Abstract—Synthetic Aperture Radar (SAR) imaging is based on airborne or satellite active microwave sensors that can capture the earth surface by emitting a signal and receiving the backscattered signal that forms the resulting image. Since microwave radiation is not interfered by sunlight and can pass through clouds, SAR imagery can be generated oblivious to weather and daylight conditions. However, the active nature of the imaging process determines that SAR images are contaminated by an inherent speckle noise that may degrade significantly the quality and usefulness of the images, and specific noise-removal processes may also filter out relevant textural information. In this article, we propose a texture-based method that can be applied for region segmentation in SAR imagery. The method is based on local analysis of the multifractal spectrum and a clustering procedure. The outcomes obtained both with synthetic and real SAR images show better region segmentation results than with state-of-the-art proposals.

Index Terms—SAR segmentation, multifractal spectrum, texture analysis

I. INTRODUCTION

Image segmentation is aimed at partitioning an image into different disjoint regions or segments. Each one is a set of pixels assumed to represent a homogeneous region possessing properties that differentiate it from other regions in the image. These regions are uniform in some sense. For instance they share some distinguishing criteria, such as luminance, color, or local texture, among others. Synthetic Aperture Radar is a coherent high resolution RADAR imaging procedure extensively used in remote sensing. Especially in research involving soil monitoring, spaceborne SAR imagery is a valuable information source, given that the available SAR satellite constellations deliver high-resolution images with wide geographic coverage, short revisit times, and in wavelengths that provide meaningful information about specific soil properties. Based on the imaging technology used, SAR images can be of one, two or four polarizations. However, several difficult tasks arise when trying to match SAR-based information together with other information sources, like optical imagery, field-based measurements, simulation models, etc. SAR imagery possesses significant difficulties in segmentation tasks, due to the specific noise contamination associated with the imaging process of capture [1]. SAR imaging is the result of evaluating the backscatter of a coherent electromagnetic source, and since the vast majority of surfaces, synthetic or natural, are extremely rough on the scale of its wavelength, several different backscatters partially reflect the same electromagnetic pulse within the range of the same pixel. The superimposition of their amplitudes and phases results in strong random amplitude fluctuations among cells of the same surface, which is known as speckle noise. This noise contamination defies most of the widespread segmentation algorithms based on local differential operators, since these operators indeed amplify the speckle, turning traditional segmentation methods ineffective.

A common approach is to remove the speckle by statistical methods. The classic approach in the literature is the use of the local density function in the categorization. Usually, each image region has unique textural properties, called texture signatures, that allows to detect regions more easily. Unfortunately, most texture signatures found on literature are not invariant to geometric or luminance transformations [2], [3].

For this reason, several SAR image segmentation approaches has been proposed. Morphological region based image analysis, for instance, is based on a texture edge detector which is used in combination with a watershed transform. This typically produces an over-segmented image, in which the regions are merged using the information generated by the texture edge detector. However, the final segmented image is still over-segmented, and further processing, usually supervised, is required [4]. Another proposals use an unsupervised segmentation algorithm for high-resolution monopolarized SAR images, where texture analysis is performed using reduced contourlet transform in sub-bands, applying after a mean shift clustering. This process computes the number of texture regions and the center of the label class. Then, the pixels are grouped by their distance to the class center pixel [5]. However, the use of the mean shift algorithm requires the user to find adequate bandwidth parameters, and thus different values result in different segmentations.

Recently, the use of multifractal properties gained widespread application in the analysis of complex phenomena, including texture analysis. In [6] the authors apply a multifractal-based technique combined with mean-shift to
highlight boundaries between regions in a synthetic SAR image. The regions are well recognized but the edges are too wide and fuzzy. In [7], another method proposed combining multifractal spectrum (MFS) analysis and the use of iterated function systems. To highlight the edges of various textures, the K-Means algorithm is applied to the MFS at each point, showing adequate results on real SAR imagery at a very high computational cost. In [8], the authors apply Bayesian classification on the MFS of SAR imagery, taking advantage of the theoretical links between multifractal theory and large deviation distribution probabilities. This relationship enables a sound estimation of a-priori probabilities of different classes. In [9], multifractal analysis is performed over different SAR images of the same oil spill in different moments, to detect how the MFS of the oil spill changes over time.

