Multiscale Segmentation and Anomaly Enhancement of SAR Imagery

Charles H. Fosgate, Hamid Krim, *Member, IEEE,* William W. Irving, *Member, IEEE,* William C. Karl, *Member, IEEE,* and Alan S. Willsky, *Fellow, IEEE*

***Abstract*— In this paper, we present efficient multiscale approaches to the segmentation of natural clutter, specifically grass and forest, and to the enhancement of anomalies in synthetic aperture radar (SAR) imagery. The methods we propose exploit the coherent nature of SAR sensors. In particular, they take advantage of the characteristic statistical differences in imagery of different terrain types, as a function of scale, due to radar *speckle*. We employ a recently introduced class of multiscale stochastic processes that provide a powerful framework for describing random processes and fields that evolve in scale. We build models representative of each category of terrain of interest (i.e., grass and forest) and employ them in directing decisions on pixel classification, segmentation, and anomalous behavior. The scale-autoregressive nature of our models allows extremely efficient calculation of likelihoods for different terrain classifications over windows of SAR imagery. We subsequently use these likelihoods as the basis for both image pixel classification and grass–forest boundary estimation. In addition, anomaly enhancement is possible with minimal additional computation. Specifically, the residuals produced by our models in predicting SAR imagery from coarser scale images are theoretically uncorrelated. As a result, potentially anomalous pixels and regions are enhanced and pinpointed by noting regions whose residuals display a high level of correlation throughout scale. We evaluate the performance of our techniques through testing on 0.3-m SAR data gathered with Lincoln Laboratory's millimeter-wave SAR.**

I. Introduction

In RECENT years, there has been a growing interest in synthetic aperture radar (SAR) imaging for applications ranging from remote sensing to surface surveillance and automatic target recognition (ATR). For applications such as these, the classification of various categories of clutter is quite important, and their delineation (i.e., segmentation) can play a key role in the subsequent analysis for target detection, recognition, and image compression. In light of typical coverage rates (exceeding 1 km^2/s) of an airborne SAR, it is of great importance to devise efficient (preferably parallelizable) algorithms capable of meeting the daunting computational demands of the resulting data collection.

Manuscript received November 1, 1995; revised July 30, 1996. This work was supported in part by ARPA under AFOSR Grant F49620-93-1-0604 and under Air Force Contract F19628-95-C-0002, by ARO under Grant DAAL03-92-6-0115, and by AFOSR under Grant F49620-95-1-0083.

C. H. Fosgate, H. Krim, and A. S. Willsky are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02215 USA (e-mail: willsky@mit.edu).

W. W. Irving is with Alphatech Inc., Burlington, MA 01803 USA.

W. C. Karl is with the Department of Electrical Engineering, Boston University, Boston, MA 02139 USA.

Publisher Item Identifier S 1057-7149(97)00461-2.

In this paper, we apply a multiscale approach to the SAR image segmentation problem that exploits the coherent nature of SAR image formation. In particular, we build on the idea of characterizing and exploiting the scale-to-scale statistical variations in SAR imagery due to radar speckle [6], [7], [10]. A SAR image measures the coherent sum of the returns from all radar scatterers within each resolution cell (pixel). As the resolution of the imagery changes, the set of complex scatterers contributing to each resolution cell changes; equivalently, the relative contribution of each scatterer to the resolution cell changes. Since this is a coherent process, there can be constructive and destructive interference among the scatterers. Consequently, there is a statistical variation in the imagery from resolution to resolution (i.e., from scale to scale). These statistics depend on the distribution of the scattering elements in the environment. More specifically, they depend on the type of the clutter being imaged.

To fully exploit this phenomenon, we employ a recently introduced class of multiscale stochastic models [1], [2]. These models provide a powerful framework for describing random fields that evolve in scale. The framework uses a pyramidal tree structure in which each node corresponds to a pixel at a particular image location and resolution. The offspring of the node correspond to the pixels, in the same location, at the next finer scale. The statistical variability of the multiscale imagery is subsequently captured by identifying a scale-recursive stochastic model for each clutter type.

In this paper, we hypothesize that the scale-to-scale variation of distinct clutter types will differ in a statistically significant manner. If this is the case, very efficient algorithms associated with the multiscale models can be applied to calculate likelihoods for the classification of individual SAR image pixels and the subsequent segmentation of SAR imagery. We demonstrate the utility of the multiscale methodology for the segmentation of regions of trees and forest from open fields and grass. Such segmentation can be useful for ATR systems. For example, if a densely forested region is identified, performing target detection in such a region is unnecessary because high-frequency SAR is incapable of providing imagery of targets under the forest canopy. Moreover, targets of interest often attempt to conceal themselves near tree lines. Thus, accurate estimation of tree lines can be used to focus attention of ATR algorithms on areas of particular interest.

Finally, by segmenting regions of natural clutter, we gain the ability to identify pixels that are anomalous with respect to the clutter classification (i.e., segmentation). That is, we identify

pixels whose prediction error residuals differ in a statistically significant manner from those expected theoretically based on the residual distribution of the corresponding multiscale model. Such identification plays a key role in ATR where the identification of man-made targets is critical. Furthermore, since man-made objects typically exhibit very bright pixels, one would expect to observe high scale-to-scale correlation in regions where targets are present. Given that the prediction error residuals are calculated at a number of scales, it is possible to exploit the scale-to-scale correlation to further enhance and detect targets.

In the next section, we describe the multiscale framework and its application to SAR image segmentation. In Section III, we describe the segmentation of SAR imagery into different clutter classes. In Section IV we outline a multiscale approach for anomalous pixel enhancement. In Section V, we evaluate the performance of our approaches when applied to 0.3-m resolution SAR imagery collected by the Lincoln Laboratory millimeter-wave SAR [5].

II. Multiscale Models of SAR Imagery

This section describes a multiscale framework for analyzing SAR imagery. Section II-A describes the mapping of multiscale imagery onto a *quadtree* structure. Section II-B describes a class of stochastic models for describing and analyzing multiscale processes that are mapped onto quadtrees. Section II-C specifies an autoregressive model for SAR imagery. Section II-D describes the estimation of parameters within the multiscale SAR models.

A. Quadtree Interpretation of SAR Imagery

The starting point for our model development is a multiscale sequence $I_L, I_{L-1}, \cdots, I_0$ of SAR images, where I_L and I_0 correspond to the coarsest and finest resolution images, respectively. The resolution varies dyadically between images at successive scales. More precisely, we assume that the finest-scale image I_0 has a resolution of $\delta \times \delta$ and consists of an $N \times N$ array of pixels (with $N = 2^M$ for some M). Hence, each coarser resolution image I_m has $2^{-m}N \times 2^{-m}N$ pixels and resolution $2^m\delta \times 2^m\delta$. Each pixel $I_m(k,l)$ is obtained by taking the coherent sum of complex fine-scale imagery over $2^m \times 2^m$ blocks, performing log-detection (computing 20 times the log-magnitude), and correcting for zero frequency gain variations by subtracting the mean value [3]. (Note that the imagery is converted to log-magnitude because the multiscale recursive models described below have empirically proven most effective when using this representation. Direct use of the complex imagery is impractical because of the variability of the phase in the imagery. Also, the log-magnitude of the imagery provides more well-behaved residuals, in a statistical sense, than the magnitude imagery. Accordingly, each pixel in image I_m corresponds to four "child" pixels in image, I_{m-1}. This indicates that a fourth-order tree, or *quadtree*, is natural for the mapping. Furthermore, each node s on the quadtree can be thought of as having associated with it a 3-tuple (m,k,l), where m denotes scale and (k,l) denotes two-dimensional (2-D) image pixel location. That is, each node s on the tree

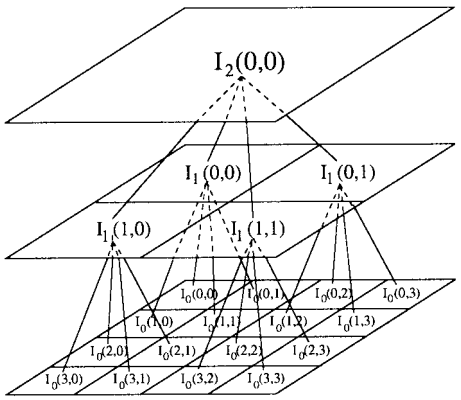


Fig. 1. Sequence of three multiresolution SAR images mapped onto a quadtree. The pixel value at scale *m* and position (k,l) is denoted by $I_{2-m}(k,l)$.

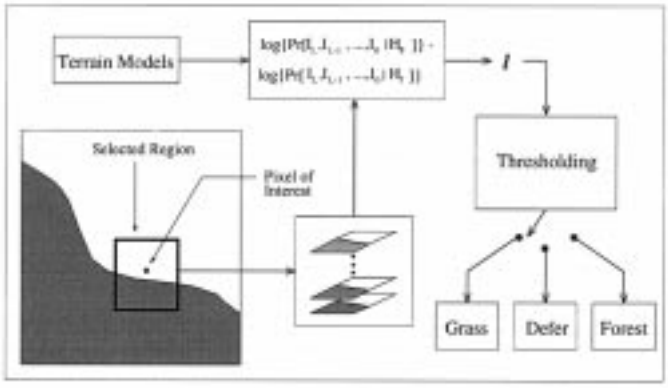


Fig. 2. Sequence of steps involved in initial pixel classification. i) Creation of multiscale sequence from the window region. ii) Evaluation of decision statistic ℓ. iii) Thresholding to determine center pixel classification.

is associated with one of the pixels $I_m(k,l)$ corresponding to pixel (k,l) of SAR image I_m. As an example, Fig. 1 illustrates a multiscale sequence of three SAR images, together with the quadtree mapping. Here the finest-scale SAR imagery is mapped to the finest level of the tree, and each coarse scale representation is mapped to successively higher levels. Furthermore, we use the notation $I(s)$ to indicate the pixel mapped to node s.

