Unsupervised SAR Image Segmentation using Recursive Partitioning

Vidya Venkatachalam[a], Robert D. Nowak[a], Richard G. Baraniuk[a] and Mário. A. T. Figueiredo[b]

[a]Computational Mathematics Laboratory
Department of Electrical & Computer Engineering
Rice University
Houston, TX, USA

[b]Instituto de Telecomunicacões, and
Instituto Superior Técnico
1049-001 Lisboa, Portugal

ABSTRACT

We present a new approach to SAR image segmentation based on a Poisson approximation to the SAR amplitude image. It has been established that SAR amplitude images are well approximated using Rayleigh distributions. We show that, with suitable modifications, we can model piecewise homogeneous regions (such as tanks, roads, scrub, etc.) within the SAR amplitude image using a Poisson model that bears a known relation to the underlying Rayleigh distribution. We use the Poisson model to generate an efficient tree-based segmentation algorithm guided by the minimum description length (MDL) criteria. We present a simple fixed tree approach, and a more flexible adaptive recursive partitioning scheme. The segmentation is unsupervised, requiring no prior training, and very simple, efficient, and effective for identifying possible regions of interest (targets). We present simulation results on MSTAR clutter data to demonstrate the performance obtained with this parsing technique.

Keywords: segmentation, multiscale, wavelets, MDL, SAR, ATR, MSTAR

1. INTRODUCTION

Segmentation of Synthetic Aperture Radar (SAR) images addresses the problem of identifying different regions of homogeneous "textural" properties within the image. Several texture segmentation schemes have been designed for this purpose; they naturally fall into two broad categories:

- Supervised — requires training prior to segmentation.
- Unsupervised — direct segmentation with no training.

In supervised schemes, we follow a two-step procedure: training followed by segmentation. In these techniques, we are required to have several sets of reliable training data for every class (texture), from which we extract a set of parameters that we label the characteristic signature for the underlying class. The signature could either be a deterministic set of parameters or a probabilistic model of the training data. Given data belonging to an unknown class, we then try to determine the class with which the data has the closest match. The key issue in these schemes is the definition of what constitutes the signature for a class. Given an intelligent and informed technique to determine the signatures, supervised approaches work well for segmentation problems.

However, there is a drawback in the supervised approach: oftentimes, sufficient data may not be available to form reliable signatures, or the data may not be accurate or completely representative. Also, in some cases, it may be

Email: {vidyav,nowak,richb}@rice.edu, Web: www.dsp.rice.edu.
This work was supported by DARPA/AFOSR grant F49620–97–1–0513, NSF grants MIP–9457438 and CCR–9973188, ONR grant N00014–99–1–0813, National Science Foundation grant MIP–9701692, Army Research Office DAAD19-99-1-0349, and the Texas Instruments Leadership University Program.

impossible to obtain training data. In these instances, we are faced with the problem of having to segment an image into textures or homogeneous regions, the nature and number of which we have very little prior knowledge of. In other words, we are faced with the problem of having to perform a "blind" segmentation.

This situation leads us to examine the second category of segmentation schemes – unsupervised techniques. In these techniques, we attempt to cluster unknown image data into groups using some predetermined criterion. The criterion used in a particular problem depends on any prior knowledge we have about the general nature of the problem. However, it is completely independent of the specific textures or classes we are trying to determine from the given data. It is important to note that unsupervised techniques are not necessarily better than supervised techniques, especially when we have very good training data. It is only that certain situations may warrant that we use unsupervised schemes to effect a reliable (in most cases, coarse) segmentation, and it is to these situations that our proposed method is targeted.

This paper addresses the following two problems:

- The characterization of SAR amplitude images using suitable Poisson field approximations.
- The use of unsupervised progressive parsing of Poisson fields using the Minimum Description Length (MDL) criterion.

In the sections to follow, we discuss both the above problems and demonstrate the effectiveness of the overall technique on MSTAR clutter data.

2. POISSON MODELS AND SAR IMAGES

We first address the issue of why we choose to model the SAR amplitude image using Poisson fields. Our main goal is to *parse*, or *segment*, the observation space into regions of (roughly) homogeneous intensity in an unsupervised manner. This problem has been studied and successfully solved by Nowak and Figueiredo[1] using Rissanen's MDL principle.[2] This technique gives good results when the data are well modeled by a spatial Poisson distribution. A significant advantage of the approach is that, since the Poisson data are integer-valued, we are able to derive MDL criteria without recourse to asymptotic approximations, which is not the case when the data is modeled by, say, Gaussian distributions. The success of the technique led us to explore the possibility of applying the same to SAR data by modeling it using Poisson field approximations.