In addition to MFS, other machine learning techniques are also widely used in SAR image segmentation. In [10] a partition clustering method is applied, which is regarded as a combinatorial optimisation problem. The authors apply the watershed transform to get initial small regions, and then an optimal clustering centroid is found using a quantum-inspired evolutionary algorithm, with which the final segmentation results are found. Another segmentation method is based on artificial immune systems (AIS), which consists of several steps. First, a joint filter is designed, which combines maximum likelihood estimator and partial non local means filter. Afterwards, the cluster centroids are found using a search algorithm with variable length of chromosomes. The last step is a multi-objective clustering paradigm in AIS and kernel mapping for segmentation [11]. The boundaries of the detected regions are very well formed and the segmented image is very clean (no oversegmented image) but instead the regions appear oversimplified and some important features are not detected. An adapted version of simple linear iterative clustering (SLIC) was proposed for SAR images. The distance measure of SLIC is modified using a similarity ratio. This ratio is used to label the pixels within the search areas, as a previous step of the construction of superpixels. Afterwards, this similarity metric is used in a clustering phase, which clusters the superpixels into wider segments [12]. It shows an improvement over SLIC but it is not better than the other methods with which it is compared. Another segmentation method is based on K-Means, median filter and Otsu thresholding. This method gives good results only when the regions have very different textures [13]. K-Means is also used as a filter to obtain a simplified image and then a morphological processing is done to remove small meaningless clusters [14]. It works well at low/medium scale of representation, but at higher scales it loses some details due to the filtering preprocessing initially applied.

In this paper we propose a segmentation method for SAR imagery based on local MFS features combined with K-Means. Instead of using the full MFS as a pixel feature vector for clustering, we characterize it with only four scalar values, namely center, width, height, and symmetry, which provides robust and precise texture characterization. The method is tested with synthetic SAR imagery generated using the well known $\mathcal{G}_\alpha$ distribution, with much better segmentation results as compared with traditional methods. Finally, we applied the method over part of a real SAR image of which we have assisted segmentation provided by an expert geographer, also with remarkable results.

II. MATERIALS AND METHODS

In this work we focus on monopolarized SAR imagery since images in this format are the most widespread SAR images, typically in higher spatial resolutions and with archival imagery available from earliest dates. Our primary interest is on IW-VV Sentinel-1 imagery (interferometric wide with vertical-vertical polarization). The purpose of the segmentation is to identify regions that correspond to geographical zones with specific soil desertification properties and to relate these regions in SAR imagery with optical imagery thus providing sensor information fusion. The area of interest is located in the El Morro region, Chubut province, Argentina (see Fig. 1). Low mountain landscapes, exhumed and covered peneplains, and volcanic plains are predominant in this region. The un-
derlying structure is defined by large basement blocks with inclined grabens, affected by the Andean orogeny (for more details see [15]). These characteristics make difficult a proper image segmentation and classification with optical imagery because spectral classes are confused and a good inter-class separability cannot be achieved.

With monopolarized SAR imagery, traditional segmentation and classification methods perform poorly, as can be seen in Fig. 2. Fig. 2(a) shows part of a IW-VV Sentinel-1 image covering the geographic area marked in Fig. 1. Fig. 2(b) shows the ground truth manually segmented by an expert geographer. In Fig. 2(c), (d) and (e) the results of applying respectively watershed-based segmentation, K-Means-based segmentation, and K-Means over the watershed-based segmentation are shown. It can be seen that these methods generate a significant over-segmentation, with no correlation with the actual supervised segmentation provided by the expert.

