ANALYSIS OF TEXTURED POLSAR DATA BY SHANNON ENTROPY

T. Eltoft, A. Doulgeris, and S. N. Anfinsen

Department of Physics and Technology, University of Tromsø, Norway

ABSTRACT

The Shannon entropy has previously been used in the analysis of polarimetric synthetic aperture radar (PolSAR) data, assuming a complex Gaussian distribution for the scattering vector. According to the maximum entropy characterization theorem, the deviation from Gaussian statistics can be quantified as a reduction of the Shannon entropy. This paper derives the Shannon entropy within the framework of the non-Gaussian, multivariate product model and discusses information theoretical interpretations. An experiment demonstrates the information contents associated with the non-Gaussian entropy component.

1. INTRODUCTION

Analysis of PolSAR data puts much emphasis on algorithms and interpretations rooted in *target decomposition theorems*, by which information about physical scattering mechanisms can be achieved. Another strategy focuses on multi-dimensional statistical modeling. This avenue adds statistical information about the scene, and is in some respect complementary to the target decomposition approach. Statistical properties are widely used for image segmentation and land cover classification (e.g.[1]). In a recent paper, Morio et al. [2] demonstrated the usefulness of applying information theoretic based contrast measures in the analysis of PolSAR data. Their analysis assumes a Gaussian distribution for the scattering vector, $\mathbf{S}$. It is well-known that among all distributions with a given covariance matrix $\mathbf{\Gamma}_s$, the Gaussian model has the maximum entropy [3]. Hence, in analysis of local *Shannon entropy* of a scene, areas with non-Gaussian statistics will have reduced entropy compared to the Gaussian model. This paper discusses Shannon entropy within the framework of the non-Gaussian, multivariate product model with focus on information theoretical interpretations, and experimental information contents.

2. THE GAUSSIAN CASE

In the literature, the complex circular multivariate Gaussian model has successfully been taken as the default hypothesis, especially in the case of low resolution SAR images covering homogenous regions. If we assume a d-dimensional, zero-mean scattering vector, $\mathbf{S}$, the probability density function (pdf) becomes

$$p_{\mathbf{S}}(\mathbf{s}) = \frac{1}{\pi^d |\mathbf{\Gamma}_s|} \exp(-\mathbf{s}^\dagger\, \mathbf{\Gamma}_s^{-1}\, \mathbf{s}), \tag{1}$$

where $|\mathbf{\Gamma}_s|$ is the determinant of the covariance matrix $\mathbf{\Gamma}_s = E\{\mathbf{S}\mathbf{S}^\dagger\}$, and the superscript $\dagger$ refers to conjugate transpose. The Shannon entropy (also known as the differential entropy) associated with the pdf of (1) is given as [4]

$$\begin{aligned} H_d^G(\mathbf{S}) &= -\int \log[p_{\mathbf{S}}(\mathbf{s})]\, p_{\mathbf{S}}(\mathbf{s}) d\mathbf{s} \\ &= \log\left[\pi^d e^d |\mathbf{\Gamma}_s|\right] = \log\left[\pi^d e^d\right] + \log[|\mathbf{\Gamma}_s|]. \end{aligned} \tag{2}$$

We introduce the dimensionless parameter δ defined as $\delta = |\mathbf{\Gamma}_s)|/\mathrm{tr}(\mathbf{\Gamma}_s)$, where $\mathrm{tr}(\mathbf{\Gamma}_s)$ is the trace of $\mathbf{\Gamma}_s$, as in [2]. This allows the entropy to be decomposed into three terms as

$$H_d^G(\mathbf{S}) = d\, \log[\pi I_0] + \log[\delta] + d, \tag{3}$$

where $I_0 = tr(\mathbf{\Gamma}_s)$ is equal to the total backscattered power, the so-called span, which is the sum of the intensities of the individual channels. The first term represents the dependence of the entropy on the disorder due to fluctuations in the intensities, and is a function of the total energy. The second term is linked to disorder due to decorrelation of the components of the backscattered vector. The third term is simply the dimension of $\mathbf{S}$.

3. THE NON-GAUSSIAN CASE

The multiplicative or product model has been widely used in non-Gaussian modeling, processing, and analysis of single- and multi-polarimetric SAR images. The model states that, under certain conditions, the complex-valued backscattered signal results as the product between a complex, circular Gaussian speckle noise component and a random scalar texture component. Several distributions could be used to model SAR texture with different spatial correlation properties and various degrees of inhomogeneity [1, 5].