We begin our study by examining the nature of SAR images in order to see how their properties can be exploited to determine a suitable Poisson approximation. It is a well-known fact that SAR images are contaminated with a particular kind of noise, commonly referred to as *speckle*. This occurs due to the interference of waves reflected from many elementary scatterers within a resolution cell on the imaged surface. This effect causes a pixel to pixel variation in intensity, which manifests itself in the granular pattern or speckle typically observed in SAR images.

Suppose we have a random SAR data set (typically complex) represented by X^*. Then, it has been observed that the pixels $y =\mid x \mid$ in the speckled SAR amplitude image $Y =\mid X \mid$ are Rayleigh distributed[3] with pdf

$$p(y) = \frac{2y}{\alpha^2} e^{-\frac{y^2}{\alpha^2}}, \;\; y \geq 0$$

where $\alpha^2 = E[Y^2]$, is the mean of the intensity image Y^2. Note that this value is *pixel dependent* since it is characteristic of the resolution cell associated with every pixel. However, we can reasonably argue that the mean intensity associated with every pixel in a homogeneous region is approximately uniform, and hence, such a region can be characterized by a single Rayleigh distribution with parameter α. The mean and variance associated with this Rayleigh distribution are given by[4] $\mu = \frac{\sqrt{\pi}\alpha}{2}$, and $\sigma^2 = \frac{(4-\pi)\alpha^2}{4}$, respectively.

The intensity image $Z = Y^2$, for a single-look SAR image, is exponentially distributed, with pdf

$$p(z) = \frac{1}{\alpha^2} e^{-\frac{z}{\alpha^2}}$$

*Upper case letters denote realizations of random variables, and lower case the actual values assumed by them.

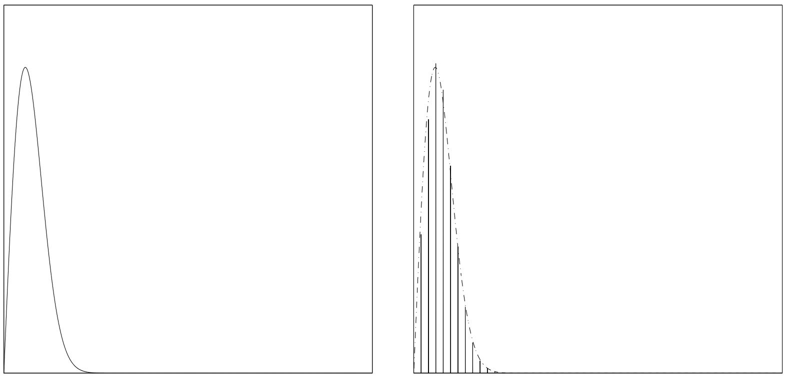
For our analysis, we could work with either the amplitude or the intensity image. Since our goal is segmentation, we choose to work with the amplitude image since it has been shown that[5] it is very difficult to segment an intensity SAR image based on the exponential distribution.

Consider a (homogeneous) region Y_R of the amplitude image Y, having a Rayleigh distribution parameterized by α, with mean μ and variance σ^2. In order to approximate Y_R using a Poisson model with parameter $\lambda^\dagger$, we match the moments of the Rayleigh and the Poisson distributions. Specifically, we match the first (mean) and second (variance) moments of the two distributions. Then, we must have $\lambda = \mu = \sigma^2$; that is,

$$\begin{aligned} \mu &= \sigma^2 \\ \Rightarrow \frac{\sqrt{\pi}\alpha}{2} &= \frac{(4-\pi)\alpha^2}{4} \\ \Rightarrow \alpha^2 &= \frac{4\pi}{(4-\pi)^2} \\ \Rightarrow \lambda &= \frac{\pi}{(4-\pi)}. \end{aligned}$$

