ROBUST CLASSIFICATION OF SAR IMAGERY

M.M. Lucini, V.F. Ruiz

Department Of Cybernetics
University of Reading
Whiteknights, Reading, RG6 6AY UK

A.C. Frery, O.H. Bustos

Centro de Informática
Universidade Federal de Pernambuco
CP 7851, 50732-970 Recife, PE – Brasil

ABSTRACT

In this work the $\mathcal{G}_A^0$ distribution is assumed as the universal model for amplitude Synthetic Aperture (SAR) imagery data under the Multiplicative Model. The observed data, therefore, is assumed to obey a $\mathcal{G}_A^0(\alpha, \gamma, n)$ law, where the parameter n is related to the speckle noise, and (α, γ) are related to the ground truth, giving information about the background. Therefore, maps generated by the estimation of (α, γ) in each coordinate can be used as the input for classification methods. Maximum likelihood estimators are derived and used to form estimated parameter maps. This estimation can be hampered by the presence of corner reflectors, man-made objects used to calibrate SAR images that produce large return values. In order to alleviate this contamination, robust (M) estimators are also derived for the universal model. Gaussian Maximum Likelihood classification is used to obtain maps using hard-to-deal-with simulated data, and the superiority of robust estimation is quantitatively assessed.

1. INTRODUCTION

Statistical modeling has provided some of the best tools for processing and understanding SAR images. Among the available models, the Multiplicative Model is probably the most successful [1]. It consists of a set of distributions that, with few parameters, are able to characterize SAR data.

Application of the $\mathcal{G}'_{\mathcal{A}}$ distribution to the MM provided a quite general and tractable extension: the Universal Model (UM [2]), and it is known to be a useful tool for describing and classifying SAR returns [3, 4].

In this paper, the $\mathcal{G}_A^0$ distribution is assumed as the UM for amplitude SAR imagery data under the multiplicative model. Thus, the return obeys a $\mathcal{G}_A^0(\alpha, \gamma, n)$ law where the parameters (α, γ) are related to the ground truth (roughness and scale, respectively) and the parameter n is related to the speckle noise.

A robust estimation of these parameters is of paramount importance as SAR images are commonly corrupted by the effect of corner reflectors. These area devices essential to data calibration that result in much higher pixel values than the rest of the image.

The authors have addressed the problem of robust estimation in the single look (noisiest) situation [5], finding M-estimators of (α, γ). It should be noted that M-estimators commonly used with symmetric distributions [6, 7] were derived in a context whereby the density functions under consideration can be highly assymmetrical. Other model widely used in signal processing is the one that uses α-stable distributions, which are highly assymetrical [8]

The approach here proposed employs Maximum Likelihood and M-estimators for the parameters (α, γ) of the $\mathcal{G}_A^0$ distribution. Maps generated using these estimators are used as input to a supervised classification based on Gaussian Maximum Likelihood. This classification scheme, using moment estimators, was proposed in [3].

Section 2 introduces the model together with the main properties of the $\mathcal{G}_A^0$ distribution. The parameter estimation methods are derived in section 3. The supervised classification results are discussed in section 4.

2. NOTATION AND MODEL DEFINITION

The Multiplicative Model is widely used in SAR images processing, and it is a common framework to explain the stochastic behavior of data obtained with coherent illumination as is the case of sonar, laser and B-scan ultrasound. This model states that the return Z is composed by the multiplicative mixture of speckle noise X and backscatter properties Y.

Under this framework, the $\mathcal{G}_A^0$ model stems from assuming that the backscatter X obeys a $\Gamma^{-1/2}(\alpha, \gamma)$ law, while the speckle noise, Y, is independent of X and follows a $\Gamma^{1/2}(n, n)$ law, where n is the number of looks of the image. Thus, the amplitude return Z has a $\mathcal{G}_A^0(\alpha, \gamma, n)$ distribution whose density is given by

$$f_Z(z) = \frac{2n^n\Gamma(n-\alpha)}{\gamma^\alpha\Gamma(-\alpha)\Gamma(n)} \frac{z^{2n-1}}{(\gamma + z^2 n)^{n-\alpha}} \mathbb{I}_{(0,\infty)}(z), \qquad (1)$$

where $-\alpha, \gamma > 0$, $n \geq 1$ and $\mathbb{I}_{(a,b)}()$ denotes the indicator function of the interval (a, b), i.e.

$$\mathbb{I}_{(a,b)}(x) = \begin{cases} 1 & \text{if} \quad x \in (a, b) \\ 0 & \text{else} \end{cases}$$

and $\Gamma(\cdot)$ is the Gamma function $\left(\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx\right)$.