The scattering vector, $\boldsymbol{S}$, is hence formulated as

$$\boldsymbol{S} = \sqrt{T}\boldsymbol{X}, \tag{4}$$

where T is a the positive scalar texture random variable with unit mean, accounting for heterogeneity in the scattering cross-section, and $\boldsymbol{X}$ is a zero mean, complex Gaussian vector variable, representing the speckle. The distribution of $\mathbf{S}$ will generally have a heavy-tailed, symmetric pdf, whose shape will depend on the actual pdf of the texture T. The Shannon entropy associated with this non-Gaussian model can be shown to consist of the following terms:

$$\begin{aligned} H_d^{nG}(\mathbf{S}) &= H_d(\mathbf{S}|T) + H(T) - H(T|\mathbf{S}) \\ &= H_d(\mathbf{S}|T) + I(\mathbf{S}, T), \end{aligned} \tag{5}$$

where

$$H_d(\mathbf{S}|T) = -\int p_T(\tau)\{\int \log[p_{\mathbf{S}|T}(\mathbf{s}|\tau)]\, p_{\mathbf{S}|T}(\mathbf{s}|\tau)\, d\mathbf{s}\}\, d\tau, \tag{6}$$

$$H(T) = -\int \log[p_T(\tau)] p_T(\tau)\, d\tau, \tag{7}$$

$$H(T|\mathbf{S}) = -\int p_{\mathbf{S}}(\mathbf{s})\{\int \log[p_{T|\mathbf{S}}(\tau|\mathbf{s})]\, p_T(\tau|\mathbf{s})\, d\tau\}\, d\mathbf{s}, \tag{8}$$

and $I(\mathbf{S}, T)$ in (5) denotes the mutual information between $\mathbf{S}$ and T. Mutual information is a nonnegative quantity, and zero if and only if the variables are statistically independent.

Now, let $\text{cov}(\mathbf{X}) = E\{\mathbf{X}\mathbf{X}^\dagger\} = \boldsymbol{\Gamma}_x$. Since the conditional variable $\mathbf{S}|T$ is Gaussian distributed, it follows from (2) and (6) that

$$\begin{aligned} H_d(\mathbf{S}|T) &= \int \{\log[\pi^d e^d |\tau \boldsymbol{\Gamma}_x|]\} p_T(\tau)\, d\tau \\ &= d\,\hat{\mu}_T + d\,\log[\pi I_0] + \log[\delta] + d, \end{aligned} \tag{9}$$

where $\hat{\mu}_T = \int \log[\tau] p_T(\tau)\, d\tau$ is the first order log-moment of T. Hence it follows that within the assumptions associated with the product model of (4), the Shannon entropy takes the form

$$H_d^{nG}(\mathbf{S}) = H_d^{G}(\mathbf{S}) + d\,\hat{\mu_T} + I(\mathbf{S}, T). \tag{10}$$

Equation (10) shows that the inclusion of radar texture, to account for variation in the scattering cross-section, adds an extra term to the Shannon entropy. Since $\text{cov}(\mathbf{S}) = \text{cov}(\mathbf{X})$, we denote $\boldsymbol{\Gamma}_s = \boldsymbol{\Gamma}_x = \boldsymbol{\Gamma}$, and rewrite (10) as

$$J(\mathbf{S}) = H_d^{G}(\mathbf{S}) - H_d^{nG}(\mathbf{S}) = -[d\,\hat{\mu_T} + I(\mathbf{S}, T)], \tag{11}$$

where $J(\mathbf{S})$ is referred to as negentropy, which is known from information theory to be a measure of deviation from Gaussian statistics, as is evident from (11). The Kullback-Leibler divergence between $p^{nG}(\mathbf{s})$ and $p^{G}(\mathbf{s})$ is defined as

$$\begin{aligned} K^{nG,G}(\mathbf{S}) &= \int p^{nG}(\mathbf{S}) \log\left[\frac{p^{nG}(\mathbf{S})}{p^{G}(\mathbf{S})}\right] d\mathbf{s} \\ &= -H^{nG}(\mathbf{S}) + H^{nG,G}(\mathbf{S}), \end{aligned} \tag{12}$$