The above analysis leads us to conclude that we need to normalize the region Y_R in the amplitude image Y such that the associated $\alpha^2 = E[Y_R^2] = \frac{4\pi}{(4-\pi)^2}$. Suppose $\alpha_0 = E[Y_R^2]$. Let $c = \frac{4\pi}{(4-\pi)^2\alpha_0}$. Then, $E[cY_R^2] = \frac{4\pi}{(4-\pi)^2}$, as desired. Thus, we need to scale the region Y_R by the factor $\sqrt{c}$ to achieve the required Poisson approximation. Fig. 1(a) shows the Rayleigh distribution with parameter $\alpha = \frac{2\sqrt{\pi}}{(4-\pi)}$, and Fig. 1(b) the Poisson distribution with parameter $\lambda = \frac{\pi}{(4-\pi)}$. Observe that the Poisson provides a good approximation to the Rayleigh distribution of Fig. 1(a) (also shown superimposed on the Poisson distribution in Fig. 1(b)).

Clearly, the above process results in a spatially varying normalization over the image Y, with areas of approximately uniform mean intensity values (α^2's) having the same normalization factor. This process can be rather cumbersome in practice, and not always practical. However, we have observed, from experimental data, that fitting a single Rayleigh density to the amplitude image Y, and adopting one normalization factor for the entire image based on this density, is much more efficient and produces minimal deterioration in the final segmentation performance. With this normalization, we can treat the scaled SAR amplitude image data as though it were acquired via a Poisson process and apply Rissanen's MDL technique to effect parsing.



(a) (b)

Figure 1. *(a) Rayleigh density with parameter* $\frac{2\sqrt{\pi}}{(4-\pi)}$, *(b) Poisson density with parameter* $\frac{\pi}{(4-\pi)}$ *with the Rayleigh density from Fig. 1(a) superimposed.*

†The Poisson random variable is parameterized by its mean and variance, both of which equal each other and assume a positive real-valued number, in this case, λ.

3. MDL ANALYSIS

Suppose we have observed data that is a realization from a spatial Poisson point process whose underlying intensity function can be well approximated by piecewise constant functions. The MDL principle essentially addresses the following problem:

> Given a set of observed data and a set of possible models, determine the model that best explains the observed data.

The first step here is to formalize the concept of the 'best' model in any given scenario. For this, we base our formulation on that of Rissanen[2] and his MDL criterion. The basic principle of the technique is provided here (for more details, refer to Nowak and Figueiredo[1]). Suppose we wish to transmit some observed data x to a hypothetical receiver. Given a probabilistic model for the data, say $p(x|\theta)$, the Shannon-optimal code length is $-\log p(x|\theta)$. In order to decode the transmitted information, the receiver also needs to know the model parameters θ. If θ is unknown *a priori*, this would mean that we need to estimate it, code it, and then transmit. Now, consider a set of K competing model classes $\{p_i(x|\theta_i)\}_{i=1}^{K}$. In each class i, the "best" model is the one that gives the minimum code length, and we express this as:

$$\widehat{\theta}_i = \arg\min_{\theta_i} \{-\log p_i(x|\theta_i)\} = \arg\max_{\theta_i} p_i(x|\theta_i).$$

Note that this is simply the *maximum likelihood* (ML) estimate within model class i. But if the class is unknown *a priori*, the "best" overall model is the one that leads to the *minimum description length*: the sum of $-\log p_i(x|\theta_i)$ with the length of the code for θ_i itself. The fundamental aspect of MDL is that unlike the ML criterion, it performs model selection by penalizing more complex model classes (those requiring longer parameter code lengths).

The critical issue in applying MDL lies in the encoding of the parameter θ_i. Appropriate parameter code lengths are usually based on asymptotic approximations; for example, the well known $(1/2)\log N$, where N is the amount of data, is an asymptotic code length.[2] Nowak and Figueiredo[1] devised a scheme whereby they avoid asymptotic approximations and obtain exact code lengths using Poisson approximations. Since our unsupervised parsing technique is based on this work, we first provide an overview of the scheme, and then explain how it applies to SAR data.

4. MDL FOR POISSON DATA

The MDL technique is general enough to be applied to a wide class of data. However, as mentioned earlier, to avoid the problem of asymptotic approximations, we only consider the special case when our data can be regarded as (at least approximately) Poisson. In the analysis that follows, we use the 2-D Haar multiscale wavelet analysis as the underlying framework.

Two approaches are explained here:

- Quadtree-based Segmentation, wherein we progressively parse the Poisson image data into dyadic partitions.
- Adaptive Recursive Segmentation, wherein we progressively partition the image data into rectangular tesselations of the plane.

