

Attributing Scatterer Anisotropy for Model Based ATR

Andrew Kim[†], Sinan Dogan[‡], John Fisher III[†], Randolph Moses[‡], and Alan Willsky[†]

[†] Stochastic Systems Group, Mass. Inst. of Tech.

[‡] Dept. of Elec. Eng., The Ohio State Univ.

ABSTRACT

Scattering from man-made objects in SAR imagery often exhibit aspect and frequency dependences which are not well modeled by standard SAR imaging techniques. If ignored, these deviations may reduce recognition performance due to model mismatch, but when appropriately accounted for, these deviations can be exploited as attributes to better distinguish scatterers and their respective targets. Chiang and Moses¹ developed an ATR system that allows the study of performance under various scatterer attributions. Kim *et. al.*² examined a nonparametric approach for exploiting non-ideal scattering using a multi-resolution sub-aperture representation. Both of these works are extended here to examine the effect of anisotropic scattering attribution for model-based ATR. In particular, predicted and extracted peak scatterers are attributed with a discrete anisotropy feature. This feature can be obtained in a computationally efficient manner by performing a set of generalized log-likelihood ratio (GLLR) tests over a pyramidal sub-aperture representation. Furthermore, an approximate probabilistic characterization of the feature set allows for a natural inclusion into the approach of Chiang and Moses which will be used to evaluate the benefit of our attribution to the X-band MSTAR data and infer the phenomenology behind anisotropic scattering.

Keywords: SAR, ATR, classification, anisotropic scattering, multiscale, model-based

1. INTRODUCTION

Scatterers composing a target in SAR imagery often exhibit nonideal scattering in the form of aspect and frequency dependences. Standard SAR image formation ignores this variability resulting in unstable scintillating reflectivity estimates complicating the recognition problem. These deviations from the ideal point scattering model should not be viewed as a hindrance and approximated away, but instead, they should be seen as an attribute which can be used to distinguish scatterers and thus their respective targets.

This paper is an extension of two separate works presented at Aerosense 1999. Chiang and Moses¹ presented a full ATR system which allowed performance comparisons to be made between systems based on different feature attributes. It was used to demonstrate the improvement in ATR performance achieved by using models based on the Geometric Theory of Diffraction (GTD) with synthetic data. Kim *et. al.*² presented a classification technique that utilizes aspect dependence by learning these dependences in a nonparametric fashion on a multi-resolution pyramid of sub-apertures.

The work presented here uses the same sub-aperture structure introduced in by Kim *et. al.*, but simplifies the information conveyed into a single scalar parameter characterizing the azimuthal concentration of unimodal scattering. Motivated by canonical scattering models, we conjecture that knowledge of the azimuthal duration of a scatterer can be used to infer properties of its geometry. Scattering models such as the physical optics model or the Geometric Theory of Diffraction (GTD) predict that for many physically large scatterers there is an inverse relation between the size of the scatterer and the duration of its response in azimuth. Thus, knowing the anisotropy of a scatterer allows one to infer properties of the physical structure of the object under investigation thereby aiding the classification procedure. Incorporating the resulting anisotropy attribution into the feature based classifier and applying it to real SAR data allows us to study the utility of and phenomenology behind azimuthal anisotropy.

The remainder of this paper is organized as follows. Section 2 presents the multi-resolution sub-aperture pyramid used to represent the SAR data. Section 3 then describes the set of hypotheses that we will consider and develops the hypothesis tests on the sub-aperture pyramid. Section 4 describes the matching algorithm that we use to evaluate our anisotropy attribution on collected SAR data. Section 5 presents experimental results demonstrating the utility of anisotropy attribution and discusses its underlying phenomenology. The paper concludes with a summary and discussion in Section 6.

Further author information: Send correspondence to Andrew Kim (E-mail: ajkim@mit.edu; WWW: <http://ssg.mit.edu/~ajkim/>).