The parameter α in equation (1) describes the roughness of the target. Small values of α ($\alpha < -15$) are usually associated to homogeneous targets, like pasture. Values of α ranging between $[-15, -5]$ usually characterize heterogeneous clutter, like forests, and large values ($-5 < \alpha < 0$ for instance) are observed when extremely heterogeneous areas are imaged. The γ parameter is related to the scale, that is, if $Z' \sim \mathcal{G}_A^0(\alpha, 1, n)$ distributed then $Z = \sqrt{\gamma}Z'$ obeys a $\mathcal{G}_A^0(\alpha, \gamma, n)$ law. The parameter n is related to the multilook processing in SAR images, and is controlled by the image generation process.

The r-th moments of the $\mathcal{G}_A^0(\alpha, \gamma, n)$ distribution are

$$E(Z_A^r) = \left(\frac{\gamma}{n}\right)^{r/2} \frac{\Gamma(-\alpha - r/2)\Gamma(n + r/2)}{\Gamma(-\alpha)\Gamma(n)},$$

if $\alpha < -r/2$ or infinite otherwise for every $n \geq 1$.

This work deals with the noisiest class of amplitude images, i.e., with raw data where no multilook processing is performed. Dropping the parameter $n = 1$, the single-look $\mathcal{G}_A^0(\alpha, \gamma)$ distribution is characterized by the following probability density function:

$$f(x, (\alpha, \gamma)) = \frac{-2\alpha x}{\left(1 + \frac{x^2}{\gamma}\right)^{1-\alpha}} \mathbb{I}_{(0,+\infty)}(x), \qquad (2)$$

where the parameters space is $-\alpha > 0, \gamma > 0$.

The multilook, intensity and complex versions of this distribution can be seen in [2].

The corresponding score function is given by

$$\mathbf{s}(x, (\alpha, \gamma)) = \begin{pmatrix} s_1(x, (\alpha, \gamma)) \\ s_2(x, (\alpha, \gamma)) \end{pmatrix} = \begin{pmatrix} \frac{1}{\alpha} + \ln\left(1 + \left(\frac{x}{\sqrt{\gamma}}\right)^2\right) \\ -\frac{\alpha}{\gamma} - \frac{1-\alpha}{\gamma + x^2} \end{pmatrix}.$$

and the cumulative distribution function is given by

$$F(x, (\alpha, \gamma)) = \left(1 - \left(1 + \frac{x^2}{\gamma}\right)^{\alpha}\right) \mathbb{I}_{(0,+\infty)}(x).$$

This distribution can be derived as the square root of the ratio of two independent random variables: one having a unitary-mean exponential distribution (which conveys the speckle noise in one look) and the other following a $\Gamma(\alpha, \gamma)$ distribution, related to the unobserved ground truth, the backscatter. Densities of this functions are illustrated in figure 1, for the situations $\alpha \in \{-1.5, -4, -10\}$ and the scale parameter γ adjusted to have unitary mean.

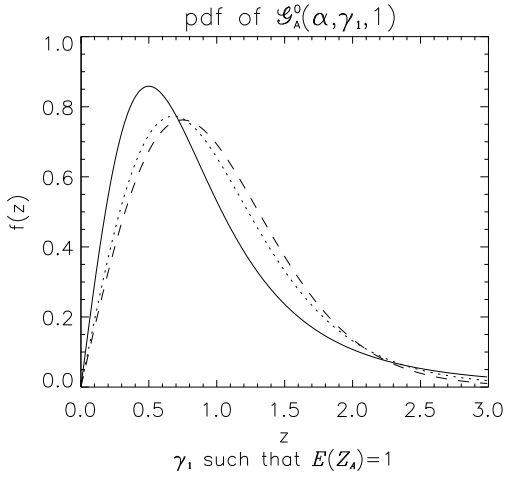

Fig. 1. $\mathcal{G}_A^0$ densities, $\alpha \in \{-1.5, -4, -10\}$ (solid line, dotted, dashes, respectively).