Both the approaches use the MDL model class selection criteria as the basic building block.

4.1. Quadtree-based Segmentation

Suppose we have Poisson image data $\{x_{k,l}\}$, $k, l = 0, \ldots, 2^J - 1$ (note that the data size is restricted to be a power of 2 in both k and l), and define $x_{J,k,l} \equiv x_{k,l}$ and for $j = J-1, \ldots, 0$

$$\begin{aligned} x_{j,k,l} &= x_{j+1,2k,2l} + x_{j+1,2k+1,2l} \\ &\quad + x_{j+1,2k,2l+1} + x_{j+1,2k+1,2l+1}. \end{aligned} \tag{1}$$

Here, $j = J$ and $j = 0$ are the highest (finest) and lowest (coarsest) resolutions (scales), respectively. Adopting a predictive coding approach, operating in scale, from coarse to fine, we start by transmitting the total count $x_{0,0,0}$ which we code using Elias' technique for arbitrarily large integers.[2] We then progressively transmit the data (in a coarse-to-fine fashion), wherein, at each stage, we take advantage of the coarser scale data already sent. Thus, at scale $j+1$ we only transmit the triple $x_{j+1,2k,2l}$, $x_{j+1,2k+1,2l}$, $x_{j+1,2k,2l+1}$, since the receiver already has the corresponding total $x_{j,k,l}$, and consequently, can infer $x_{j+1,2k+1,2l+1}$ from Eq. 1.

The quantity of interest to us is the conditional probability $p(x_{j+1,2k,2l}, x_{j+1,2k+1,2l}, x_{j+1,2k,2l+1}|x_{j,k,l})$ which in this case, follows a multinomial distribution,[6,7] with parameters $\rho_{j+1,2k,2l} = \frac{\lambda_{j+1,2k,2l}}{\lambda_{j,k,l}}$, $\rho_{j+1,2k+1,2l} = \frac{\lambda_{j+1,2k+1,2l}}{\lambda_{j,k,l}}$, and $\rho_{j+1,2k,2l+1} = \frac{\lambda_{j+1,2k,2l+1}}{\lambda_{j,k,l}}$, where the $\{\lambda_{j,k,l}\}$ are the intensities underlying the Poisson counts.

Consider two alternative model classes:

Model Class 0: This model assumes a homogeneous Poisson process, and hence, there is no split, requiring no encoding. The description length L_0 is given by the negative of the log of the multinomial probability $p(x_{j+1,2k,2l}, x_{j+1,2k+1,2l}, x_{j+1,2k,2l+1}|x_{j,k,l})$.

Model Class 1: Here, we assume non-homogeneous regions, and split into 4 subregions. The parameters are coded and transmitted progressively with $\log_2(x_{j,k,l}+1)$, $\log_2(x_{j,k,l} - x_{j+1,2k,2l}+1)$, and $\log_2(x_{j,k,l} - x_{j+1,2k,2l} - x_{j+1,2k+1,2l}+1)$ bits, respectively. Their sum gives us the description length L_1.

We transmit the data according to Model Class 0 if $L_0 < L_1$; else, we use Model Class 1 (note that the model class we choose when $L_0 = L_1$ is irrelevant). The optimal MDL parsing of the data is obtained according to this same criterion, applied at each scale and location; i.e., if $L_0 < L_1$, we transmit the parameters of Model Class 0, else, we transmit the parameters of Model Class 1. The underlying piecewise constant intensity (scale J) function is then reconstructed/estimated in a successive refinement scheme starting from the initial $\lambda_{0,0,0} = x_{0,0,0}$, and building on this using the sequence of splitting proportion estimates transmitted, these being the parameters (ρ's) of the selected model class. This MDL rule can be shown to be equivalent to a Bayesian selection criterion. Complete details can be obtained from Nowak and Figueiredo.[1]

4.2. Adaptive Recursive Segmentation

The quadtree-based segmentation approach outlined above is limited by the fact that the parsing is restricted to a fixed quadtree. In general, the best locations for parsing may not coincide with the dyadic partition enforced by this multiscale analysis. The scheme presented here considers an adaptive recursive approach that allows for splits at arbitrary locations.

Again, we begin with a Poisson data set $\{x_{k,l}\}$, $k, l = 0, \ldots, N-1$, but in this case, N need not be a power of 2. We now have greater freedom in how we split the data. However, for purposes of computational tractability, we restrict the splitting to rectangular tesselations of the plane, giving rise to the following 3 possible model classes.