2. SUB-APERTURE ANALYSIS

The foundation of our analysis is the sub-aperture pyramid which we present in this section. This structure is motivated by the scattering physics involved in SAR and presents information in a way that allows for simple and intuitively reasonable hypothesis tests. Because of the linear structure of the aperture, we will associate it with an interval of the real line throughout this paper. In particular, the full-aperture will be denoted by the interval $[0, 1)$.

2.1. Definition

The intuitive idea of the sub-aperture pyramid is to generate an over-complete covering of the full-aperture with sub-apertures that can be arranged on a pyramidal structure. These sub-apertures will be used to represent both our set of candidate hypotheses and to form our reflectivity estimates. The prototypical sub-aperture covering that we will use throughout this paper is the half-overlapping half-aperture pyramid shown in the lower portion of Figure 1.

We take as a sub-aperture pyramid a set S of sub-apertures with the following structure. The set S is partitioned into smaller sets S_m which correspond to a particular degree of anisotropy. For reasons which will become apparent, we associate m with scale. S_0 refers to the set consisting of the largest sub-apertures, and S_M refers to the set of the smallest sub-apertures. A second subscript on S denotes a specific sub-aperture at the given scale. To obtain the necessary structure on the sub-aperture pyramid for what follows later, the following conditions are imposed on S :

- (S1) $\forall S_{m,i}, S_{m,i} = [a, b)$ for some $0 \leq a < b \leq 1$,
- (S2) $\forall S_{m,i}$ with $m \geq 1$, $\exists S_{m-1,j}$ such that $S_{m,i} \subset S_{m-1,j}$, and
- (S3) $\forall S_{m,i}, \exists$ a partition $\mathcal{P}(m, i) \subset S_M$ of $S_{m,i}$.

The first condition simply restricts the sub-apertures to be a single connected interval. This is motivated by our search for concentrated unimodal scattering in azimuth. The second condition asserts that each sub-aperture, except those in S_0 , has a parent which contains it. This allows us to construct a telescopic hypothesis test on a tree which will not only afford computationally efficiency but also robustness. The third condition requires the existence of a partition of each sub-aperture by coarsest scale sub-apertures. This will allow for the set of measurements given by S_M to form a sufficient statistic for all the measurements in S . Herein, the term *sub-aperture pyramid* will always refer to one satisfying conditions (S1)-(S3).

Any sub-aperture can be used to form a SAR image. The images formed with smaller values of m have a finer cross-range imaging resolution because of their larger apertures. This is our motivation for associating scale with m . Each sub-aperture $S_{m,i}$ generates an associated reflectivity estimate $q_{m,i}$. The collection of reflectivity estimates from the sub-apertures in S_M is denoted as \mathbf{q}_M . The measured reflectivity $q_{m,i}$ is not normalized with respect to the aperture length, i.e.

$$q_{m,i} = \int_{S_{m,i}} a(s) ds \quad (1)$$

where $a(s)$ is the azimuthal response of the scatterer*. Thus, when interested in the normalized reflectivity estimate, one should divided $q_{m,i}$ by the sub-aperture length $L_{m,i} = \lambda(S_{m,i})$, where λ denotes Lebesgue measure.

2.2. Interpretation and Motivation

Different types and sizes of scatterers will yield different aspect dependencies. The motivation for using the sub-aperture pyramid is that it is expected to reveal distinguishing aspect dependences in the scattering. For example, a small metal sphere will have a strong response in all directions and thus produce a strong reflectivity estimate from each of the sub-apertures. However, as depicted in Figure 1, a flat plate produces a significantly stronger response when oriented broadside with respect to the radar as compared to off-broadside orientation. Thus, the reflectivity estimates will vary across the sub-apertures with the largest estimate coming from the sub-aperture oriented broadside to the plate. Furthermore, because various sized sub-apertures are used, the duration of the broadside flash is captured in this representation. In particular, while the sub-aperture is contained within the main-lobe of the response, the reflectivity estimate will be consistently large, but as the sub-aperture is expanded, the additional energy received will be relatively insignificant and result in a lower reflectivity estimate when normalized

*By azimuthal response, we mean the 1-D cross-range uncompressed signal for a given down-range location. The signal $a(s)$ is assumed to have already been appropriately demodulated to have zero phase modulation for the inspected cross-range location.

with respect to the sub-aperture length. While this framework will capture general degrees of anisotropy, it is not overly-sensitive to azimuthal dependencies in that slight deviations on the scatterer geometry are modeled. This relieves of the burdens associated with such models that requires an excessive number of parameters.