B. Multiscale Stochastic Models

In this Section, we describe a general multiscale modeling framework [1], [2] and its applicability to the SAR quadtree representation. Under this framework, a multiscale process is mapped onto nodes of a qth-order *tree*, where q depends upon how the process progresses in scale. A qth-order tree is a connected graph in which each *node*, starting at some root node, branches off to q child nodes. As described above, the appropriate representation for a multiscale SAR image sequence is $q = 4$, a quadtree. Each level of the tree (i.e., distance in nodes from the root node) can be viewed as a distinct scale representation of a random process, with the resolutions proceeding from coarse to fine as the tree is traversed from top to bottom (root node to terminal nodes).

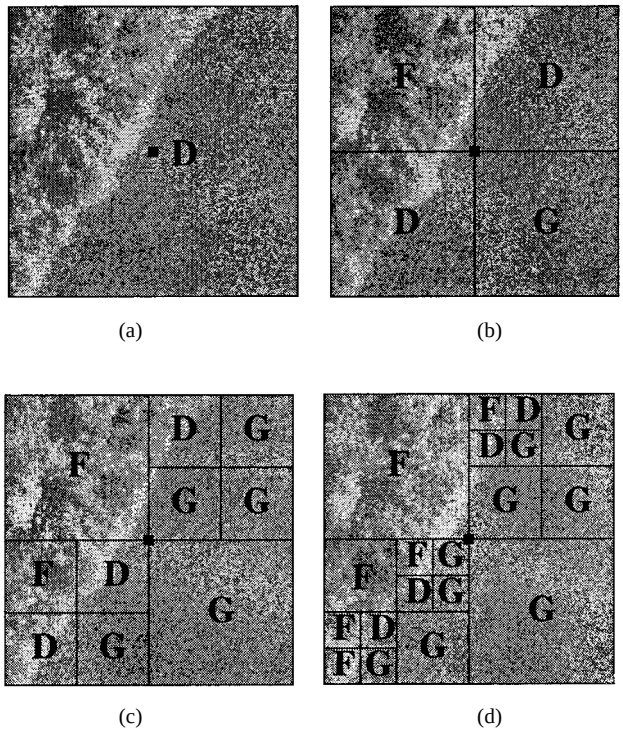


Fig. 3. Boundary pixel refinement of typical SAR window region. (a) Window region deferred due to boundary presence. (b) Region divided into quadrants, with ternary classification results marked for each subregion. (c), (d) Regions still classified as boundary further subdivided and reclassified. Classifications G, F, and D refer to grass, forest, and defer, respectively. Note that $9/16$ of the window region is classified as grass; thus, the pixel is classified as grass.

A coarse-scale shift operator, $\bar{\gamma}$ is defined to reference the *parent* of node s, just as the shift operator z allows referencing of previous states in discrete time-series. The state elements at these nodes may be modeled by the coarse-to-fine recursion

$$\mathbf{x}(\mathbf{s}) \;=\; \mathbf{A}(\mathbf{s})\mathbf{x}(\mathbf{s}\bar{\gamma}) \;+\; \mathbf{B}(\mathbf{s})\mathbf{w}(\mathbf{s}). \tag{1}$$

In this recursion, $\mathbf{A}(s)$ and $\mathbf{B}(s)$ are matrices of appropriate dimension and the term $\mathbf{w}(\mathbf{s})$ represents white driving noise. The matrix $\mathbf{A}(s)$ captures the deterministic progression from node $s\bar{\gamma}$ to node s, i.e., the part of $\mathbf{x}(s)$ predictable from $\mathbf{x}(s\bar{\gamma})$, while the term $\mathbf{B}(s)\mathbf{w}(s)$ represents the unpredictable component added in the progression. An attractive feature of this framework is the efficiency it provides for signal processing algorithms. This stems from the Markov property of the multiscale model class, which states that, conditioned on the value of the state at any node s, the processes defined on each of the distinct subtrees extending away from node s are mutually independent.

For the application of segmenting different types of clutter in SAR imagery, a multiscale model can be constructed for each clutter class. To specify each model, it is necessary to determine the appropriate coefficients in the matrices $\mathbf{A}(s)$ and $\mathbf{B}(s)$, and the statistical properties of the driving noise $\mathbf{w}(s)$. Once the models have been specified, a likelihood ratio test can be derived to segment the imagery into the clutter classes.

Consider the problem of segmenting regions of forest and grass in SAR imagery. For each pixel in the image, we choose between two hypotheses: The pixel is part of a grass (H_g) or forest (H_f) region. The log-likelihood ratio test for classifying each pixel based on multiscale imagery is given by

$$\begin{aligned}\ell = &\log\left[p_{I_L,I_{L-1},\ldots,I_0|H_g}(I_L, I_{L-1},\ldots,I_0 \mid H_g)\right] \\ &- \log\left[p_{I_L,I_{L-1},\ldots,I_0|H_f}(I_L, I_{L-1},\ldots,I_0 \mid H_f)\right].\end{aligned} \tag{2}$$

By invoking the state space interpretation of multiscale imagery and exploiting the Markov property associated with the multiscale models, the log-likelihood ratio test for the two competing hypotheses can also be written as

$$\begin{aligned}\ell = &\sum_s \log\left[p_{\mathbf{x}(s)|\mathbf{x}(s\bar{\gamma}),H_g}(\mathbf{X}(s) \mid \mathbf{X}(s\bar{\gamma}), H_g)\right] \\ &- \sum_s \log\left[p_{\mathbf{x}(s)|\mathbf{x}(s\bar{\gamma}),H_f}(\mathbf{X}(s) \mid \mathbf{X}(s\bar{\gamma}), H_f)\right].\end{aligned} \tag{3}$$

Here, $p_{\mathbf{x}(s)|\mathbf{x}(s\bar{\gamma}),H_g}$ and $p_{\mathbf{x}(s)|\mathbf{x}(s\bar{\gamma}),H_f}$ are the conditional distributions for $\mathbf{x}(s)$ given $\mathbf{x}(s\bar{\gamma})$ for the two hypothesized models. In the next section, we will show that this likelihood test can be efficiently computed in terms of the distributions for $\mathbf{w}(s)$ under the two hypotheses.

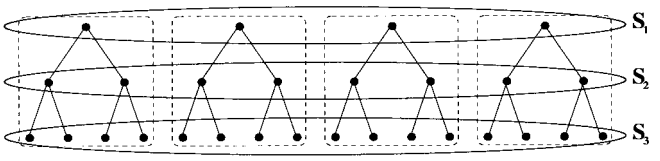


Fig. 4. Dyadic tree example illustrating the sets of nodes ($\mathbf{S}_i$) representing each scale of the process and the independent set of nodes (dashed boxes) representing subregions of the entire process.

C. Scale-Autoregressive SAR Model

In this paper, we focus on a specific class of multiscale models, namely scale-autoregressive models [6], [7] of the form

$$\begin{aligned}I(s) \;=\; &a_1(s)I(s\bar{\gamma}) + a_2(s)I(s\bar{\gamma}^2) \\ &+ \cdots + a_R(s)I(s\bar{\gamma}^R) + w(s)\,,\; a_i(s) \in I\!R\end{aligned} \tag{4}$$

where $w(s)$ is white driving noise. For homogeneous regions of texture, the prediction coefficients (the $a_i(s)$ in (4)) are constant with respect to image location for any given scale. That is, the coefficients, $a_1(s), \cdots, a_R(s)$, depend only on the scale of node s (denoted by $m(s)$), and thus will be denoted by $a_{1,m(s)}, \cdots, a_{R,m(s)}$. Furthermore, the probability distribution for $w(s)$ depends only on $m(s)$. Thus, specifying both the scale-regression coefficients and the probability distribution for $w(s)$ at each scale completely specify the model.

Following the procedure of state augmentation used in converting autoregressive time series models to state space models, we associate to each node s a R-dimensional vector of pixel values, where R is the order of the regression in (4). The components of this vector correspond to the SAR image pixel associated with node s and its first $R-1$ ancestors. Specifically, we define

$$\mathbf{x}(s) \;=\; [I(s) \quad I(s\bar{\gamma}) \quad \cdots \quad I(s\bar{\gamma}^{R-1})]^T. \tag{5}$$

The recursion in (1) takes on the form

$$\mathbf{x}(s)=\begin{bmatrix} a_{1,m(s)} & a_{2,m(s)} & \cdots & a_{R-1,m(s)} & a_{R,m(s)} \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}\mathbf{x}(s\overline{\gamma}) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} w(s). \tag{6}$$

Thus, for a model of the form (4) or equivalently (6), ℓ in (3) can be calculated using

$$p_{\mathbf{x}(s)|\mathbf{x}(s\overline{\gamma})}(\mathbf{X}(s) \mid \mathbf{X}(s\overline{\gamma})) = p_{w(s)}(W(s)) \tag{7}$$

where

$$W(s) = I(s) - \left[a_{1,m(s)}I(s\overline{\gamma}) + \cdots + a_{R,m(s)}I(s\overline{\gamma}^R)\right] \tag{8}$$

and $p_{w(s)}(W(s))$ is the probability density function for $w(s)$. By substituting (7) into (3), the test statistic can be written

$$\ell = \sum_s \log\left[p_{w(s)|H_g}(w(s) \mid H_g)\right] - \sum_s \log\left[p_{w(s)|H_f}(w(s) \mid H_f)\right]. \tag{9}$$

The likelihood test can thus be computed directly from the prediction error residuals, $w(s)$.