Model Class 0: Here, we assume homogeneous rectangular regions with constant Poisson intensity. Hence, we have no splitting.

Model Class 1: Here, we assume non-homogeneous regions within the rectangle, and split it into four sub-rectangles defined by a common vertex, each characterized by a different constant intensity. The advantage over the quadtree approach where the split locations were fixed, is that we can now look for the best possible such split.

Model Class 2: Again, we assume non-homogeneous regions, but the rectangle is split either horizontally or vertically into two sub-rectangles, each with its own intensity. As in Model Class 1, we can select the best possible location to effect the splitting.

In reality, we actually have under Model Classes 1 and 2, several sub-classes, each corresponding to a split at a specific location. However, if we assume that every split location in the sub-classes is *a priori* equiprobable, then each index requires the same code length which can then be dropped from any comparisons. The description lengths L_i's for each of these classes can be derived from the associated multinomial probabilities - for complete details, refer Nowak and Figueiredo.[1]

The MDL-based parsing process is as follows: we start by encoding the total count $s_N = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} x_{k,l}$. Then, from the full data set, we compute the description lengths for all possible splits. Unless we have L_0 (description length for Model Class 0) less than all other L_i's, we split, either using Model Class 1 or Model Class 2 (the actual splitting location is of course determined by the corresponding L_i that is minimum). This process is recursively applied to rectangular blocks in the image data (at the start, the block is the entire image), and every time one rectangular block is split (into 2 or 4 sub-rectangles), the criterion is again applied to the resulting sub-regions. At each stage, we need to transmit the s_i's (total Poisson counts of the sub-rectangles) corresponding to the splits chosen. The parsing process stops when no further splits are indicated by the MDL criterion, which implies the selection of Model Class 0 for every sub-block. The final estimate of the intensity field is piecewise flat, with the rectangular regions defined by the parsing; the corresponding intensities are the ML estimates based on the data inside each region. Thus, we have an efficient adaptive recursive partitioning scheme, which we can now apply to parse Poisson data.

5. SIMULATION RESULTS

In this section, we present the results of applying the above two parsing techniques to SAR data. Since this is an *unsupervised* approach, we do not require any prior training.

We use, as an illustrative example, the MSTAR clutter data HB06170,[8] which is a SAR image containing forest and field imagery (Fig. 2). The SAR amplitude data is Rayleigh distributed (see Fig. 3(a)), and the corresponding intensity data is exponentially distributed (see Fig. 3(b)), which suggests that we have a single-look image. We normalized the amplitude data shown in Fig. 2, to approximate a Poisson process using the approach outlined in Section 2. The resulting normalized image is shown in Fig. 4. We can now assume that this normalized image is well approximated by Poisson fields, and apply the MDL-based parsing techniques to it. Fig. 5 shows the reconstructed image obtained by applying the fixed quadtree-based parsing technique, while Fig. 6 shows the reconstructed image using the adaptive recursive parsing scheme. The data shown in Figs. 5 and 6 represent the different (homogeneous) regions, each characterized by a constant Poisson intensity determined by the fixed (Fig. 5) and adaptive (Fig. 6) tree-based parsing algorithms. The main difference between the two approaches is the flexibility in the adaptive approach which results in a finer delineation of the forest boundary, and in far fewer regions. In both cases, observe that we obtain a fairly good segmentation, especially in the homogeneous field region. We tested this approach on several SAR images with uniformly good results. This suggests that this technique yields a good and reliable coarse segmentation, which can then be fed to more sophisticated segmentation algorithms to effect better final results.

6. CONCLUSIONS

We have presented a technique that performs unsupervised parsing of SAR data using the MDL criterion, specifically adapted to Poisson processes. For this, we showed first how we can approximate a SAR amplitude image by a Poisson process by using a suitable normalization, and then explained how we can apply MDL-based parsing to the normalized image. Two parsing approaches were enunciated — a fixed quadtree-based approach, and an adaptive recursive approach. Both these are based on a progressive multiscale refining hypothesis, with the adaptive approach giving us greater flexibility in selecting the optimal partitions.

We illustrated the effectiveness of this technique in providing coarse segmentations of SAR images through an example. It is important to stress that our main goal here was not to obtain the best possible segmentation; rather, we wanted to obtain, with no prior training, a reliable coarse segmentation in a reasonably efficient manner.