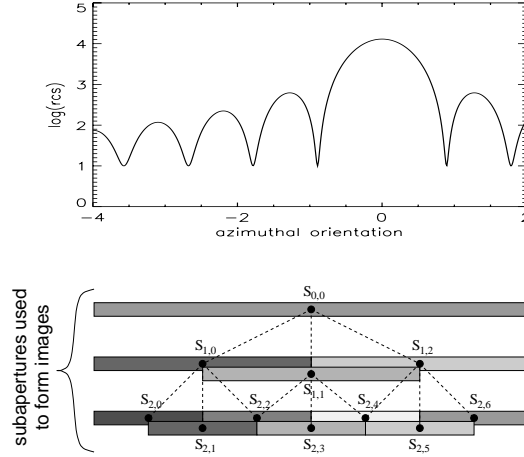


Figure 1. The response of a .5m x .5m flat plate and a depiction of the reflectivity estimate for each of the sub-apertures. Lighter shaded sub-apertures convey larger reflectivity estimates.

A slightly different viewpoint of this sub-aperture feature set comes from considering the cross-range resolution versus azimuthal[†] resolution trade-off. Recall that the cross-range resolution is inversely proportional to the aperture length. Thus, at lower levels of the sub-aperture pyramid, spatial resolution has been exchanged for azimuthal resolution, i.e. the ability to better observe anisotropic phenomena. This is the classic time–frequency resolution tradeoff in Fourier analysis, and each level of the pyramid represents the data under a particular cross-range–azimuth resolution. The presence of multiple resolutions is attractive because we expect the best representation for different objects to be nonunique as the importance between resolution in the two domains is balanced.

3. ANISOTROPIC SCATTERING MODELS

Having presented the sub-aperture pyramid, we now proceed to formulate our hypothesis testing problem for anisotropy. The hypotheses will be drawn directly from the sub-aperture pyramid. Two models will be presented here. The first is a simple single scatterer model with an intuitive sufficient statistic. This test however is susceptible to the influence of neighboring scatterers. This motivates the second model which explicitly accounts for the contributions from neighbors. The tests presented in this section are for a fixed scattering location which we assume to be specified. These locations could come from a peak extraction process or a pre-specified grid of points to produce an image of anisotropy.

3.1. Single Scatterer Model

Each sub-aperture $S_{m,i}$ defines an associated scattering hypothesis $H_{m,i}$ over the aperture s via

$$H_{m,i} : a(s) = 1_{S_{m,i}}(s) \quad (2)$$

where $1_A(\cdot)$ denotes the indicator function over the set A . Thus, our hypotheses correspond to scattering responses that are uniform over the sub-aperture in question and zero elsewhere. Naturally, this is an idealization for anisotropic scattering, but because we are only looking for a general characterization of anisotropy, it will serve our purposes here. Although we will call this a test of anisotropy, the ideal isotropic scattering hypothesis is included in our hypothesis set if the full-aperture is included in the sub-aperture pyramid. The set of all possible hypotheses associated with the sub-aperture pyramid will be denoted as \mathcal{H} .

[†]For clarity, we will use the term “cross-range” when referring to the cross-range in the spatial (image) domain, and we will exclusively use the term “azimuthal” when referring to the corresponding dimension in the sensor domain.