D. Identification of Multiscale Models of SAR Imagery

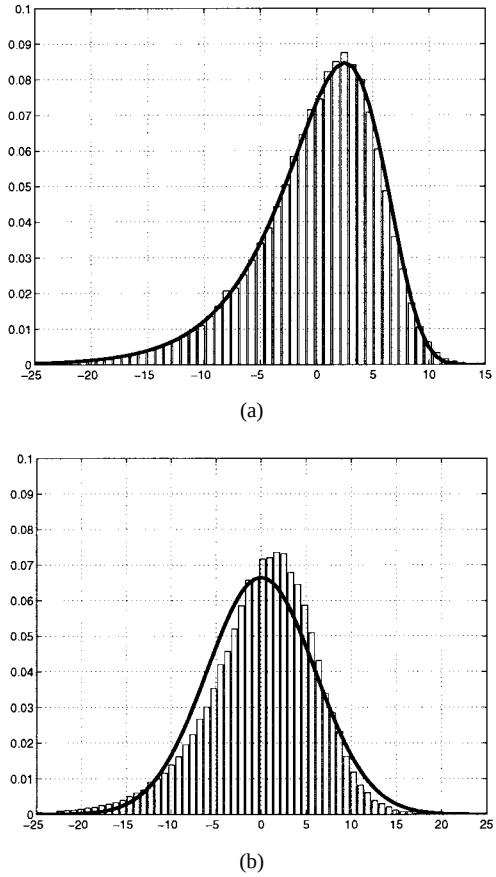
In order to use the previously described multiscale methodology, we need to identify the model parameters for each clutter class of interest, namely the model order R, the model coefficients $a_{1,m(s)}, \cdots, a_{R,m(s)}$ for each scale, and the probability distribution for $w(s)$ at each scale. To accomplish this, we choose a homogeneous training region of SAR imagery representative of each clutter class being modeled. This region is subsequently processed to produce a sequence of images, $I_L, I_{L-1}, \cdots, I_0$. The regression coefficients for each scale m are obtained by a standard least-squares minimization

$$\mathbf{a}_m = \arg\min_{\mathbf{a}_m \in I\!R^R}\left\{\sum_{\{s \mid m(s)=m\}} \left[I(s) - a_{1,m}I(s\overline{\gamma}) - \ldots - a_{R,m}I(s\overline{\gamma}^R)\right]^2\right\} \tag{10}$$

where

$$\mathbf{a}_m = \left[a_{1,m}\ a_{2,m}\ \cdots\ a_{R,m}\right]^T.$$

The regression length R, may be selected in a manner similar to that by which standard autoregressive (AR) model orders are chosen. For instance, one may also increase R until the statistical assumption of decorrelation throughout scale of $w(s)$ is satisfied and/or until algorithmic performance is acceptable. In [6] it was experimentally shown that for a regression length of one for the grass model, the residuals are spatially decorrelated. We found that by increasing the regression order to $R = 3$ for both grass and forest, we could achieve a lower probability of misclassification in homogeneous regions of terrain. Minimal performance gains were noted for larger regression orders as the scale-coefficients become negligible. Hence, the results presented in Section V were achieved using a third order regression ($R = 3$) for both the grass and the forest models.



(a)

(b)

Fig. 5. Histograms of residuals in prediction of second-finest resolution for (a) Grass model. (b) Forest model. Solid line represents (a) log-Rayleigh distribution and (b) Gaussian distribution.

To obtain a statistical characterization of the prediction error residuals (the $w(s)$ in (4)) of the model at scale m, we evaluate the residuals in predicting scale m of the homogeneous test region. In particular, we use the $\mathbf{a}_m$ found in (10) to evaluate all $\{\mathbf{w}(s)|m(s)=m\}$ in (8). We then choose a theoretical distribution that provides a good fit to the normalized histogram of these residuals. In Section V-A, we define the theoretical distributions chosen for the models for grass and forest and graphically illustrate their accuracy in statistically representing the prediction error residuals.

III. Natural Clutter Segmentation

In this section, we describe a procedure for the segmentation of SAR imagery consisting primarily of natural clutter. The starting point for this development is the construction of multiscale SAR models for different clutter types as described in the preceding section. We illustrate the approach by focusing on

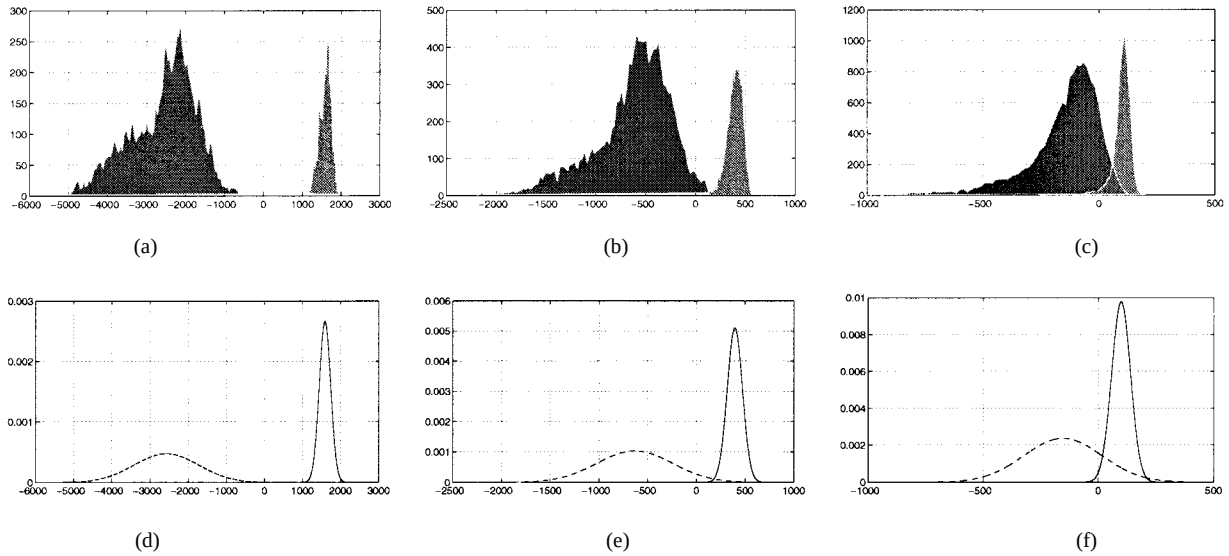


Fig. 6. Statistical results used in determination of threshold values for each region size. (a) Histograms of values of ℓ from 128 pixel square homogeneous regions of forest (dark) and grass (light.) (d) Gaussian estimates of ℓ for each terrain category (solid line for grass and dashed for forest). (b), (e) Similar results for 64 pixel square regions. (c), (f) Similar results for 32 pixel square regions.

distinguishing forested regions from grass. This methodology, however, may be extended to include delineation of additional clutter types.

A. Basic Structure of the Segmentation Procedure

We employ the multiscale models constructed for forest and grass in the preceding section to classify individual pixels and subsequently segment regions of clutter. One could also postulate a spatial random field model for each clutter category, capturing, for example, the fact that the classification of a given pixel is very likely to be the same as its neighbors [4]. The use of such a model, however, would likely increase the computational complexity of the classification algorithm considerably. Consequently, we employ a simpler approach that exploits the efficiency of multiscale likelihood calculation for scale-autoregressive models. Specifically, we classify each individual pixel based on a test window of pixels surrounding it.

The size of the window used in the pixel-by-pixel classification must be judiciously chosen. A larger window provides a more accurate classification of homogeneous regions. Using larger windows, however, increases the likelihood that the window contains a clutter boundary. Thus, keeping the window size as small as possible is also desirable. As demonstrated in the next section, by examining the empirical distribution of ℓ in (2) over windows of various sizes for homogeneous regions of grass and forest, we can determine the trade-off between classification accuracy and window size. This, in turn, allows us to choose the smallest window size that yields adequate performance in classifying homogeneous regions of clutter.

Whenever a clutter boundary is present within a test window, the validity of the center pixel classification is questionable. This effect results in a classification bias near boundaries. To address this problem, we devise a method to detect the proximity of grass–forest boundaries as well as a procedure to refine the subsequent classification. Terrain boundary proximity is detected via a simple modification of the decision made based on the test statistic ℓ. Specifically, rather than comparing ℓ to a single threshold to decide on a grass-or-forest classification, we compare ℓ to the two thresholds a and b as follows:

$\ell > a$	Classify as grass,
$a > \ell > b$	Defer decision (possible boundary presence),
$\ell < b$	Classify as forest.

The resulting test structure is illustrated in Fig. 2, in which the box designated “defer” corresponds to the test statistic falling between the two thresholds. In such cases, a refinement procedure, described below, is used to classify the pixel (hence, the designation *defer*). As we describe in Section V-A, the choice of thresholds a and b are determined by examining empirical distributions of ℓ for windows containing boundaries with varying percentages of forest and grass.

B. Deferred Pixel Classification

For pixels where the classification decision has been deferred, it is necessary to determine the appropriate classification. The structure of the multiscale likelihood calculations allows us to perform this additional task as a replication of the classification procedure depicted in Fig. 2 at a hierarchy of scales. Recall that the objective of this process is to classify the center pixel of a window as either grass or forest. Consequently, in a region that is likely to contain a boundary, it is necessary to determine on which side of the boundary the center pixel lies. Under the assumptions that, at most, one terrain boundary resides within the window and that this boundary is relatively smooth, the center pixel may be classified with a high level of confidence by merely determining which of the two hypothesized clutter types occupies the majority of the window.

The majority rule decision is accomplished in a recursive manner by first independently considering the four quadrants of the initial window region. A classification procedure similar (the only difference being the thresholds a and b) to the one described in Section III-A is performed on each of these *quadrants*, with the classification assigned to the entire corresponding region. This divide-and-conquer approach is repeated recursively on each deferred subquadrant until one of two criteria is met.

1) From the classification of the subregions of the initial window, a majority rule decision for the windowed region may be determined.
2) The subregional size becomes so small that a statistically significant decision may not be made.

In the latter case, a majority rule decision is made based solely upon the regions that have been classified up to that point. In the event that a majority rule decision may not be made at either of these points (e.g., half of the classified area is grass, and the other half forest), we assign a center pixel classification of grass (note that it should matter very little which classification is assigned in this case since the center pixel likely lies very near a terrain boundary). Clearly, the assumptions made above will not always be met. We have found experimentally however, that by restricting the size of the initial window, we may not only reduce the likelihood that the window contains a terrain boundary, but also increase the likelihood that the boundary meets the above criteria.