A. Synthetic Images

In order to model SAR images, the $G^0_x$ distribution [16] is a good option because it is able to characterize all kind of textures that may possibly arise. $G^0_x$ is governed by three parameters: $\alpha$, related to texture of the region, $\gamma$, related to the brightness and $L$, the number of looks. Two synthetic images are generated to test our method. The first one consists of a foreground circle with $G^0_x$ distribution parameters different to the background. The second one is a lattice of sixteen regions, where one of the $G^0_x$ distribution parameters vary from left to right and other from top to bottom. In most cases, coarseness or smoothness of SAR images, caused by the variability or uniformity of image tone, have textural properties with valuable information for analysis.

B. Multifractal Spectrum

As stated in [17] the MFS is a vector of fractal dimensions of some image measure $\mu$ (for example, the luminance). A remarkable property of MFS is its invariance under illumination and affine geometric transformations. This invariance makes MFS a robust textural feature as compared with texture nation and affine geometric transformations. This invariance is remarkable property of MFS is its invariance under illumi-

\[ f(\epsilon) = \{ \text{dim}(E_\epsilon) : \epsilon \in \mathbb{R} \} \]

where $N(r, E_\epsilon)$ is the smallest number of sets of diameter less than $r$ that cover $E_\epsilon$. Thus, a point categorization is obtained of the image with a MFS denoted as

The MFS $f(\epsilon)$ defined by (4) is the natural extension of fractal dimension.

C. A MFS Characterization

A typical MFS has the following properties (as shown in Fig. 3):

- **Width** The width of the curve is calculated as maximum epsilon with $f(\epsilon) > 0$ ($\epsilon_{\text{max}}$) minus the minimum epsilon with $f(\epsilon) > 0$ ($\epsilon_{\text{min}}$). The width of the MFS indicates the range of local fractal dimensions in the image.

- **High** The maximum value ($f(\epsilon)_{\text{max}}$) of the curve.

- **Center** The $\epsilon$ that results in $f(\epsilon)_{\text{max}}$ is called $\epsilon_{\text{center}}$.

- **Symmetry** The symmetry indicates the skew of the MFS. It is calculated as $(\epsilon_{\text{max}} - \epsilon_{\text{center}})/(\epsilon_{\text{center}} - \epsilon_{\text{min}})$. Values above one arise when the the MFS is skewed towards the left (as shown in the figure).

D. Local MFS Computation

The experiments in this paper are based on a direct computation of the MFS at each pixel. The density $f(x)$ is found at a given pixel $x$ by linear fitting of a scatter plot of $\log\mu(B(x, r))$ against $\log r$. For each $r$ value, $\mu(B(x, r))$ can be computed taking into account one of several pixel features found within the ball of decreasing radius $r$. For instance the maximum luminance, the luminance range, the accumulated luminance, and many other. In the experiments presented here, $\mu(B(x, r))$ evaluates the accumulated luminance by means of a prior averaging filtering with a given window size. All these features are non-decreasing monotonic functions of $r$, and the slope of the linear fit in log-log space provides a good estimation of $f(x)$.

In images, the possible values for $f(x)$ are within 0 (isolated pixels) and 2 (dense regions), but these extreme values are infrequent in SAR images. $f(x)$ values are discretized in a
the final \( f(\epsilon) \) MFS.

E. Image Segmentation

The image segmentation algorithm follows these steps, using the software described in Table I:

1) A square window of a given size slides over the image. The default size is \( 32 \times 32 \) but the user can modify this value (smaller windows yield faster but less accurate results).

2) For each window position, the MFS is calculated. The amount of MFS coefficients (i.e., the amount of \( \epsilon(x) \) bins) and the pixel feature (in our case, cumulative luminance over a given window size) are set by default but can be redefined by the user.

3) For each MFS, the four features of the spectra (maximum, center, width, and symmetry, see section II-C) are computed. For illustration purposes we represent these features in four separate images.

4) The four features are squared and added together, generating the enhanced texture image.

5) The enhanced texture image is clustered using the K-Means algorithm (MFS cluster).

6) If only one cluster is found, then the averaging filter window size is lowered, and the whole process is repeated from step 2.

7) If required, a majority rule filter is applied over the MFS cluster image to merge together small clusters and to remove small artifacts (see below).