3. INFERENCE FOR THE $\mathcal{G}_A^0$ MODEL

Let $\Theta = \{(\alpha, \gamma) : \alpha < -1.0, \gamma > 0\}$ and $\mathbb{R}$ be the set of real numbers. Let $X_1, \ldots, X_N$ be a sample of independent identically distributed random variables with $\mathcal{G}_A^0(\alpha, \gamma)$ distribution. The maximum likelihood estimators of (α, γ) based on $X_1, \ldots, X_N$, denoted $\widehat{\alpha}_{\mathrm{ML},N}$ and $\widehat{\gamma}_{\mathrm{ML},N}$ respectively, are given by

$$\widehat{\alpha}_{\mathrm{ML},N} = -\left(\frac{1}{N}\sum_{k=1}^{N} \ln\left(1 + \left(\frac{X_k}{\sqrt{\widehat{\gamma}_{\mathrm{ML},N}}}\right)^2\right)\right)^{-1} \qquad (3)$$

and

$$\widehat{\gamma}_{\mathrm{ML},N} = \left(1 + 2\left(\frac{1}{N}\sum_{k=1}^{N} \ln\left(1 + \left(\frac{X_k}{\sqrt{\widehat{\gamma}_{\mathrm{ML},N}}}\right)^2\right)\right)^{-1}\right) \cdot \frac{1}{N}\sum_{k=1}^{N} \frac{X_k^2}{1 + \left(\frac{X_k}{\sqrt{\widehat{\gamma}_{\mathrm{ML},N}}}\right)^2}. \qquad (4)$$

In order to define the M-estimators a contamination model is assumed. That is, $X_1, \ldots, X_N$ is taken to be a sequence of independent identically distributed random variables with cumulative distribution function $F(\cdot, (\alpha, \gamma), \epsilon, z)$ given by

$$F(x, (\alpha, \gamma), \epsilon, z) = (1 - \epsilon) F(x, (\alpha, \gamma)) + \epsilon \delta_z(x) \qquad (5)$$

where $\delta_z(x) = \mathbb{I}_{[z,+\infty)}(x)$, with z a "very large" value with respect to most of the values typically assumed by a random variable with distribution $\mathcal{G}_A^0(\alpha, \gamma)$. Equation (5) describes a contamination that occurs at random with probability ϵ, i.e., in average $N\epsilon$ out of N samples will be "very large" fixed values (z), while $N - N\epsilon$ samples will obey the $\mathcal{G}_A^0(\alpha, \gamma)$ distribution. Other contamination hypothesis may include different distributions for the departure of the model, and/or spatial dependence among observations.

The M-estimators of the parameters (α, γ) based on $X_1, \ldots, X_N$ are given by the solutions $\hat{\alpha}_{\mathrm{M},N}, \hat{\gamma}_{\mathrm{M},N}$ of the system

$$\begin{array}{l} \sum_{k=1}^{N} \psi_{b_1}\left(s_1(X_k, (\widehat{\alpha}_{\mathrm{M,N}}, \widehat{\gamma}_{\mathrm{M,N}})) - c_1(\widehat{\alpha}_{\mathrm{M,N}}, \widehat{\gamma}_{\mathrm{M,N}}, b_1)\right) = 0 \\ \sum_{k=1}^{N} \psi_{b_2}\left(s_2(X_k, (\widehat{\alpha}_{\mathrm{M,N}}, \widehat{\gamma}_{\mathrm{M,N}})) - c_2(\widehat{\alpha}_{\mathrm{M,N}}, \widehat{\gamma}_{\mathrm{M,N}}, b_2)\right) = 0 \end{array} \qquad (6)$$

where $\psi_{b_i} : \mathbb{R} \times \Theta \to [-b_i, b_i]$, $i = 1, 2$ are the Huber's functions:

$$\psi_{b_i}(x) = \begin{cases} -b_i & \text{if } x \le -b_i \\ x & \text{if } -b_i < x < b_i \\ b_i & \text{if } x \ge b_i. \end{cases} \qquad (7)$$