A reasonable choice of features to test these hypotheses would be all the measured sub-aperture reflectivities $\{q_{m,i}\}$. From the definition of the $q_{m,i}$ in Eq. (1) and partition property (S3), it is sufficient to consider the subset $\mathbf{q}_M \subset \{q_{m,i}\}$ since all sub-aperture reflectivities $q_{m,i}$ can be computed from \mathbf{q}_M by summing all the $\mathbf{q}_{M,j}$ which form a partition of $S_{m,i}$. Thus, we will take \mathbf{q}_M as our feature vector. The value of this feature vector under hypothesis $H_{m,i}$ is $\mathbf{b}(m,i)$ whose j^{th} element is given by

$$\begin{aligned} b(m,i)_j &= \int_{S_{M,j}} 1_{S_{m,i}}(s) ds \\ &= \lambda(S_{M,j} \cap S_{m,i}), \end{aligned} \quad (3)$$

i.e. it is the portion of the response one expects to see over the j^{th} sub-aperture at scale M . We now define our scattering model conditioned on anisotropy hypothesis $H_{m,i}$ as the signal plus noise model,

$$q_{M,j} = \int_{S_{M,j}} A 1_{S_{m,i}}(s) + \eta(s) ds, \quad (4)$$

where A is the scattering amplitude of the signal and $\eta(s)$ is circularly complex white Gaussian noise with spectral density σ^2 . This leads to the model

$$\mathbf{q}_M = A\mathbf{b}(m,i) + \boldsymbol{\epsilon}, \text{ with } \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, 2\sigma^2\Lambda), \quad (5)$$

where Λ is the noise covariance structure inherited from the sub-aperture pyramid. The noise in the measured reflectivities in Eq. (5) are characterized as zero-mean circularly complex Gaussians with covariances dictated by the amount of sub-aperture overlap. The elements of its covariance matrix Λ are given by $[\Lambda]_{i,j} = \lambda(S_{M,i} \cap S_{M,j})$, which for the half-overlapping half-aperture pyramid in Figure 1 is

$$\Lambda = \frac{1}{2^M} \begin{bmatrix} 1 & .5 & & & \\ .5 & \ddots & \ddots & & \\ & \ddots & \ddots & .5 & \\ & & & .5 & 1 \end{bmatrix}.$$

To classify the anisotropy of a scatterer from our vector of sub-aperture measurements \mathbf{q}_M , we apply a log-likelihood ratio (LLR) test to the model in Eq. (5) where each log-likelihood is compared to the full-aperture hypothesis. Because there is the unknown reflectivity parameter A , we actually use a generalized LLR (GLLR) test where for each hypothesis, we take A to be its maximum likelihood (ML) estimate under that hypothesis. Thus, for $H_{m,i}$, we take $\hat{A} = q_{m,i}/L_{m,i}$. This produces the GLLR

$$\ell_{m,i} = \frac{1}{4\sigma^2} \left[\frac{1}{L_{m,i}} |q_{m,i}|^2 - |q_{0,0}|^2 \right]. \quad (6)$$

Thus, the most likely sub-aperture in this case is the one whose average energy is largest. We note the similarity here to the approach taken by Chaney *et. al.*³ in which they replace, within the image, the standard reflectivity estimate with the maximum sub-aperture reflectivity estimate $q_{m,i}/L_{m,i}$, thus using normalized reflectivity (instead of normalized energy) as their criterion for choosing anisotropy. Their approach however is based on intuitive arguments and not a derived statistic.

Though simple and intuitive, the GLLR in Eq. (6) is susceptible to the effects of close proximity neighboring scatterers which are not included in our model in Eq. (4). Recall that the images formed by the smaller sub-apertures have a coarser resolution. Thus, if a scatterer were located outside the finest resolution cell but within a coarser resolution cell, then the finest resolution reflectivities $q_{0,j}$ would be 0, but $q_{m,i}$ would be large if the resolution cell associated with scale m included the scatterer. We illustrate this with the example shown in Figure 2. Here a scatterer with amplitude A is located outside the finest resolution cell of size δ but is contained within the coarser resolution cell of size 2δ associated with half-aperture estimates. If this scatterer is isotropic, then its response is the complex exponential illustrated. Integrating over the full aperture gives a 0 reflectivity estimate as expected, but integrating over a half aperture produces a normalized reflectivity with magnitude $A/\sqrt{2}$. Eq. (6) would then classify the center of the resolution cell as anisotropic, even though no scatterer is present.