This "progressive refinement" around boundary regions is depicted in Fig. 3. In each frame, the pixel of interest is exaggerated by the solid square box in the center of the image. Fig. 3(a) represents the original windowed region about the center pixel. Based on the test illustrated in Fig. 2, this window has been identified as potentially containing a grass–forest boundary. Note that the correct center pixel classification for this region is grass, yet due to the boundary proximity, the classification has been deferred. Figs. 3(b)–3(d) display the successive subdivision and reclassification of the entire windowed region (F, G, and D represent classification as forest, grass, and defer, respectively), where at each stage only the regions deferred at the preceding stage are subject to further examination. Note that in Fig. 3(d), we have met at least one of the stopping criteria (i.e., all subregions have been classified), and may classify the center pixel according to the majority rule of the terrain classifications within the window. We see that $9/16$ of the region has been classified as grass; hence, the center pixel is ultimately correctly classified as grass.

We mentioned above that the structure of the multiscale likelihood calculations allows us to perform the refinement procedure with minimal further computational cost. We illustrate this point for a process defined on the dyadic tree displayed in Fig. 4. In this figure, the nodes at each scale of the process are represented by the sets $\mathbf{S}_i, \quad i = 1, 2, 3$. As previously mentioned, ℓ may be calculated as in (9) by summing the individual log-likelihood differences. Hence, for the process in Fig. 4 and for the discrimination between terrain

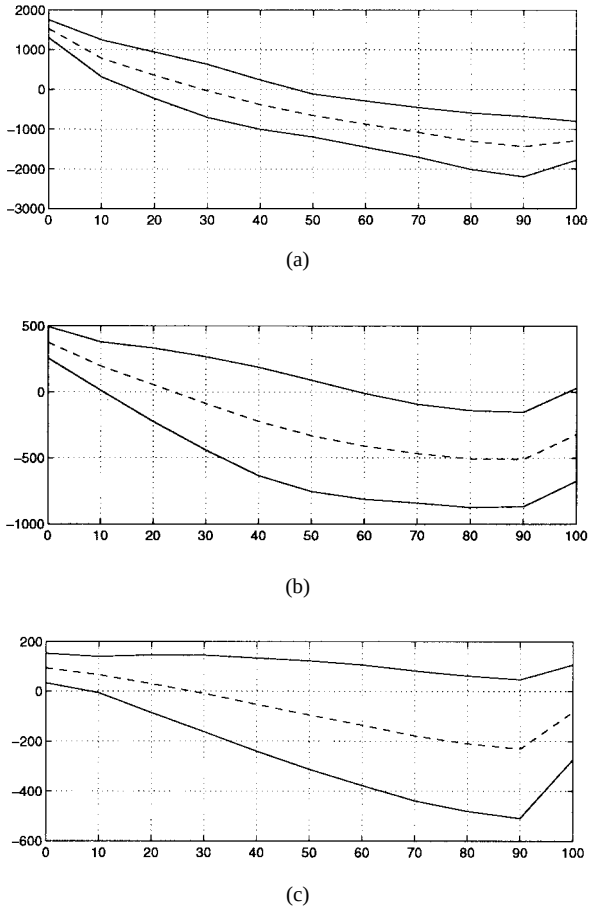


Fig. 7. Mean value of the test statistic ℓ versus percentage forest in window region. Frame (a) displays plots of the test statistic mean (dashed line) as well as plots of its plus and minus 2σ points (solid lines) for a 128-pixel window. (b) and (c) display similar data for 64- and 32-pixel window sizes, respectively.

models M_1 and M_2, ℓ may be calculated as

$$\ell = \sum_{s \in \mathbf{S_2}, \mathbf{S_3}} \log \left[p_{w(s)|M_1}(w(s)|M_1) \right] - \log \left[p_{w(s)|M_2}(w(s)|M_2) \right]. \tag{11}$$

Since we do not have a full set of measurements (i.e., the tree does not proceed up to the root node), the measurements may be split into independent sets, as distinguished by the dashed boxes in Fig. 4. Clearly, whereas the evaluation of ℓ for the entire region involves the calculation of individual log-likelihoods over the entirety of sets $\mathbf{S}_2$ and $\mathbf{S}_3$, the evaluation of ℓ over each subregion involves the summation over *subsets* of $\mathbf{S}_2$ and $\mathbf{S}_3$. Hence, ℓ may be evaluated in a computationally attractive manner by merely summing existing quantities. The ideas presented in this example are easily extended to the quadtrees used in SAR image representation.

IV. Anomaly Enhancement

The methods described in the previous section address two of the principal objectives of this paper, namely, terrain classification and terrain boundary estimation. In this section, we discuss a method aimed at the third objective discussed in Section I, namely, the use of terrain classification to enhance

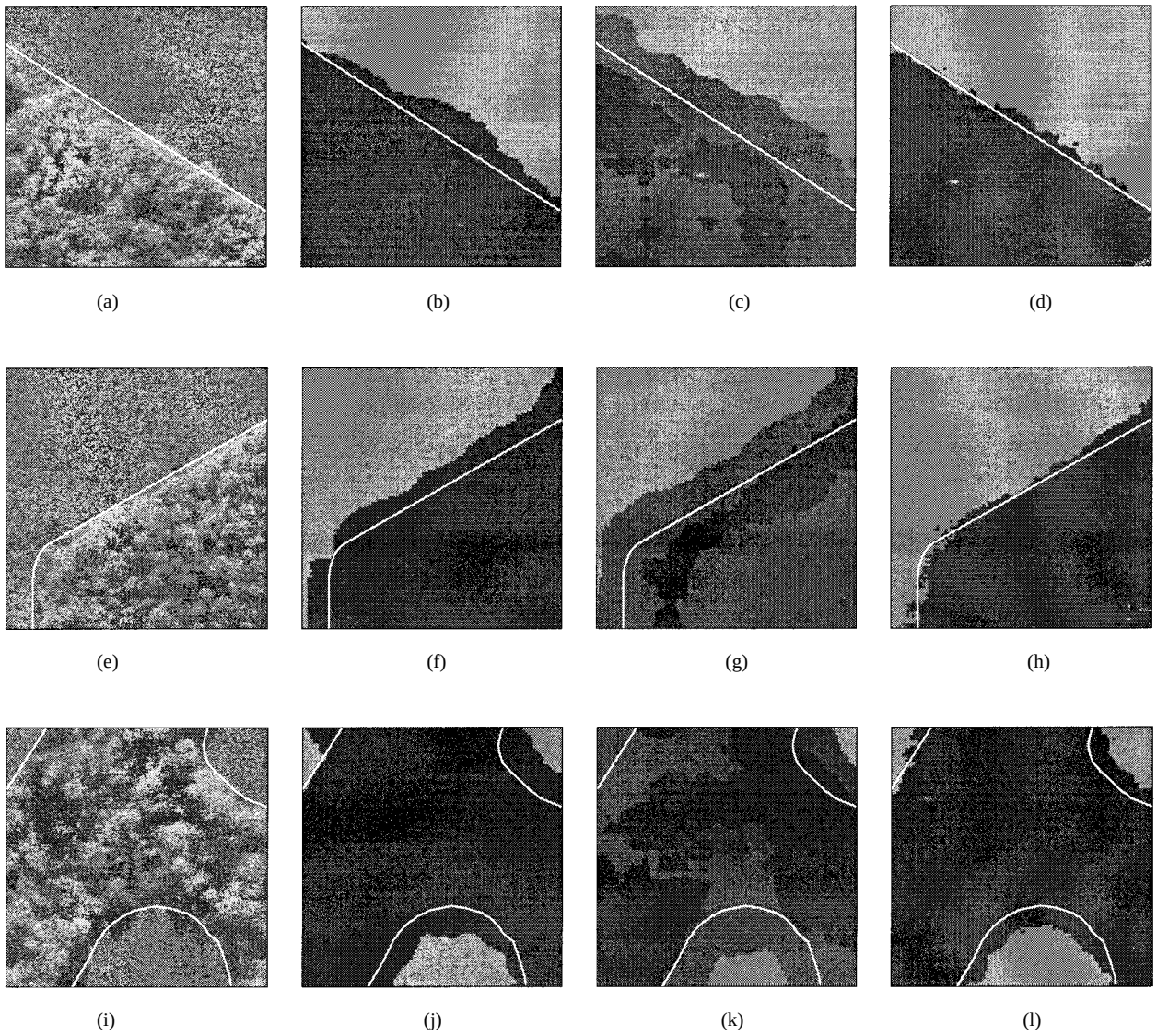


Fig. 8. Segmentation results for three (top row to bottom row) 512 pixel square images of 0.3-m resolution stripmap SAR data. The black line in each frame represents the eyed estimate of the grass–forest boundary. (a), (e), f) Original SAR image. (b), (f), (j) Segmentation results without deferral. (g), (h), (l) Preliminary classifications using double threshold (dark green = forest, olive green = grass, light green = deferred decision). (d), (h), (l) Final segmentation after refinement procedure.

anomalies (i.e., to make anomalies due to cultural clutter and targets to stand out from their background). Enhancement of cultural clutter image regions provides the potential for improvement in the detection of potential target regions (i.e., man-made objects) through simple thresholding. That is, higher probabilities of target detection are possible by thresholding the enhanced imagery.

A commonly accepted method to enhance anomalous pixel regions is the constant false alarm rate (CFAR) procedure [8], [9]. In this procedure, a local sample mean and standard deviation are estimated from an annular region around each pixel to characterize the second-order statistics of the background. In particular, the CFAR statistic $\chi(k,l)$ for pixel $I(k,l)$ in the finest-scale log-detected (decibel) image is defined as

$$\chi(k,l) = \frac{I(k,l) - \hat{\mu}(k,l)}{\hat{\sigma}(k,l)} \tag{12}$$

where $\hat{\mu}(k,l)$ and $\hat{\sigma}(k,l)$ represent the background sample mean and standard deviation, respectively. Anomalous pixels are identified as outliers with respect to this estimated background distribution.