REFERENCES

1. R. Nowak and M. Figueiredo, "Unsupervised progressive parsing of poisson fields using minimum description length criteria," in *IEEE Int. Conf. on Image Proc. — ICIP '99*, pp. 26–29, vol. II, (Kobe, Japan), Oct. 24-28 1999.
2. J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, World Scientific, Singapore, 1989.
3. J. Lee and I. Jurkevich, "Speckle filtering of synthetic aperture radar images: A review," *Remote Sensing Reviews* **8**, pp. 313–340, 1994.
4. E. B. Manoukian, *Modern Concepts and Theorems of Mathematical Statistics*, Springer-Verlag, New York, 1986.
5. J. Lee and I. Jurkevich, "Segmentation of sar images," *IEEE Trans. Geoscience and Remote Sensing* **27**(6), pp. 674–680, 1989.
6. K. Timmermann and R. Nowak, "Multiscale modeling and estimation of poisson processes with application to photon-limited imaging," *IEEE Trans. on Information Th.,* **45**, pp. 846–862, April 1999.

7. E. Kolaczyk, "Bayesian multi-scale models for poisson processes," *Journal of American Statistical Association* , Sep. 1999.
8. *MSTAR (Public) Clutter DBA7B02AA 2-CD set* , September 1995. Further information available at `www.mbvlab.wpafb.af.mil/public/MBVDATA`.

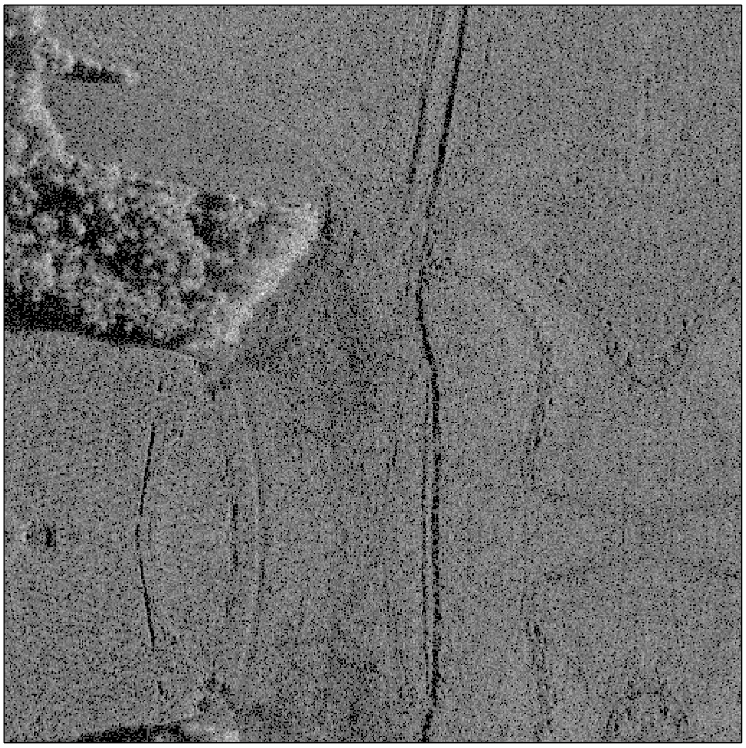
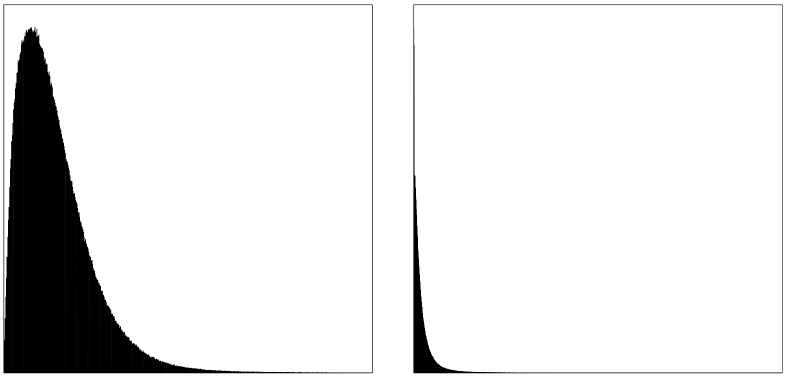


Figure 2. *Original SAR amplitude data HB06170 containing forest and field imagery.*



(a) (b)

Figure 3. *Histogram plots of the amplitude and intensity of the data in Fig. 2: (a) Approximates a Rayleigh distribution with parameter* $\alpha = 58.36$*, (b) Approximates an exponential distribution with parameter* $\alpha = 58.36$.