The functions $c_i : \Theta \times (0, +\infty) \to \mathbb{R}$ are defined so that the sequence of estimators $(\widehat{\alpha}_{\mathrm{M,N}}, \widehat{\gamma}_{\mathrm{M,N}})_N$ is asymptotically unbiased, Fisher consistent of (α, γ). That is,

$$\begin{cases} \int \psi_{b_1}\left(s_1(x, (\alpha, \gamma)) - c_1(\alpha, \gamma, b_1)\right) f(x; (\alpha, \gamma)) dx &= 0 \\ \int \psi_{b_2}\left(s_2(x, (\alpha, \gamma)) - c_2(\alpha, \gamma, b_2)\right) f(x; (\alpha, \gamma)) dx &= 0 \end{cases} \qquad (8)$$

for any $(\alpha, \gamma) \in \Theta$, $b_1 > 0$ and $b_2 > 0$. The constants b_1 and b_2 are the *tuning parameters* of the M-estimators.

A comparison among different estimators for roughness and scale parameters of the $\mathcal{G}_A^0$ distribution was performed in [9], where no contamination was considered, and the $\widehat{\alpha}_{\mathrm{ML}}$ estimator was the best estimator in almost all cases in terms of mean square error and bias, assuming γ known. In [5] it is shown that if the sample is contaminated by a percentage of "outliers" the behavior of classical estimators is unreliable, while the performance of M-estimators is good.

4. RESULTS

In this section three classification procedures in SAR images where corner reflectors are present will be compared using Gaussian Maximum Likelihood: the first one using the result of filtering data by an adaptive Lee filter as an input, the second and third ones using the ML- and M-estimates of the parameters (α, γ) of the $\mathcal{G}_A^0$ distribution as input, respectively. In each situation the size of the filtering window was 7×7.

Although there exist several supervised classification techniques (see, for instance, [10]), the one chosen here is one of the most widely used techniques and it is simple to implement.

By using stochastic simulation the three procedures were tested in a simulated SAR image where four different classes were contaminated by corner reflectors. The phantom of four classes (figure 2 left) and the distributions for each class were proposed in [3] as a benchmark for classification procedures assessment. It is used here in order to assess the effect of corner reflectors on the classification, and to show that robust procedures overcome the influence of such contamination.

The results are presented in thematics maps, were each pixel is labeled by a color representing the class it belongs to. The assessment is performed visually and presenting tables (confusion matrices) that summarize the results and allow the comparison among procedures. In [3] different results were obtained mainly due to two facts: (i) no contamination is present, and (ii) estimators based on moments were used.

The data set shown in figure 2 left, was obtained sampling independent observations from four different $\mathcal{G}_A^0(\alpha, \gamma)$ laws, characterized by two values of α and two values of γ. The values of α were chosen to represent heterogeneous ($\alpha_1 = -5.0$, top row of the image) and extremely heterogeneous ($\alpha_2 = -1.5$, bottom row of the image) areas. For each α, the two different values of γ are $\gamma_1 = 2 \cdot 10^5$ and $\gamma_2 = 4 \cdot 10^5$. As it was already mentioned, the number of looks is $n = 1$, the noisiest situation.

Some corner reflectors were introduced in the data set, in order to evaluate the behavior of these procedures when the data are not free of this kind of spurious values.

In figure 2 right it is shown the result of the simulation before the introduction of the corner reflectors, being the colored spots the training areas.

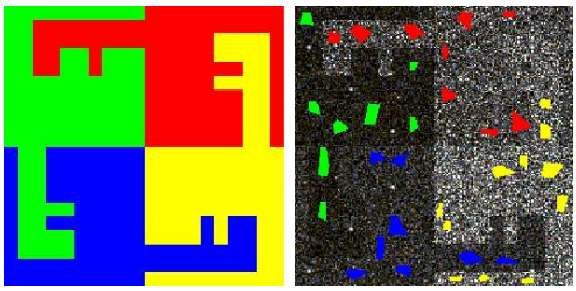

Fig. 2. Phantom and return with training data (colored spots).