A. Multiscale CFAR Statistic

The ideas above suggest a method for identifying potentially anomalous pixels using what can be thought of as a multiresolution CFAR statistic. This statistic may be viewed as a synthesis of our multiscale modeling approach with the method described in [10]. Specifically, suppose a region has been identified as consisting of grasslike terrain either by the segmentation method described in the preceding section or by some other means (such as from prior maps of the region of surveillance). Now, consider a particular pixel in the finest-scale SAR imagery of that region. As described in

Section III, we can associate with that pixel a set of multiscale SAR imagery over a window centered at that pixel. If we let s_c denote the node associated with the central pixel under investigation, then using the parameters associated with the grass model, the residual $w(s_c)$ in (8) represents the error in predicting the central pixel at the finest resolution based on the SAR imagery at R coarser resolutions. From our multiscale model we have a theoretical mean μ_c and variance p_c associated with this residual, and can thus compute a statistic, $\zeta(s_c)$, that is analogous to the CFAR statistic

$$\zeta(s_c) = \frac{w(s_c) - \mu_c}{\sqrt{p_c}}. \tag{13}$$

As with the CFAR statistic, this statistic exploits the characteristics of the background imagery. In particular, when the pixel classification is that obtained using the segmentation procedure in Section III, we note that it is based on a window of imagery surrounding the pixel of interest. As a result, the classification characterizes the background. Accordingly, we subtract the background mean associated with the corresponding terrain model from the statistical value (each model assumes zero-mean residuals), and normalize by the corresponding terrain model standard deviation.

The multiscale segmentation algorithm computes the prediction error residuals required to determine the statistic (13). The variance p_c may be found from the theoretical distribution specified by the multiscale terrain models. Furthermore, because the residuals are computed at a set of scales, the scale-to-scale behavior of the CFAR statistic can be characterized and used to improve detection performance. Specifically, for a node s that is m tree levels up from the finest resolution of the process, we can also compute $w(s)$ as in (8) and compare it to the corresponding theoretical variance p_m from scale m of the multiscale model. Specifically, generalizing (13), we can compute

$$\zeta(s) = \frac{w(s)}{\sqrt{p_m}} \tag{14}$$

for sth on the m level from the bottom of the tree. Note that in (14) we disregard the mean since for all scales of each model we assume zero-mean residuals.

B. Combined Multiscale CFAR Statistic

By considering the set of random variables acquired by evaluating (14) at multiple scales, $\{\zeta(s), \zeta(s\overline{\gamma}), \cdots, \zeta(s\overline{\gamma}^{P-2})\}$, we allow for several methods of obtaining statistics to enhance the identification of anomalies. For example, one possible statistic is the sum of squared values of these random variables

$$\begin{aligned} c_1(s_c) = {} & \zeta^2(s_c) + \zeta^2(s_c\overline{\gamma}) + \zeta^2(s_c\overline{\gamma}^2) \\ & + \cdots + \zeta^2(s_c\overline{\gamma}^{P-2}) \end{aligned} \tag{15}$$

where P represents the number of scales in the state-augmented multiscale model.

As with the standard CFAR statistic, $c_1(s_c)$ will enhance any behavior that deviates significantly from that expected for nonanomalous pixels. However, in contrast to the standard CFAR statistic, we can consider alternatives since we have a *set* of statistics rather than one. In particular, we can consider looking for specific scale-to-scale signatures in the residuals $w(s_c)$. Indeed, as the results in [6] and [10] indicate, the presence of a few dominant scatterers in man-made objects will result in scale-to-scale variations in SAR imagery with a more deterministic behavior across scale and thus do not obey a simple white noise driven autoregressive model. If we can predict this scale-to-scale behavior, we can seek telltale signatures in the resulting residuals. A simple example of this is the statistic

$$c_2(s_c) = \left[\zeta(s_c) + \zeta(s_c\overline{\gamma}) + \zeta(s_c\overline{\gamma}^2) + \cdots + \zeta(s_c\overline{\gamma}^{P-2})\right]^2. \tag{16}$$

Comparing (15) and (16), the difference is that in (16) we sum before squaring. This corresponds to looking for a consistent over- or underprediction of pixels at successive resolutions, as one might expect for a single strong scatterer. In such a case the statistic $c_2(s_c)$ would give a significantly larger value than $c_1(s_c)$. In addition, $c_2(s_c)$ would attenuate the effects of other types of scale-to-scale behavior. For example, if the signs of the successive values of ζ change throughout scale, then $c_2(s_c)$ will be much smaller than $c_1(s_c)$. Thus the statistic in (16) is far more selective in what it enhances, which will be of value if we have accurate models for the type of scatterer distributions we expect in objects of interest.

The two statistics $c_1(s_c)$ and $c_2(s_c)$ both provide means of identifying and enhancing pixels that exhibit anomalous behavior. They both involve a squaring, in some sense, of the scale-to-scale statistics nevertheless. This will enhance pixels that are outliers in the positive and negative tails of the model distribution. Due to the nature of pixels corresponding to man-made imagery, we are only concerned with the positive tail. In other words, we are only interested in those values of $c_2(s_c)$, for example, that display a consistent overprediction. Hence, we consider the third statistic

$$c_3(s_c) = \left[\zeta(s_c) + \zeta(s_c\overline{\gamma}) + \zeta(s_c\overline{\gamma}^2) + \cdots + \zeta(s_c\overline{\gamma}^{P-2})\right]. \tag{17}$$

This statistic will enhance only those pixels that display a consistent overprediction and not those displaying a consistent underprediction. As a result, we may simply threshold this statistic, as with the standard CFAR statistic, to identify potential targets of interest.

V. Experimental Results

In order to evaluate the performance of our approaches to SAR image segmentation and anomaly enhancement, we have applied them to 0.3-m resolution horizontal-horizontal (HH) polarization SAR data gathered over Stockbridge, New York, with Lincoln Laboratory's millimeter-wave SAR [5]. We independently evaluate the segmentation and anomaly enhancement performance in Sections V-A and V-B, respectively.

A. Segmentation Performance

We begin by describing the construction of the models on which the subsequent experiments are based in Section V-

A1. In Section V-A2, we discuss the details of the algorithm design, namely the setting of window sizes and decision thresholds. In Section V-A3 we then illustrate the performance of our algorithm for the main objectives of the paper, namely classification and segmentation.

1) Model Construction: The first step in applying our methods is the construction of multiscale models for SAR imagery of grass and forest from homogeneous regions of terrain. We have chosen to use a third-order regression for each model and to build models for the prediction of each of the three finest resolution images ($\delta \times \delta$, $2\delta \times 2\delta$, and $4\delta \times 4\delta$, with $\delta = 0.3m$). As described in Section III-B, this implies a third-order model with a four-level tree, and for the prediction of each of the three finest resolution images there are three coefficients, a_1, a_2, a_3, to be specified. Using the method described in Section II-D, we determined the coefficient values given in Table I(a) and (b). Note that the coefficients for the forest model are consistently larger, indicating higher scale-to-scale correlation. This is consistent with what one would expect based on the interpretation of scale-to-scale correlation as capturing information about the distribution of scatterers in the image: Grassy regions tend to have larger numbers of equivalued scatterers and, thus, since greater numbers of scatterers migrate out of each resolution cell as we move from scale to scale, we would expect that SAR imagery of grass to have less scale-to-scale correlation.

To complete the models, we specify the distributions for $w(s)$ in (4) for each scale and each terrain type. For imagery of grass, a log-Rayleigh distribution

$$p_{w(s)}(W(s)) = k\,exp[k\,W(s) - \gamma - exp(k\,W(s) - \gamma)]$$
$$k = \frac{ln(10)}{10}$$
$$\gamma \approx 0.577\,215\,66 \text{ (Euler's constant)}$$

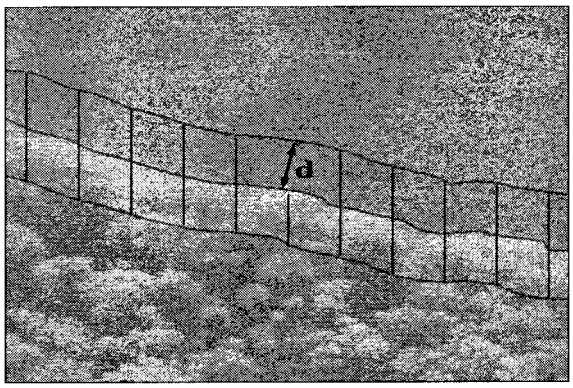
provides a good fit at each scale, while Gaussian distributions are a better choice at each scale of the forest model. In particular, the standard deviations chosen for the Gaussian densities representing $w(s)$ for each of the three scales predicted are as follows:

$$\delta \times \delta:\ 5.3724, \quad 2\delta \times 2\delta:\ 6.1811, \quad 4\delta \times 4\delta:\ 6.6056.$$

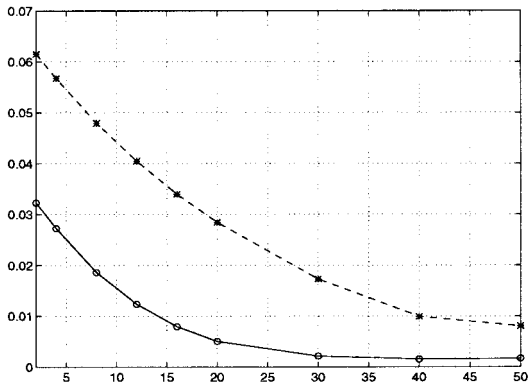
Rather than following statistical goodness-of-fit tests to determine the theoretical distributions for our residuals, we relied on the results of Irving *et al.* [6]. In their work, it was shown that the log-Rayleigh distribution provides an acceptable fit for regions of grass. We found, however, that the log-Rayleigh distribution did *not* provide as good a fit for regions of forest as did a Gaussian distribution. Examples showing histograms and the resulting fits are provided in Fig. 5.

2) Algorithm Specification: In order to specify our classification and segmentation algorithm, we need to set

1) the window size around each pixel used to perform each classification;
2) the thresholds at each stage in the hierarchical procedure.



(a)



(b)

Fig. 9. (a) Illustration of swath (hashed) of width d ignored in misclassification probability calculations. (b) Probability of misclassification versus swath width on either side of boundary for grass–forest boundaries with average illumination (solid) and all levels of illumination (dashed).