where $H^{nG,G}(\mathbf{S})$ denotes the cross entropy between $p^{nG}(\mathbf{s})$ and $p^{G}(\mathbf{s})$. For the product model, some simple algebra gives

$$H^{nG,G}(\mathbf{S}) = \log[\pi^d |\boldsymbol{\Gamma}|] + d\mu_T, \tag{13}$$

which, when the mean of T is one, becomes exactly $H^{G}(\mathbf{S})$. The Kullback-Leibler divergence becomes $K^{nG,G}(\mathbf{S}) = H^{G}(\mathbf{S}) - H^{nG}(\mathbf{S})$, and hence is a proper measure of non-Gaussianity.

4. ENTROPY IN COMPLEX SYMMETRIC MULTIVARIATE ELLIPTICAL DISTRIBUTIONS

Let $\mathbf{Z}$ be a complex, elliptical random vector in $\mathbb{C}^d$ with mean vector $\boldsymbol{\mu}_z$ and covariance matrix $\boldsymbol{\Gamma}$. The pdf of $\mathbf{Z}$ can then be expressed as

$$p_{\mathbf{Z}}(\mathbf{z}) = c_d |\boldsymbol{\Gamma}|^{-1} h\{(\mathbf{z} - \boldsymbol{\mu}_z)^\dagger \boldsymbol{\Gamma}^{-1} (\mathbf{z} - \boldsymbol{\mu}_z)\}, \tag{14}$$

where the non-negative, real-valued function $h(\cdot)$ is called *density generator*, and c_d is a normalizing constant [7]. It can be shown that the pdf of the squared radial random variable $U = (\mathbf{Z} - \boldsymbol{\mu}_z)^\dagger \boldsymbol{\Gamma}^{-1} (\mathbf{Z} - \boldsymbol{\mu}_z)$ can be written as

$$g_U(u) = c_d \frac{\pi^d u^{k-1}}{\Gamma(d)} h(u), \tag{15}$$

Model	Texture pdf, $p_T(\tau)$	Density generator, $c_k h(u)$
Gaussian	$p_T(\tau) = \delta(u-1)$	$c_d h(u) = \frac{1}{\pi^d}\exp(-u)$
Gamma	$p_T(\tau) = \frac{\alpha^\alpha}{\Gamma(\alpha)}\tau^{\alpha-1}\exp(-\alpha\tau)$	$c_k h(u) = \frac{1}{\pi^d}\frac{2\alpha^\alpha}{\Gamma(\alpha)}\left(\frac{u}{\alpha}\right)^{\left(\frac{\alpha-d}{2}\right)}\mathcal{K}_{\alpha-d}(2\sqrt{\alpha u})$
Fisher	$p_T(\tau) = \mathcal{B}(\alpha,\lambda)\frac{\alpha}{\lambda-1}\frac{\left(\frac{\alpha\tau}{\lambda-1}\right)^{\alpha-1}}{\left(\frac{\alpha\tau}{\lambda-1}+1\right)^{\alpha+\lambda}}$	$c_d h(u) = \mathcal{B}(\alpha,\lambda)\frac{\Gamma(d+\lambda)}{\pi^d}\frac{\alpha^d}{(\lambda-1)^d}\mathcal{U}(d+\lambda, d-\alpha+1, \frac{\alpha u}{\lambda-1})$
Inverse Gamma	$p_T(\tau) = \frac{(\alpha-1)^\alpha}{\Gamma(\alpha)}\tau^{\alpha+1}\exp(-\frac{\alpha-1}{\tau})$	$c_d h(u) = \frac{1}{\pi^d}\frac{\Gamma(\alpha+d)}{\Gamma(\alpha)}\frac{(\alpha-1)^\alpha}{(\alpha+u-1)^{\alpha+d}}$