The rationale of this algorithm is to have only few texture descriptors that can be clustered together without oversegmentation, and that these features are unlikely to be similar when the speckle texture is different. If the spectrum of each region has at least one of these features with a different value from other regions, then the regions can be segmented apart. However, the MFS may depend on the pixel feature used to compute the Hölder exponent (maximum luminance, luminance range, accumulated luminance, etc.). If all the values of the features of the spectrum curves of all the regions are the same, the regions cannot be separated with the current pixel averaging parameter of the MFS algorithm, and a smaller value of the pixel averaging parameter must be used, or an increase in the spectrum size should be explored.

F. K-Means clustering and majority filter

The K-Means algorithm clusters a given data set in a given number \( K \) of clusters [18]. K-Means is extensively used in image processing for segmentation of color image data or as a multilevel thresholding method. For instance in [19], the performance of K-Means as a multilevel grayscale thresholding is compared with the Otsu multilevel method. Despite being equivalent in this sense, K-Means considered to be more efficient and fast. The grayscale K-Means multilevel thresholding starts with \( K \) random gray levels (cluster centroid initialization), within the \( n \)-bit grayscale range \( 0 \leq K \leq 2^{(n-1)} \). Each image pixel is then assigned to the nearest centroid, by minimum distance (assignment stage). New values for the cluster centroids are computed as the average of each cluster after the assignment (update stage). Assignment and update stages are iterated until pixels no longer change their assigned cluster. This procedure is guaranteed to converge to the same stable result regardless of the initial assignment. Finally, each class is characterized by its corresponding centroid. Majority filter may be required in cases where this procedure generates small intertwined clusters. This filter processes over a sliding window of a given size (usually twice the size of the MFS window) assigning to each pixel the class of the most frequent class.

III. Results

In Fig. 4(a), a synthetic image composed by a circular foreground against a flat background is shown. Each region is generated with different values of \( \alpha \) and \( \gamma \) parameters of \( G^\alpha_I \) distribution, the foreground has \( \alpha = -2 \) and \( \gamma = 1 \), the background has \( \alpha = -3 \) and \( \gamma = 2 \), and \( L = 4 \) in the whole image. The MFS was computed at eleven \( \epsilon(x) \) values, with an averaging window size of six and a sliding window of \( 32 \times 32 \). It can be observed in Figs. 4(b)–(f) that the MFS features discriminate correctly foreground from background, especially the center of the MSF (Fig. 4(c)), since there appears to be a strong relationship between the \( \gamma \) parameter of the \( G^\alpha_I \) distribution and the monofractal dimension of the synthetic speckle.

In Fig. 5(a), a synthetic image with a grid of 16 regions is shown. Each region is generated with different values of the \( \alpha \) and \( \gamma \) parameters of \( G^\alpha_I \) distribution. By rows, the \( \alpha \) value varies from \(-4\) to \(-7\), and by columns the \( \gamma \) parameter varies from \( 4 \) to \( 7 \) (see Table II). The MFS was computed at \( 50 \) \( \epsilon(x) \) values, with an averaging window size of three and a sliding window of \( 32 \times 32 \). It can be observed that only the MFS maximum exhibits meaningful results (see Fig. 5(c)). The changes in \( \gamma \) parameter have a strong influence in the monofractal dimension of the generated speckle. However, this is barely noticeable with changes in the \( \alpha \) value. After K-Means is applied over the enhanced texture with sixteen clusters, the regions start to be noticeable (each cluster is

### Table I: Libraries and Applications Used in the Research

<table>
<thead>
<tr>
<th>Library/Tool</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imfractal</td>
<td>A library to compute (multi)fractal dimensions of images written in Python.</td>
</tr>
<tr>
<td>GDAL</td>
<td>Read and write of the satelital images.</td>
</tr>
<tr>
<td>ArrayFire</td>
<td>General-purpose library that simplifies the process of developing software that targets parallel and massively-parallel architectures including CPUs, GPUs, and other hardware acceleration devices.</td>
</tr>
<tr>
<td>SAGA GIS</td>
<td>GIS platform for scientific analysis and modeling.</td>
</tr>
<tr>
<td>ESA SNAP</td>
<td>Application that supports viewing, editing, and analysis of geospatial data (specialized in Sentinel data).</td>
</tr>
</tbody>
</table>
Fig. 4. MFS processing of a synthetic image of a circle, where the background has $\alpha = -2$ and $\gamma = 1$ and the foreground has $\alpha = -3$ and $\gamma = 2$ and four looks. The processing was made with a MFS of eleven values and an averaging of six pixels.