On row (i) of Fig. 6, we display histograms of the value of the statistic ℓ computed over homogeneous regions of both terrain categories using windows of varying size. In addition we have also displayed, on row (ii), the corresponding Gaussian fits to the histograms. For computational convenience, we limited our interest to window sizes that were powers of 2 (128, 64, and 32 pixels square). By noting the statistical properties of ℓ for these window sizes, we gain the ability to choose both the largest and smallest window sizes for our subdividing refinement procedure simultaneously. As mentioned in Section III-A, our choice of initial window size involves a trade-off between misclassification probability in homogeneous terrain regions and the likelihood of the window containing a terrain boundary. From the histograms and associated Gaussian fits in Fig. 6(a) and (d) it is apparent that a 128 pixel square window size yields a sufficiently small probability of misclassification[1] in homogeneous regions of terrain (approximately 1.5×10^{-5}). Furthermore, for the 32 pixel square window size (Fig. 6(c) and (f)), there is certainly still distinguishing information between the two classes yet a

[1]Although the Gaussian distributions on row 2 do not seem to provide adequate fits for the histograms in row 1, they *do* provide the ability to estimate the probability of misclassification (at each hierarchical level of the refinement procedure) for larger sample spaces. Otherwise, from the information given, we would empirically estimate the probability of misclassification for frames (a) and (b) to be zero.

reduction in statistical significance, as is seen by the increased overlap in the histograms. Further subdivision in window size (i.e., 16 pixel-square) provided minimal distinguishing information between the two classes. Accordingly, for the results presented in this paper we have chosen an original window size of 128 $\times$ 128 and considered subdivisions down to a size of 32 $\times$ 32 for subsequent reclassification of deferred pixels.

If we were only considering classification of homogeneous regions and thus only needed to set a single threshold level for each size window, then Fig. 6 would provide all of the information required to accomplish this. However, as we have indicated, we also wish to ensure good performance in classifying pixels near tree lines and, consequently, have improved boundary estimation. In order to see how the presence of a boundary within a window changes the distribution of the resulting test statistic, in Fig. 7 we display plots of the mean of ℓ plus and minus 2σ calculated for windows containing varying percentages of forest[2]. An immediate point to note is that the presence of even a modest amount of "contamination" by forest pixels can change the behavior of the test statistic significantly. As a result, an effective bias is introduced in the resulting segmentation when no deferred classifications are allowed. As a consequence, it is essential that grass–forest decisions at the level of the full 128 $\times$128 pixel region be made only if there is overriding evidence for one of these two hypotheses. Similarly, the decisions at the subsequent 64 $\times$ 64 and 32 $\times$32 pixel regions must exhibit similar characteristics.

The thresholds were first approximated using Fig. 7, then refined after experimentation to maximize performance. For example, as mentioned above, we necessitate overriding evidence of homogeneity to classify definitively as either grass or forest. As a result, we started with grass and forest thresholds for the largest 128 $\times$ 128 pixel region size (a and b) of 1800 and $-$2000, respectively. These thresholds are approximately the highest and lowest 2σ points in Fig. 7(a). We found experimentally, however, that by relaxing these thresholds we were able to obtain a gain in performance. Hence, we obtained the pairs of thresholds in Table II. As before, a value above the *grass* threshold leads to classification as grass, whereas a value below the *forest* threshold leads to classification as forest. A value in between the two thresholds again results in a deferred classification. If a deferred decision is made at the 128 $\times$ 128 or 64 $\times$ 64 pixel levels, we continue down to the next level of classification. If it is made at the 32 $\times$ 32 region, we classify that region as defer and stop. However, as we have described in the preceding section, these classifications are used solely to determine the classification of the center pixel under test.

3) Classification and Segmentation Performance: We apply our segmentation technique to SAR imagery to investigate its performance in terms of accuracy of classification and segmentation. The accuracy of classification of forest and grass pixels embedded in homogeneous regions of the same terrain type is completely characterized by two numbers, namely, the probabilities of each type of classification error. Note that if we knew *a priori* that a region is homogeneous, then, based on Fig. 6(a) we would already know that the probability of misclassification of either type is extremely small. In fact, for the Gaussian probability distributions shown in Fig. 6(d), with a single threshold set at a value of 910, the probabilities of misclassifying forest as grass and grass as forest are both approximately equal to 1.5×10^{-5}. However, since we wish to use our algorithm in contexts in which there may be boundaries, a fairer test involves using the complete hierarchical decision procedure. To calculate the misclassification probabilities for this algorithm, we have performed tests on 20 512–pixel square segments of SAR imagery, ten consisting of homogeneous imagery of forest and ten of grass. We note that this set of imagery used in performance testing is completely independent of the training set used in model creation. The resulting misclassification probabilities were approximately 0.005 for misclassifying grass as trees and 0.011 for misclassifying trees as grass. The corresponding 2σ confidence intervals for these probabilities are approximately 0.01 for misclassifying grass as trees and 0.02 for misclassifying trees as grass.

The next set of experiments focuses on evaluating the performance of our algorithm in terms of segmentation (i.e., in terms of the accuracy in estimating the boundary between forest and grass). Fig. 8 displays the results of applying our algorithm to three images, top row, middle row, and bottom row. The black line in each frame represents a hand-picked estimate of the boundary, and dark and light regions represent terrain classification as forest and grass, respectively. Each row displays the segmentation results for the SAR image pictured in the left-most frame of each row. The other three frames of each row display the classification results using (in order) a single threshold for the full 128 $\times$ 128 window, two thresholds (a and b) at the full window size, and the full hierarchical approach described in Section III. Comparing the second and fourth frames for each image, we first note that restriction to a single threshold in the 128 $\times$128 window likelihood test leads to a bias in the estimated boundary; the tree line is pushed into the grassy region. However, when the full hierarchical system is implemented, the final classification presents a more accurate identification of the boundary.

In addition, we have computed a quantitative measure of boundary estimation accuracy by defining a parameter d and computing the empirical probability of misclassification of pixels that are more than d pixels away from the true boundary. Thus for $d = 0$ we consider the misclassification probability for all pixels in the image. As d increases we ignore a swath of the image, of width d, around the boundary (see Fig. 9(a)). We would thus expect that the misclassification probability would decrease with increasing d, approaching the homogeneous misclassification probability. The faster the drop in misclassification probability with d, the smaller the effective error in estimating the boundary. If this algorithm were to be followed by a target detection step aimed at looking for targets that are partially obscured by a tree-line, producing an accurate determination of the boundary would enhance the performance

[2]The upswing in the mean value of ℓ in each case at approximately 90% forest may be attributed to the large number of bright scatterers generally found along terrain boundaries. The correlation structure of these scatterers tends to drastically alter the statistical value of ℓ. As the window passes this highly illuminated region (i.e., becomes 100% forest), these scatterers no longer affect the value of ℓ.

