. Gao. Statistical modeling of SAR images: A survey. *Sensors*, 10(1):775–795, 2010.
13. S. Kullback and R. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86, 1951.
14. F. Liese and I. Vajda. On divergences and informations in statistics and information theory. *IEEE Transactions on Information Theory*, 52(10):4394–4412, 2006.
15. M. Mejail, J. C. Jacobo-Berlles, A. C. Frery, and O. H. Bustos. Classification of SAR images using a general and tractable multiplicative model. *International Journal of Remote Sensing*, 24(18):3565–3582, 2003.

16. C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar Images*. Artech House, Boston, 1998.
17. D. M. Pianto and F. Cribari-Neto. Dealing with monotone likelihood in a model for speckled data. *Computational Statistics and Data Analysis*, 55:1394–1409, 2011.
18. D. Puig and M.A. Garcia. Pixel classification through divergence-based integration of texture methods with conflict resolution. In *International Conference on Image Processing*, volume 3, pages 1037–1040, 2003.
19. A. Rényi. On measures of entropy and information. In University of California Press, editor, *Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability*, volume 1, pages 547–561, 1961.
20. C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.
21. M. Silva, F. Cribari-Neto, and A. C. Frery. Improved likelihood inference for the roughness parameter of the GA0 distribution. *Environmetrics*, 19(4):347–368, 2008.
22. W. B. Silva, C. C. Freitas, S. J. S. Sant’Anna, and A. C. Frery. Classification of segments in PolSAR imagery by minimum stochastic distances between Wishart distributions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, in press.
23. K. L. P. Vasconcellos, A. C. Frery, and L. B. Silva. Improving estimation in speckled imagery. *Computational Statistics*, 20(3):503–519, 2005.
24. B.C. Vemuri, Meizhu Liu, S.-I. Amari, and F. Nielsen. Total bregman divergence and its applications to DTI analysis. *IEEE Transactions on Medical Imaging*, 30(2):475–483, 2011.

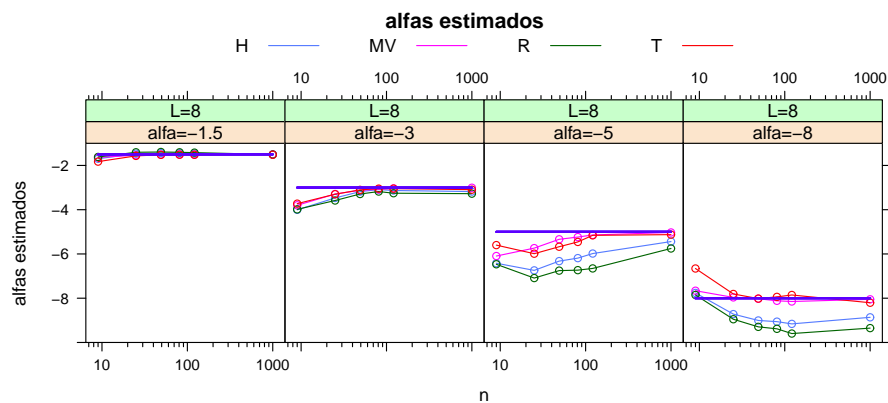
10 J. Casseti, J. Gambini, A. Frery



(a) $L = 1$



(b) $L = 3$



(c) $L = 8$

Figura 2. Estimación del parámetro $\hat{\alpha}$ para las distancias de Hellinger, Rényi y Triangular para muestras generadas con distintos valores de α_0 y del número equivalente de looks L . La línea recta azul es el valor del parámetro con que la muestra fue generada.