The Gaussian Maximum Likelihood algorithm was used to classify the data shown in Figure 2. The first procedure used the Lee filtered data with a sliding window of size 7×7 (Figure 3 left) as an input. It should be noted that the Lee filter is known to fail in providing reliable data for the classification [3] of speckled imagery. In this work, with contaminated data, this classification technique only identified two classes (Figure 3), and therefore its application is again unacceptable.

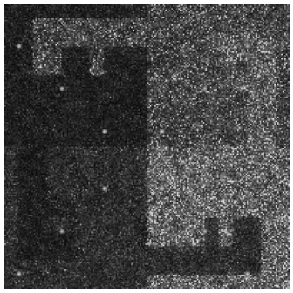

Fig. 3. Lee filtered image.

Figure 4 (left and right) shows the maps that resulted of applying the ML estimators of (α, γ) given by equations (3) and (4) respectively, to the image shown in Figure 2 right. The confussion matrix of this classification is shown in table 1, the overall accuracy is 54.24 %, and $\kappa = 0.39$. The classification map is shown in Figure 5. It can be seen that this classification procedure was not able to eliminate the corner reflectors, some of them were classified as urban area. It can not distinguish between the different scales γ (brightness) in extremely heterogeneous areas.

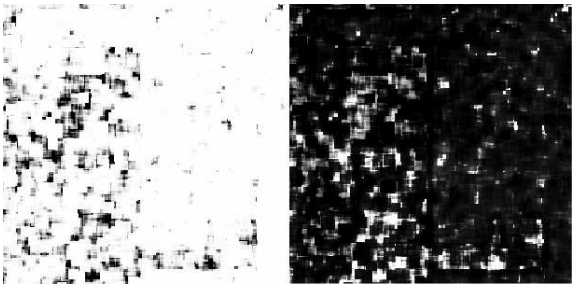

Fig. 4. Left: $\hat{\alpha}_{\mathrm{ML}}$ map, right: $\hat{\gamma}_{\mathrm{ML}}$ map.

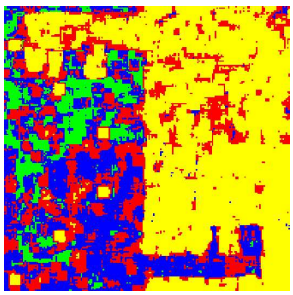

Fig. 5. Gaussian Maximum Likelihood Classification using the $(\hat{\alpha}_{\mathrm{ML}}, \hat{\gamma}_{\mathrm{ML}})$ map.

Figure 7 shows the maps that resulted of applying the M estimators of (α, γ) given by equation (6) to the original image (Figure 6). The confusion matrix of this classification is shown in Table 4, being the overall accuracy $81,12\%$ and $\kappa = 0.74$. It can be seen that this matrix is better than the one obtained with

Class/True	C_1	C_2	C_3	C_4
C_1	22.28	22.57	34.70	1.88
C_2	0.00	37.74	0.89	0.00
C_3	2.60	39.68	64.41	0.00
C_4	75.12	0.00	0.00	98.12

Table 1. Confusion Matrix for the Gaussian Maximum Likelihood Method using the $(\hat{\alpha}_{\mathrm{ML}}, \hat{\gamma}_{\mathrm{ML}})$ map as input.

Class/True	C_1	C_2	C_3	C_4
C_1	92.15	0.00	15.51	32.97
C_2	0.37	88.14	6.44	0.36
C_3	0.37	11.86	78.04	0.18
C_4	7.10	0.00	0.00	66.48

Table 2. Confusion Matrix for the Gaussian Maximum Likelihood Method using the map $(\hat{\alpha}_{\mathrm{M}}, \hat{\gamma}_{\mathrm{M}})$ as input

the $(\hat{\alpha}_{\mathrm{ML}}, \hat{\gamma}_{\mathrm{ML}})$ input. As it was expected, the presence of corner reflectors did not spoil the classification, as happened in the ML procedure. More features and borders are retained with this procedure than using the ML estimators. From the figures presented here, it is possible to note that M-estimators are able to separate regions with different roughness and scale (α, γ) even in the presence of corner reflectors, what leads to a better classification. The numerical problems reported in [3] were overcome for both, ML and M-estimators.