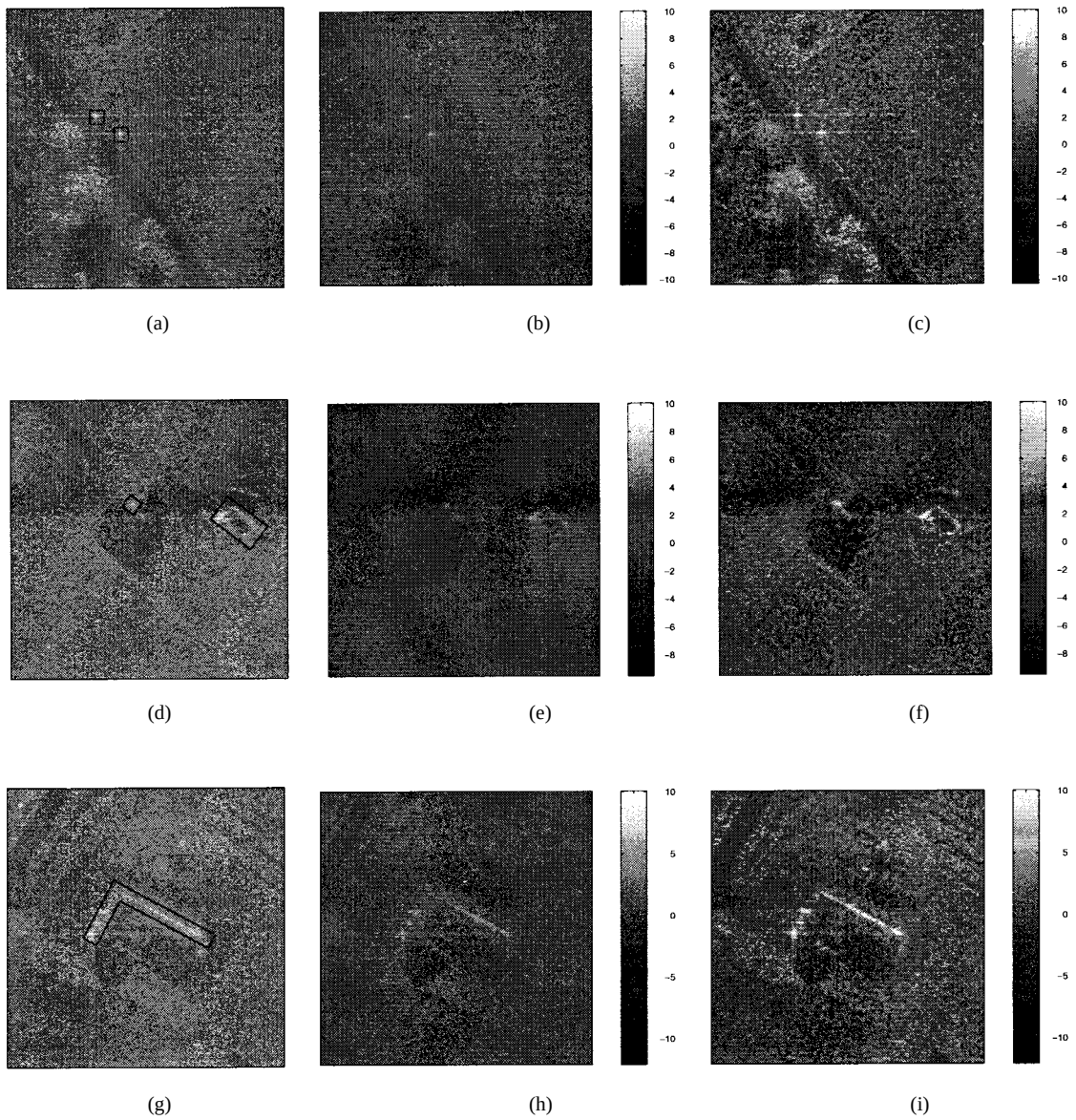


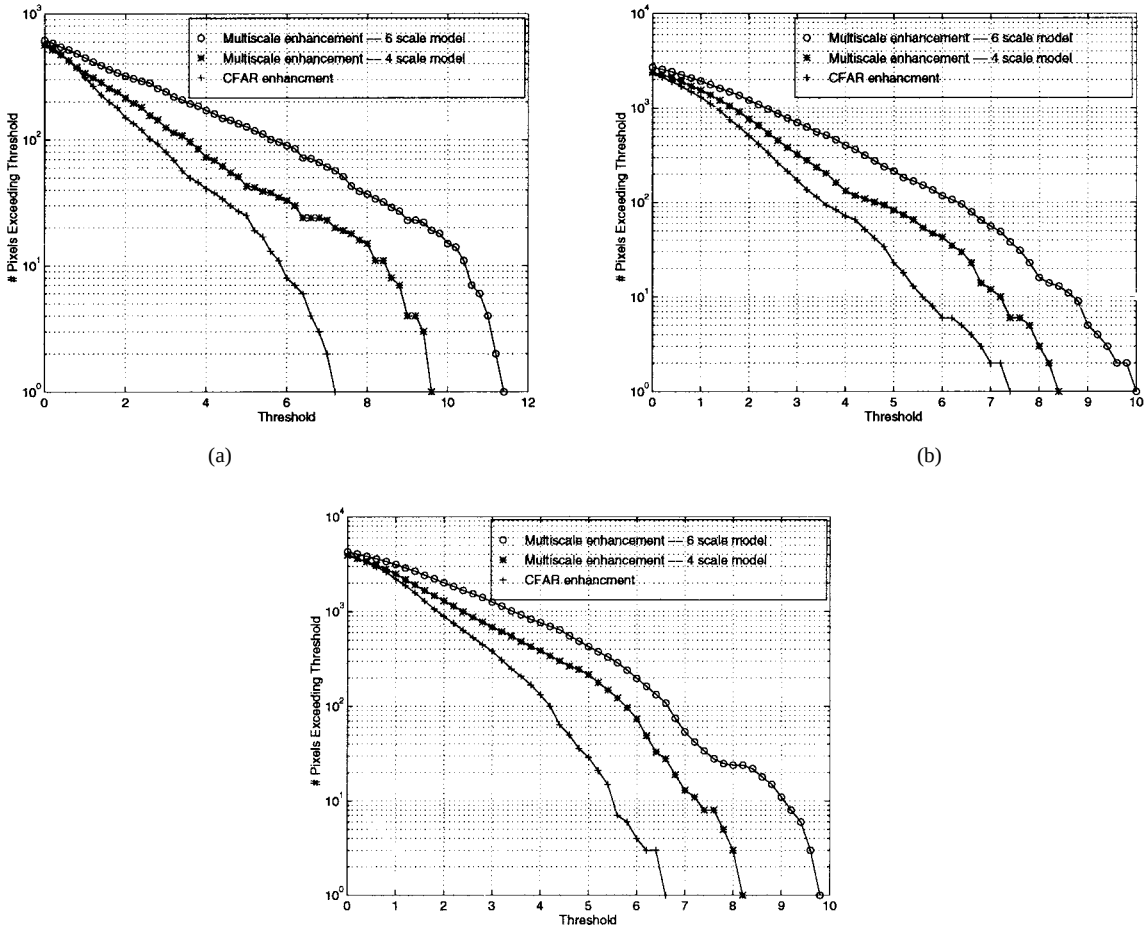
Fig. 10. Anomaly enhancement results for three (top row to bottom row) 512-pixel square images of 0.3-m resolution stripmap SAR data. (a) Original SAR image. (b) CFAR Enhanced image. (c) Multiscale enhanced image.

of that subsequent detection step. In Fig. 9(b) we display the results for our algorithm based on regions of Lincoln SAR imagery containing grass–forest boundaries with average tree line illumination (solid line) and regions containing boundaries with all levels of illumination. From these tests, we see that for boundaries with average illumination we are able to determine the boundary within approximately 7 pixels ($2m$) with a probability of misclassification of 0.02. However, if all illuminations are considered, we can isolate the boundary to within 27 pixels ($8m$) with the same probability of misclassification.

B. Anomaly Enhancement

We describe here the methods we have chosen to measure the performance of the multiscale anomaly enhancement technique described in Section IV. We will focus on performance in the ATR environment, and accordingly, compare it to the standard CFAR enhancement algorithm used in practice. As test data, we use three HH polarized 0.3-m foot resolution SAR images, each containing a different type of cultural clutter. We illustrate the potential of our approach by presenting side-by-side comparisons with CFAR-enhanced imagery. We further demonstrate the potential by providing a qualitative measure of pixel enhancement over target regions. While neither of these methods yields the definitive performance measure of a receiver operating characteristic (ROC) curve, they *do*, however, demonstrate the promise that a multiscale-based anomaly enhancement technique affords.

1) Statistic Specification: The three 512–pixel/square SAR images used as test data are displayed in Fig. 10, left column. Each image contains at least one man-made object. The highly reflective portion of each object is outlined by the black boxes



(a)

(b)

(c)

Fig. 11. Anomaly enhancement performance plots for the three images in Fig. 9(a)–(c). Number of pixels exceeding a threshold value is plotted versus threshold.

TABLE I
MODEL COEFFICIENTS FOR THIRD-ORDER REGRESSION IN SCALE. (a) GRASS MODEL COEFFICIENTS. (b) FOREST MODEL COEFFICIENTS.

Grass model coefficients			
Resolution	a_1	a_2	a_3
$\delta \times \delta$	0.5263	0.0720	−0.0029
$2\delta \times 2\delta$	0.3135	0.0313	−0.0064
$4\delta \times 4\delta$	0.2278	0.0169	−0.0006

(a)

Forest model coefficients			
Resolution	a_1	a_2	a_3
$\delta \times \delta$	0.5842	0.1257	0.0669
$2\delta \times 2\delta$	0.5005	0.1222	0.0683
$4\delta \times 4\delta$	0.4584a	0.1292	0.0250

(b)

in each frame. Since these highly reflective regions typically set man-made objects apart from natural clutter, we focus on the performance of each algorithm exclusively over these outlined image regions.

TABLE II
THRESHOLD VALUES FOR VARIOUS WINDOW SIZES. THESE VALUES SERVE AS THRESHOLDS IN THE TERNARY HYPOTHESIS TEST FOR CLASSIFICATION AS GRASS, FOREST, OR DEFER AT EACH HIERARCHICAL LEVEL OF THE ALGORITHM.

Region Size	Grass Threshold	Forest Threshold
128	1000	-1600
64	500	-800
32	50	0

For each image displayed, we evaluate both the CFAR statistic and the multiscale statistic, $c_3(s_c)$. The prediction error residuals were calculated using a third-order scale-autoregressive grass model spanning four scales ($P = 4$). In each row, the second and third frames display the enhanced images using the CFAR enhancement method and the multiscale enhancement method, respectively. In each of the three rows, it appears that the multiscale enhancement technique increases the visibility of cultural clutter regions over the CFAR enhancement method. We see below that for these examples this is in fact the case.

The *blocky* appearance in each of the multiscale enhanced images in Fig. 10 is an artifact of the multiscale structure of the calculation of $c_3(s_c)$. This is apparent by considering, for

TABLE III
PEAK AND AVERAGE VALUES OF EACH NORMALIZED STATISTIC OVER TARGET REGIONS (OUTLINED IN BLACK) FOR IMAGERY IN FIG. 10(a), (d), AND (g).

	Normalized Statistic					
	CFAR		$c_3(s_c)$, $(P=4)$		$c_3(s_c)$, $(P=6)$	
Image	Peak	Average	Peak	Average	Peak	Average
Figure 10-a	7.40	0.91	9.61	1.13	11.32	2.13
Figure 10-d	7.38	0.65	8.50	0.94	9.97	1.62
Figure 10-g	6.62	1.03	8.19	1.32	9.82	1.97

example, two adjacent pixels at nodes s_1 and s_2 in the finest-scale image for which $s_1\overline{\gamma}=s_2\overline{\gamma}$. Clearly, $c_3(s_1)$ and $c_3(s_2)$ share the same ancestry and will, as a result, display a higher level of correlation than finest-scale pixels, s_3 and s_4, for example, for which $s_3\overline{\gamma}^2=s_4\overline{\gamma}^2$. Although the blocky nature of this statistic *does* degrade visual quality, we find it to be of minimal consequence in the identification (not localization) of regions containing objects of interest.

2) Normalization: To perform a qualitative comparison of the two statistics, we must account for enhancement of natural clutter by each statistic due to differences in variance. For image regions that are statistically consistent with their background, the CFAR statistic will have zero-mean and unit variance. The statistic, $c_3(s_c)$, on the other hand, will have zero mean but a variance of $P-1$ (provided that the residuals are decorrelated throughout scale). We could normalize this statistic by $P-1$ to ensure a fair comparison, yet if the residuals are not truly white throughout scale, we will not accomplish our goal. Hence, we estimate the variance. We characterize the performance over grass regions since most forested regions will be rejected by the segmentation algorithm. Thus, we estimate the mean and standard deviation of each statistic for SAR imagery of grass. We then normalize each enhanced image such that each statistic will have zero-mean and unit variance over regions of natural clutter (grass). As a result, we may directly compare each normalized image to determine which provides better enhancement of anomalous regions.