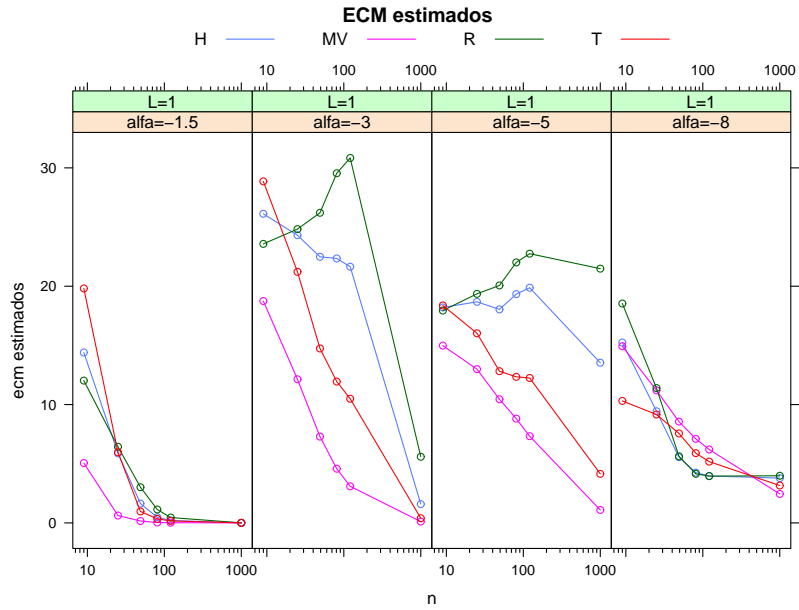


Figura 3. Error cuadrático medio para cada uno de los métodos, utilizando 1000 experimentos, para cada tamaño de muestra y para cada α_0 , con $L = 1$.

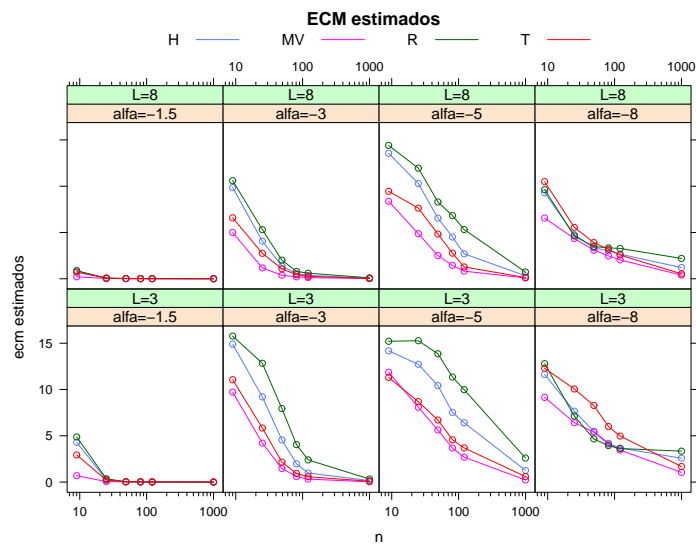


Figura 4. Error cuadrático medio para cada uno de los métodos, utilizando 1000 experimentos, para cada tamaño de muestra y para cada α_0 , con $L = 3$ y $L = 8$.

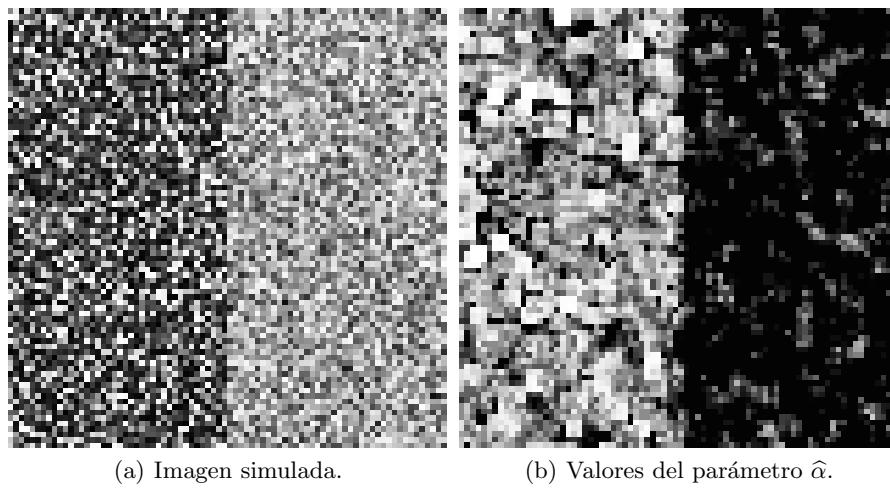


Figura 5. Resultado de aplicar el método de estimación de la distancia de Hellinger (b), y con la distancia de R enyi (c) para cada pixel de la imagen, tomando muestras de 9 elementos, a una imagen simulada.