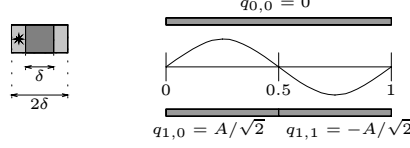


Figure 2. Illustration of a scatterer not in the finest resolution cell that can produce a false anisotropy classification in Eq. (6).

The problem above is a consequence of not modeling the influence of neighboring scatterers. One of the ways in which the neighboring scatterer manifests itself is through its corruption of the estimated reflectivity as the size of the resolution cell varies. Choosing $\hat{A} = q_{m,i}$ is the maximum likelihood reflectivity estimate for the resolution cell associated with $S_{m,i}$. Alternatively, we may instead choose the best reflectivity estimate constrained to lie in the finest resolution cell which is the full aperture estimate $q_{0,0}$. Choosing $\hat{A} = q_{0,0}$ for all hypotheses can be shown to produce the GLLR statistic

$$\ell_{m,i} = \frac{1}{4\sigma^2} \left[\frac{1}{L_{m,i}} |q_{m,i}|^2 - \frac{1}{L_{m,i}} |q_{0,0} - q_{m,i}|^2 - |q_{0,0}|^2 \right]. \quad (7)$$

This statistic is identical to that in Eq. (6) except for the extra term comparing the reflectivity estimates $q_{0,0}$ and $q_{m,i}$. Recall that Eq. (6) compared the average energy in a sub-aperture to the full-aperture. This new GLLR accounts for the average energy outside the current sub-aperture as well. Viewed differently, under our scattering model in Eq. (4), the values of $L_{m,i}q_{0,0}$ and $q_{m,i}$ should simply be noisy perturbations of each other. The new term penalizes when this is not the case. Under the example in Figure 2, since the contribution of each of the half-apertures would be the same, the GLLR is equal to zero for all hypotheses, which is reasonable, since there is no underlying scattering at the focused location.

3.2. Multiple Scatterer Model

The modification in Eq. (7) addresses the problem when a neighboring scatterer is isotropic, however, when the neighboring scatterer is anisotropic, problems such as that illustrated by the example in Figure 2 can still arise as its contribution will not integrate out over the full-aperture estimate. To alleviate this problem we generalize the model in Eq. (4) to account for multiple scatterers.

First, we must generalize the sub-aperture scattering model to account for the modulations produced by neighboring scatterer. Again, we assume that the sub-apertures have been formed while being focused on cross-range location y_0 . The possibility of other scatterers are considered at discrete locations $y_0 + k\Delta_p$, where Δ_p specifies a cross-range sampling resolution. To model the effect of a scatterer at location $y_0 + k\Delta_p$ on the measurements at location y_0 , we simply modulate the focused response $1_H(\cdot)$ to account for the shift in the image domain. The observed effect over the sub-apertures is then given by

$$b^k(m, i)_j = \int_{S_{M,j}} e^{j2\pi \frac{k\Delta_p}{\Delta_r} s} 1_{H_{m,i}}(s) ds \quad (8)$$

where Δ_r is the null-to-null resolution associated with the full aperture. Incorporation of the neighboring scatterers is now modeled via superposition, i.e.

$$\mathbf{q}_M = \sum_k A_k \mathbf{b}^k(H_k) + \boldsymbol{\epsilon} \quad (9)$$

where the noise model is the same as in Eq. (5). The summation over k in Eq. (9) should be over a range that at least includes all scatterers contained in the largest resolution cell. Thus, if L_* is the smallest aperture length, then we need to consider $k \in \{-K, \dots, K\}$ where

$$K\Delta_p \geq \frac{\Delta_r}{2L_*} = \text{size of coarsest resolution cell} \\ K = \left\lceil \frac{\Delta_r}{2L_*\Delta_p} \right\rceil \quad (10)$$

and $k = 0$ corresponds to the resolution cell under investigation. Rewriting Eq. (9) as a matrix equation we get