3) Performance Measures: As an initial measure, we compare the peak and average enhanced pixel values over the areas of interest (outlined in black) for each of the SAR images in Fig. 10, from top to bottom of the left-most row. The peak value provides a measure of probability of detection, since each image will ultimately be thresholded to detect anomalous regions. The average value, on the other hand, yields a measure of overall target enhancement. We evaluate three statistics: CFAR, and $c_3(s_c)$ for multiscale models spanning four and six scales. The results listed in Table III indicate the potential of a multiscale enhancement technique in increasing visibility of anomalous (i.e., man-made) regions. We see that for each SAR image, the peak value over the target is greater in each multiscale enhanced image than in the CFAR enhanced image. Furthermore, an increase in the number of model scales (4 versus 6) provides better anomalous pixel visibility.

We evaluate a second performance measure that further supports our claim that a multiscale anomaly enhancement technique increases the visibility of cultural clutter. This measure evaluates the number of target pixels exceeding a given threshold, versus threshold. Clearly, this measure will provide a performance curve that begins at some value (corresponding to the total number of target pixels) and monotonically decreases to zero. The rate at which the curve decreases provides a measure of performance (i.e., a "higher" curve is indicative of increased ability for detection).

We evaluate this measure for the three normalized statistics listed in Table III over each of the test images in Fig. 10. The results for the three images Fig.10(a), (d), and (g) are, respectively, displayed in Fig. 11(a)–(c). It is apparent that the multiscale technique provides better enhancement of the target regions than the CFAR algorithm. For example, for the L-shaped structure in Fig. 10(g) the plotted results in Fig. 11(c) show that using the multiscale enhancement technique with a four-scale model, we may set a threshold value of 8 and still detect the target. With the threshold set at this value, the CFAR algorithm will not detect this target. By allowing for higher threshold values, the multiscale enhancement technique will subsequently reduce the clutter false alarm rate and improve detection performance. Furthermore, it is clear from Fig. 11 that the performance of the multiscale enhancement technique improves when using the larger model size (six scales).

VI. CONCLUSIONS AND EXTENSIONS

In this paper, we have described a methodology for the classification of background clutter using multiscale models of SAR imagery that exploit the differences in interscale variability and predictability of images of different types of terrain. In addition, we have outlined a method by which we employ the prediction error residuals that are essentially a byproduct of this segmentation, and use them to enhance anomalous pixel regions for man-made object detection.

A. Segmentation

We believe that the segmentation results that we have presented, as well as those described in [6] and [10], demonstrate the promise of a multiresolution approach to SAR image segmentation. Much remains to be done in order to fully exploit the advantages of multiscale modeling and analysis. In particular, to develop a fully automatic system, one must develop a method to adapt to the nature of the SAR imagery (e.g., to different squint or depression angles, types of vegetation, or the presence of bright tree lines). Possible directions for future work are i) use of more complex models for forest residuals (e.g., product model); ii) adaptation of a full system to combine our approach to segmentation and anomaly enhancement with the discrimination method described in [6]; and iii) applying the methods presented in this paper toward SAR image compression.

B. Anomaly Enhancement

For each SAR image tested, we observed consistently higher peak and average enhanced image values over the target area for the multiscale enhancement technique. In addition, we measured the number of pixel values over the target regions that exceeded a threshold value for various thresholds. In order to ensure a similar probability of false alarm for grass regions,

we normalized each enhanced image such that for grasslike terrain, each statistic would have zero-mean and unit variance. We noted that for each man-made object tested, more pixels exceeded the threshold over the target region in the multiscale enhancement technique than in the CFAR technique, for all thresholds.

REFERENCES

[1] M. Basseville *et al.*, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 766–784, 1992.

[2] P. W. Fieguth, W. C. Karl, A. S. Willsky, and C. Wunsch, "Multiresolution optimal interpolation and statistical analysis of TOPEX/POSEIDON satellite altimetry," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, no. 2, pp. 280–292, 1995.

[3] C. H. Fosgate, "Multiscale segmentation and anomaly enhancement of SAR imagery," Master's thesis, Mass. Inst. Technol., Cambridge, June 1996.

[4] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, no. 6, 1984.

[5] J. C. Henry, "The Lincoln Laboratory 35 GHz airborne polarimetric SAR imaging system," in *Proc. IEEE Nat. Telesyst. Conf.*, Atlanta, GA, Mar. 1991, p. 353.

[6] W. W. Irving, "Multiresolution approach to discriminating targets from clutter in SAR imagery," in *Proc. SPIE Symp.*, Orlando, FL, 1995.

[7] _____, "Multiscale stochastic realization and model identification with applications to large-scale estimation problems," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, Sept. 1995.

[8] D. E. Kreithen, S. D. Halverson, and G. J. Owirka, "Discriminating Targets from Clutter," *Lincoln Lab. J.*, vol. 6, no. 1, 1993.

[9] L. M. Novak, G. J. Owirka, and C. M. Netishen, "Performance of a high-resolution polarimetric sar automatic target recognition system," *Lincoln Lab. J.*, vol. 6, no. 1, 1993.

[10] N. S. Subotic, L. M. Collins, J. D. Gorman, and B. J. Thelen, "A multiresolution approach to target detection in synthetic aperture radar data," in *Proc. Asilomar Conf.*, Monterey CA, 1994.

Charles H. Fosgate received the B.S. degree in electrical engineering from Northeastern University, Boston, MA, in 1984, and the M.S. degree in electrical engineering from Massachusetts Institute of Technology (MIT), Cambridge, MA, in 1996.

From 1994 to 1996, he was a Research Assistant in the Stochastic Systems Group of the Laboratory for Information and Decision Systems at MIT. He joined the technical staff at Hughes Network Systems, MD, in 1996, and is currently working in the area of wireless communications.

Hamid Krim (S'89–M'91) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, and from Northeastern University, Boston, MA.

As a member of the Technical Staff at Bell Laboratories, he has worked in the area of telephone and digital communication systems/subsystems. In 1991, he became an NSF post-doctoral scholar at Foreign Centers of Excellence (LSS Supelec/University of Orsay, Paris, France). He is currently a Research Scientist in the Massachusetts Institute of Technology (MIT) Laboratory for Information and Decision Systems, and also with the Brown–Harvard–MIT Center for Intelligent Control at MIT, Cambridge, performing and supervising research. His main interests are in statistical signal/image processing (estimation/detection) with keen interest in various practical applications.

William W. Irving (S'94–M'95) received the B.S. degree in 1987, the M.S. degree in 1991, and the Ph.D. degree in 1995, all in electrical engineering, and all from The Massachusetts Institute of Technology (MIT), Cambridge.

He was a Staff Associate at Lincoln Laboratory, MIT, where he developed theory and algorithms for automatic detection and recognition of stationary ground targets in high-resolution synthetic aperture radar imagery. Currently, he is a Senior Member of the Technical Staff at Alphatech, Burlington, MA, and is a part-time Lecturer in the Department of Electrical Engineering at MIT. He is developing a theory for SAR ATR performance as part of Wright Laboratory's Moving and Stationary Target and Recognition Program.

William C. Karl (M'91) received the Ph.D. degree in electrical engineering and computer science in 1991 from the Massachusetts Institute of Technology, Cambridge, MA, where he also received the M.S., E.E., and B.S. degrees.

He held the position of Staff Research Scientist with the Brown–Harvard–MIT Center for Intelligent Control Systems and the MIT Laboratory for Information and Decision Systems form 1992 to 1994. He joined the faculty of Boston University in 1995, where is currently Assistant Professor of electrical, computer, and systems engineering. Since January 1996, he has also held a joint appointment in the Department of Biomedical Engineering. His research interests are in the areas of multidimensional and multiscale statistical signal and image processing and estimation, geometric modeling and estimation, and medical signal and image processing.

Dr. Karl is currently an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and a Guest Editor of the *International Journal of Pattern Recognition and Artificial Intelligence*. In 1993, he was organizer and Chair of the geometry and estimation session of the Conference on Information Sciences and Systems at Johns Hopkins University, Baltimore, MD. In 1994, he was on the technical committee for the Workshop on Wavelets in Medicine and Biology, part of the International Conference of the IEEE Engineering in Medicine and Biology Society. In 1996, he was an invited speaker at the JASON Winter Study on the Applications of Wavelets to Military Problems.

Alan S. Willsky (S'70–M'73–SM'82–F'86) received both the B.S. degree and the Ph.D. degree from The Massachusetts Institute of Technology (MIT), Cambridge, in 1969 and 1973, respectively.

From 1974 to 1981, he served as Assistant Director of the MIT Laboratory for Information and Decision Systems. He is also a founder and Member of the board of directors of Alphatech, Inc. He joined the MIT faculty in 1973 and is currently Professor of electrical engineering. His current research interests are in problems involving multidimensional and multiresolution estimation and imaging, discrete-event systems, and the asymptotic analysis of control and estimation systems. He is the author of the research monograph *Digital Signal Processing and Control and Estimation Theory*, and is coauthor of the undergraduate text *Signals and Systems.*

Dr. Willsky has held visiting positions at Imperial College, London, UK, the Université de Paris-Sud, France, and the Institut de Recherche en Informatique et Systèmes Aléatoires, Rennes, France. He has given several plenary lectures at major scientific meetings, including the 20th IEEE Conference on Decision and Control, the 1991 IEEE International Conference on Systems Engineering, the 1991 SIAM Conference on Applied Linear Algebra, and the 1992 Inaugural Workshop for the National Centre for Robust and Adaptive Systems, Canberra, Australia. He was program chairman for the 17th IEEE Conference on Decisoion and Control, has been as Associate Editor of several journals including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and has served as a member of the Board of Governors and Vice President for Technical Affairs of the IEEE Control Systems Society. He was Program Chairman for the 1981 Bilateral Seminar on Control Systems held in China, and was special Guest Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY (Special Issue on Wavelet Transforms and Multiresolution Signal Analysis, 1992). He was awarded the 1979 Alfred Noble Prize by the ASCE and the 1980 Browder J. Thompson Memorial Prize awared by the IEEE for a paper excerpted from his monograph. He is a Distringuished Member of the IEEE Control Systems Society.