Table 1. Summary of some previously used texture models in SAR polarimetry

where $h(u) = 2^d h_r(2u)$, and $h_r(u)$ denotes the density generator of a real variable with the corresponding real distribution, and of the same dimension [6]. Now, assume $p_{\mathbf{Z}}(\mathbf{z}) = p_{\mathbf{Z}_0}(\mathbf{\Gamma}^{-1/2}(\mathbf{z} - \boldsymbol{\mu}_z))$. It then follows that that the Shannon entropy of Z is given as

$$H(\mathbf{Z}) = \frac{1}{2}\log[|\mathbf{\Gamma}|) + H(\mathbf{Z}_0), \quad (16)$$

where $H(\mathbf{Z}_0)$ is the corresponding entropy of the standardized vector $\mathbf{Z}_0 = \mathbf{\Gamma}^{-1/2}(\mathbf{Z} - \boldsymbol{\mu}_z)$. Using (14), we find that

$$\begin{aligned} H(\mathbf{Z}_0) &= -E\{\log[|\mathbf{\Gamma}|^{-\frac{1}{2}} c_d h(\mathbf{Z}_0^\dagger \mathbf{Z}_0)]\} \\ &= \frac{1}{2}\log(|\mathbf{\Gamma}|) - E\{\log(c_d h(U))\}, \quad (17) \end{aligned}$$

where

$$E\{\log(c_d h(U))\} = \int_0^\infty \log[c_d h(u)] g_U(u) du. \quad (18)$$

Combining (16), (17), and (18), we get

$$H(\mathbf{Z}) = \log[|\mathbf{\Gamma}|] - \int_0^\infty \log[c_d h(u)] g_U(u) du. \quad (19)$$

The product model introduced in Eq.(5) is noted to be a subclass of the family of complex symmetric elliptical distribution. By comparing Eqs.(2) and (19), we can write the Shannon entropy in the non-Gaussian product model as

$$H_d^{nG}(\mathbf{S}) = H_d^G(\mathbf{S}) - (\log[\pi^d e^d] + \int_0^\infty \log[c_d h(u)] g(u) du), \quad (20)$$

with $h(u)$ and $g(u)$ properly defined.

4.1. The density generator in product models

Let $\mathbf{X}$ be a multivariate Gaussian variable with mean $\boldsymbol{\mu}_x$ and covariance matrix $\mathbf{\Gamma}$. Then $p_{\mathbf{X}}(x)$ can be written as

$$p_{\mathbf{X}}(\mathbf{x}) = c_d|\mathbf{\Gamma}|^{-1} h_G\{(\mathbf{x} - \boldsymbol{\mu}_x)^\dagger \mathbf{\Gamma}^{-1}(\mathbf{z} - \boldsymbol{\mu}_x)\}, \quad (21)$$

where $h_G(\cdot)$ refers to the Gaussian density generator. It then follows that the density generator of the variable $\mathbf{S} = \sqrt{T}\mathbf{X}$ is given as

$$h(u) = \int_0^\infty \tau^{-k} h_G\left(\frac{u}{\tau}\right) p_T(\tau) d\tau, \quad (22)$$

where $p_T(\tau)$ is the pdf of the texture variable T.
Table 1 gives the density generators associated with some of the most commonly used texture models in SAR polarimetry, where the entries $\mathcal{K}(\cdot)$, $\mathcal{B}(\cdot)$, $\mathcal{U}(\cdot)$, and $\Gamma(\cdot)$ refer to the modified Bessel function of the second kind, the Beta function, the confluent hypergeometric function of the first kind (also known as the Kummer U function), and the Gamma function, respectively. The corresponding pdfs of the radial variable U are obtained from $c_d h(u)$ using Eq.(15).

5. ANALYSIS

Fig.1 shows the various entropy terms associated with a Radarsat-2 scene over the San Francisco area. The entropy calculations are performed by fitting a product model with Fisher distributed texture to the data within a sliding window of size 17×17. The parameters of the texture model are estimated from sample log-cumulants. The figure shows in the upper left panel the Pauli image, and the consecutive panels show the disorder due to fluctuations in intensity (upper right), the disorder due to decorrelation of the polarimetric channels (lower left), and reduction in Shannon entropy due to non-Gaussianity (lower right). As expected, in areas dominated by buildings and man-made structures, the non-Gaussian entropy term clearly has significance.