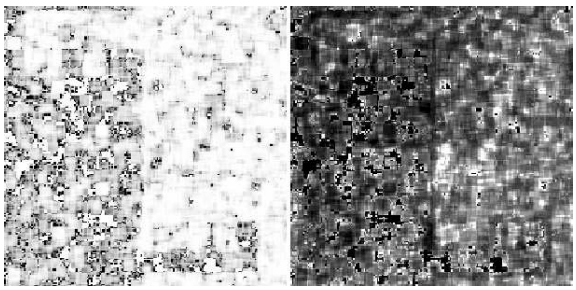

Fig. 6. Left: $\hat{\alpha}_{\mathrm{M}}$ map, right: $\hat{\gamma}_{\mathrm{M}}$ map.

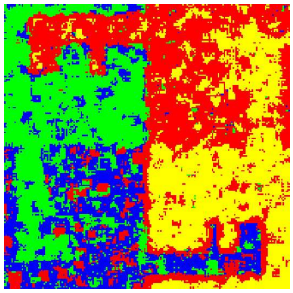

Fig. 7. Gaussian Maximum Likelihood Classification using the $(\hat{\alpha}_{\mathrm{M}}, \hat{\gamma}_{\mathrm{M}})$ map.

5. CONCLUSIONS

In this work Gaussian Maximum likelihood classification was applied to features extracted from a simulated image with four different classes representing four different ground covers contaminated by corner reflectors. Three classification procedures were compared: one using the result of filtering data by an adaptive Lee filter as an input, the second and third ones using the ML-estimates and M-estimates of the parameters (α, γ) of the $\mathcal{G}_A^0$ distribution as input, respectively. The method involving the Lee filter as an input failed to classify the image. Among the other two methods, the M-estimator was the best, in the sense that it classified more accurately the different classes, preserved more details and it was not as corrupted by the presence of corners as the ML one. The results here obtained motivate the use of M-estimators in signal processing techniques, because returns as those caused by corner reflectors are present in the images.

Since speckle noise and corner reflector returns also appear in ultrasound B-scan, sonar and laser images, the procedure here presented has potential application in all these techniques.

6. REFERENCES

[1] C. Oliver and S. Quegan, *Understanding synthetic aperture radar images*, Artech House, Boston, 1998.

[2] 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*, vol. 35 (3), pp. 648–659, 1997.

[3] M. E. 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*, in press.

[4] F. Cribari-Neto, A. C. Frery, and M. F. Silva, "Improved estimation of clutter properties in speckled imagery," *Computational Statistics and Data Analysis*, vol. 40, no. 4, pp. 801–824, 2002.

[5] 0. H. Bustos, M.M Lucini, and A.C. Frery, "M-estimators of roughness and scale for $\mathcal{G}_A^0$-modelled SAR imagery," *EURASIP Journal on Applied Signal Processing*, vol. 1, pp. 105–115, 2002.

[6] F.R. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel, *Robust Statistics-The Approach Based on Influence Functions*, John Wiley & Sons, New York, 1986.

[7] Marazzi A. and Ruffieux C., "Implementing M-estimators of the gamma distribution," *Robust Statistics, Data Analysis and Computer Intensive Methods, Lectures Notes in Statistics*, vol. 109, pp. 277–296, 1996.

[8] C.L. Nikias and M. Shao, *Signal Processing with Alpha Stable Distributions and Applications*, John Wiley & Sons, Inc., New York, 1995.

[9] M. E. Mejail, A. C. Frery, J. Jacobo-Berlles, and O. H. Bustos, "Approximation of distributions for SAR images: proposal, evaluation and practical consequences," *Latin American Applied Research*, vol. 31, pp. 83–92, 2001.

[10] J.A. Richards, *Remote Sensing Digital Image Analysis. An Introduction*, Springer-Verlag, 2 edition, 1994.