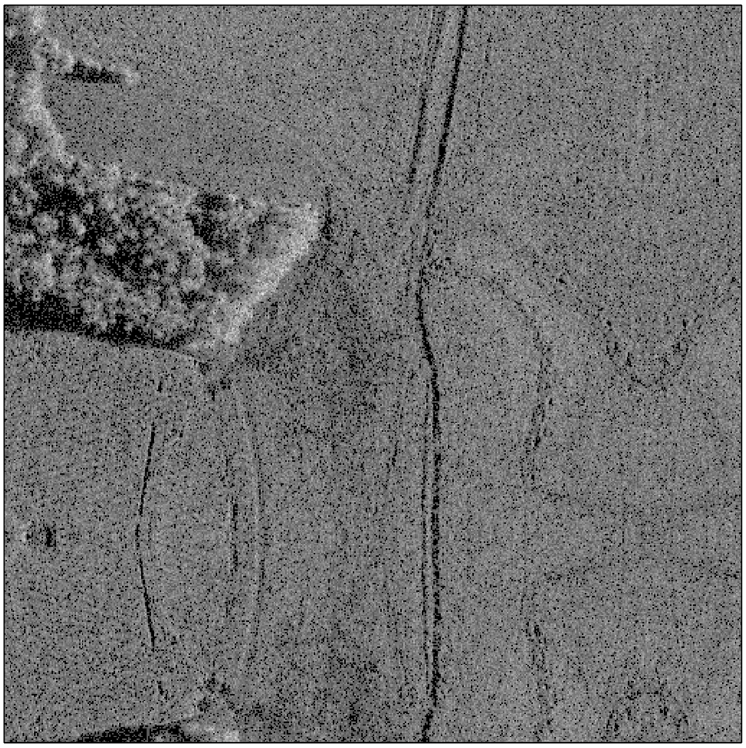


Figure 4. *The SAR data of Fig. 2, normalized to approximate a Poisson process.*

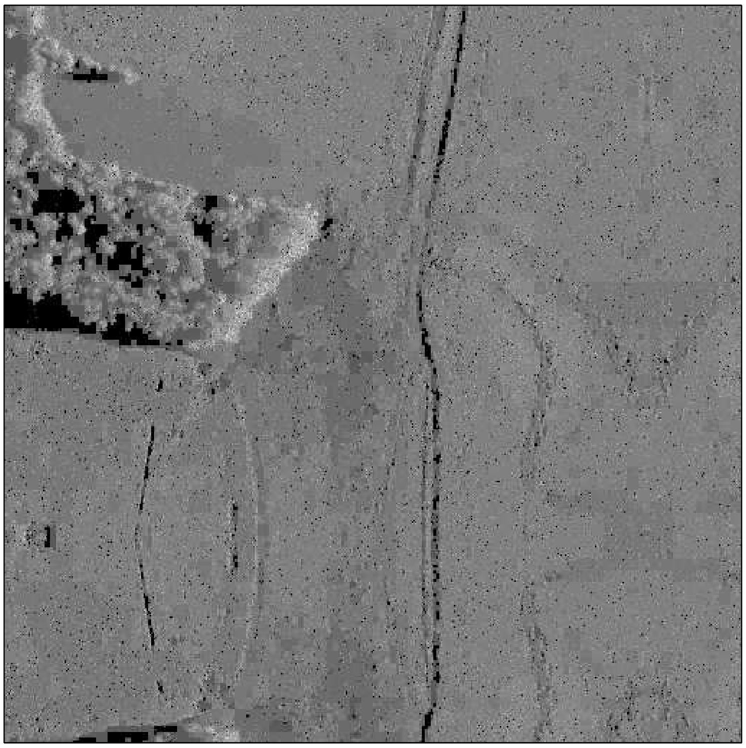


Figure 5. *Unsupervised segmentation of the data in Fig. 4 using the multiscale quadtree algorithm of Section 4.1.*

Figure 6. *Unsupervised segmentation of the data in Fig. 4 using the adaptive recursive parsing algorithm of Section 4.2.*