represented with a different color) (see Fig. 5(e)). Finally, a majority filter is applied to the K-Means image, with a kernel size three times bigger than sliding window to group small clusters together. In order to measure the performance of the method, we compute the Jaccard index for each segmented image, given by:

$$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}$$ (5)

where $X$ is the ideal set of pixels with a given class and $Y$ is the segmentation result set of pixels with the same class. It is a measure of similarity for the two sets of data, with a range from 0 to 1. The higher the value, more similar the two populations are. Table III shows the results. It can be seen that the classes in the expected regions (Fig. 5(b)) and the segmented image (Fig. 5(f)) have a high correlation, but the detection loses its accuracy as the $\alpha$ value increases negatively. The results appear to be much more satisfactory than the region segmentation provided by the current state of the art (see Fig. 5(e)).

In Fig. 6 we show the results of processing the image of Fig. 2(a) with the proposed workflow, using 50 $\epsilon(x)$ values and an averaging filter of width two. It can be noticed that, the centers, maxima, width and asymmetries are more variate than with the synthetic images (see Figs. 6(a)–(d)). The enhanced texture is clustered with $K = 2$, and then a majority filter of width 6 is applied. The results (Fig. 6(g)) is much more consistent with the ground truth provided by an expert supervised segmentation (Fig. 2(b)).

| TABLE II |
| PARAMETERS USED IN EACH REGION OF THE GRID $4 \times 4$ (SEE FIG. 5). |
| FORMATTED USED $(\alpha, \gamma)$. |
| 
| (-4,4) | (-5,4) | (-6,4) | (-7,4) |
| (-4,5) | (-5,5) | (-6,5) | (-7,5) |
| (-4,6) | (-5,6) | (-6,6) | (-7,6) |
| (-4,7) | (-5,7) | (-6,7) | (-7,7) |

| TABLE III |
| JACCARD INDEXES OF THE THREE SEGMENTED IMAGES IN FIGURES 4(G), 5(F) AND 6(G). EACH INDEX IN GRID COLUMN CORRESPONDS TO EACH REGION IN 5(B). |
| 
| Fig. 4(g) | Fig. 5(f) | Fig. 6(g) |
| 0.9181 | 0.5590 | 0.6745 | 0.7296 |
| 0.9202 | 0.4545 | 0.5469 | 0.54198 |
| 0.2010 | 0 | 0 | 0 |

0.8034
IV. DISCUSSION, CONCLUSIONS, AND FURTHER WORK

We presented a SAR image segmentation method based on local features of the multifractal spectrum, together with clusterization. The performance of the method was analyzed with two synthetic images with controlled speckle properties, and with an actual SAR image over a geographic region within which expert land cover recognition was available. As can be seen in the results, the method is very sensitive to $\gamma$ changes. Besides, experimental results show that raising the spectrum size (the amount of $\epsilon(x)$ evaluations per pixel) produces an increase in the algorithm sensitivity to distinguish apart similar $\alpha$ parameter values in the $G^\alpha$ distribution, and that this parameter is strongly related to the local monofractal dimension. The averaging filter and the sliding window sizes both increase the spectrum accuracy. On the other hand, these factors also increase significantly the computation times. For this reason, we are currently devising quality measures that can be evaluated without supervision, in a way such that an exploration of these parameter values can be performed automatically to achieve the best results with the minimum computation.

ACKNOWLEDGMENTS

We wish to thank Dr. Hector del Valle for his assistance with ground truth validation and for comments that greatly improved the quality of the manuscript. This work was partially funded by a AO-SAOCOM grant provided by the National Aerospace Activities Agency of Argentina (CONAE).

REFERENCES