$$\begin{aligned} \mathbf{q}_M &= [\mathbf{b}^{-K}(H_{-K}) \quad | \quad \cdots \quad | \quad \mathbf{b}^K(H_K)] \begin{bmatrix} A_{-K} \\ \vdots \\ A_K \end{bmatrix} + \boldsymbol{\epsilon} \\ &= B\mathbf{A} + \boldsymbol{\epsilon} \end{aligned} \quad (11)$$

where B and \mathbf{A} are appropriately defined. Thus, we can use weighted least squares (WLS) to implicitly estimate the values of the interfering A_k and account for their contribution to \mathbf{q}_M . In order to have our least squared error minimization correspond to ML, we need to have our inner product effectively whiten the noise. This is accomplished by using the inner product weighted by the inverse of the noise covariance, i.e. $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{2\sigma^2} \mathbf{u}^T \Lambda^{-1} \mathbf{v}$. We estimate the hypothesis as the one which minimizes the norm of $\boldsymbol{\epsilon}$ when the WLS estimate of \mathbf{A} is used. The ML estimate for \mathbf{A} is obtained from this model as

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} \{ \|\boldsymbol{\epsilon}\|_{\Lambda^{-1}}^2 \} = (B' \Lambda^{-1} B)^{-1} B' \Lambda^{-1} \mathbf{q}_M \quad (12)$$

which simplifies to Eq. (6) in the case of limiting the model order to $K = 0$.

Note that the hypotheses are now in a higher dimensional space. In particular, the hypothesis set is now \mathcal{H}^{2K+1} since each location k has a scatterer associated with it. For the initial work presented in this paper, we will restrict ourselves to hypotheses where only the $k = 0$ (the focused resolution cell) is allowed to vary. Thus, our hypothesis space is effectively still \mathcal{H} . We note that this constraint will be relaxed in the future in order to appropriately take into account the scattering of neighboring scatterers.

To classify the anisotropy of the $k = 0$ scatterer, we use ML, i.e. we choose the sub-aperture hypothesis which minimizes the weighted norm of $\boldsymbol{\epsilon}$ in Eq. (9)

$$\begin{aligned} \|\boldsymbol{\epsilon}\|_{\Lambda^{-1}}^2 &= \|\mathbf{q}_M - B\hat{\mathbf{A}}\|_{\Lambda^{-1}}^2 \\ &= \frac{1}{2\sigma^2} \mathbf{q}_M' [\Lambda^{-1} - \Lambda^{-1} B (B' \Lambda^{-1} B)^{-1} B' \Lambda^{-1}] \mathbf{q}_M \end{aligned} \quad (13)$$

where the weighting matrix in the last line is independent of the data and thus can be precomputed for each of the candidate hypotheses.

The model in Eq. (9) is too unconstrained for the test given by the minimization in Eq. (13) to work. In particular, if one chooses K sufficiently large to account for all scatterers in the coarsest resolution cell, then the model order is greater than the number of sub-aperture measurements and $\boldsymbol{\epsilon}$ can be made zero for all hypothesis of $k = 0$. Thus, we need to regularize the model. In order for the error $\boldsymbol{\epsilon}$ to be made small under the incorrect model, many of the values in \mathbf{A} generally have to be made unreasonably large and result in an unrealistic scenario. Thus, we impose a 2-norm regularization penalty in the estimation of \mathbf{A} . In particular, instead of minimizing the weighted squared error to estimate \mathbf{A} , we take

$$\begin{aligned} \hat{\mathbf{A}} &= \arg \min_{\mathbf{A}} \{ \|\boldsymbol{\epsilon}\|_{\Lambda^{-1}}^2 + \gamma \mathbf{A}' R \mathbf{A} \} \\ &= (B' \Lambda^{-1} B + \gamma R)^{-1} B' \Lambda^{-1} \mathbf{q}_M \\ &= P \mathbf{q}_M \end{aligned} \quad (14)$$

where P is defined accordingly, γ is the regularization parameter, and R is the regularization matrix that penalizes the energy in all A_k other than $k = 0$, i.e.