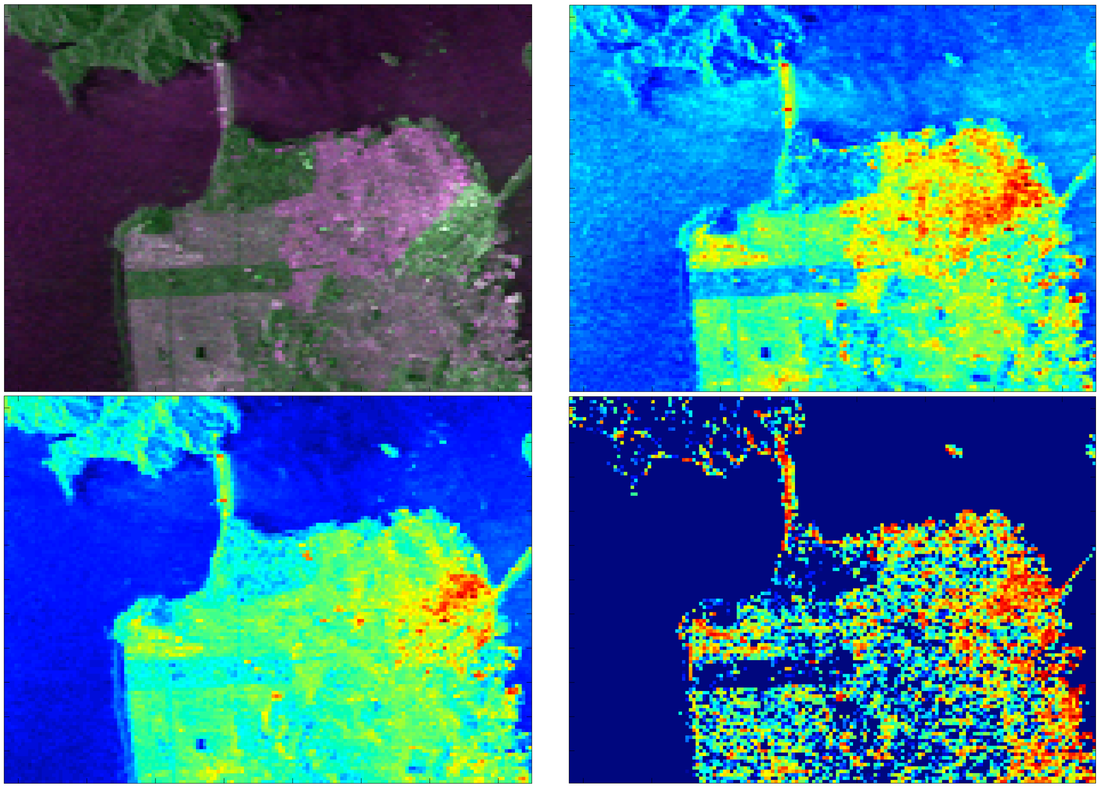

Fig. 1. Pauli image (upper left), intensity disorder (upper right), polarization disorder (lower left), negentropy (lower right).

6. REFERENCES

[1] A. Doulgeris, S. Anfinsen, and T. Eltoft, ”Classification with a non-Gaussian model for polsar data,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 46, no. 10, pp 2999-3009, Oct. 2008.

[2] J. Morio, P. Réfrégier, F. Goudail, P. C. Dubois-Fernandez, and X. Dupuis, ”Information theory-based approach for contrast analysis in polarimetric and/or interferometric SAR images,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 46, no. 8, pp 2999-3009, Aug. 2008.

[3] F.D. Neeser and J.L. Massey, ”Proper complex random processes with applications to information theory,” IEEE Trans. Inf. Theory, vol. 39, no. 7, pp. 1293-1302, Jul. 2003.

[4] P. Réfrégier and F. Goudail, ”Entropy of partially polarized light and application to statistical processing techniques,” *J. Opt. Soc. Am. A*, vol. 21, no. 11, Nov. 2004.

[5] C. C. Freitas, A. C. Frery, and A. H. Correia, ”The polarimetric G distribution for SAR data analysis,” *Environmetrics*, vol. 16, no. 1, pp. 1331, Feb. 2005.

[6] E. Ollila and V. Koivonen, ” Generalized complex elliptical distributions,” *Proceedings of the 3rd IEEE Sensor Array and Multichannel Signal Processing Workshop*, Sitges, Spain. 18-21 July 2004, pages 460-464.

[7] R. B. Arellano-Valle, J. E. contreras-Reyes, and M. G. Genton, ”Shannon entropy and mutual information for multivariate skew-elliptical distributions,” *Scand. J. Statist.*, doi: 10.1111/j.1467-9469.2011.00774.x, February 2012