$$R = \mathbf{I} - \mathbf{e}_K \mathbf{e}_K' = \text{diag}(1, \dots, 1, 0, 1, \dots, 1).$$

This produces the following value for the weighted error norm as

$$\begin{aligned} \|\boldsymbol{\epsilon}\|_{\Lambda^{-1}}^2 &= \|\mathbf{q}_M - B\hat{\mathbf{A}}\|_{\Lambda^{-1}}^2 \\ &= \frac{1}{2\sigma^2} \mathbf{q}_M' [\Lambda^{-1} - 2\Lambda^{-1} B P + P' B' \Lambda^{-1} B P] \mathbf{q}_M \end{aligned} \quad (15)$$

which can be used for our hypothesis test.

3.3. Telescopic Testing

The sub-aperture pyramid which we use to form our measurements and base our hypotheses is convenient not only for providing anisotropy information, but also for providing an efficient means of performing the hypothesis tests. We can obtain an efficient approximation to the test by evaluating only a small subset of the candidate hypotheses. Due to the nested structure of the sub-apertures (condition (S2) in Section 2), we can perform the tests in a telescopic fashion by traversing down the tree of sub-apertures as depicted in Figure 3. We expect the likelihoods to increase as the hypothesized sub-apertures “shrink down to” the correct sub-aperture, and then to decrease as the hypothesized sub-apertures “shrink beyond” the correct sub-aperture. This motivates performing the hypothesis test in the following manner:

- step 1) Start with the set of largest sub-aperture(s) at scale $m = 0$. Find the most likely hypothesis at that scale and denote it as H_{0,i^*} .
- step 2) Consider those hypotheses at scale $n = m + 1$ for which $S_{n,j} \subset S_{m,i^*}$. Find the one which has the highest likelihood and denote it as H_{n,j^*} .
- step 3) If the parent is more likely (i.e. $\gamma_{m,i^*} > \gamma_{n,j^*}$), then stop and return H_{m,i^*} as the estimated hypothesis.
- step 4) If $m = M$, we are at the bottom of the tree so stop and return H_{n,j^*} as the estimated hypothesis. Otherwise, set $m = n$ and $i^* = j^*$ and goto step 2.

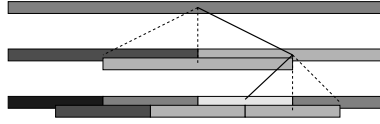


Figure 3. Illustration of how the anisotropy testing can be done in a decision directed fashion by starting with the largest aperture and at each scale, inspecting only the children of the most likely sub-aperture.

The intuition described above can be justified under the single scattering models. In particular, consider the expected value of $\ell_{n,j}$ when the true hypothesis is $H_{m,i}$ and the proportion of overlap between $S_{n,j}$ and $S_{m,i}$ is given by $\alpha = \frac{\lambda(S_{n,j} \cap S_{m,i})}{\lambda(S_{m,i})}$. For Eq. (6) in which $\hat{A} = q_{m,j}$ is used, the expected value is

$$\begin{aligned} \mathbb{E}[\ell_{n,j} \mid H_{m,i}] &= \mathbb{E}\left[\frac{1}{L_{n,j}} |q_{n,j}|^2 - |q_{0,0}|^2 \mid H_{m,i}\right] \\ &= \left(\frac{\alpha^2}{L_{n,j}} - 1\right) |A|^2 + (2L_{n,j} + 1)\sigma^2 \\ &\approx \left(\frac{\alpha^2}{L_{n,j}} - 1\right) |A|^2 \end{aligned} \quad (16)$$

where the approximation is for high SNR. From this, we see the intuitive behavior described above. When $S_{n,j}$ contains $S_{m,i}$, the overlap is $\alpha = 1$ and the GLLR increases with decreasing $L_{n,j}$. When the sub-aperture becomes too small, i.e. $S_{n,j} \subset S_{m,i}$, the overlap is $\alpha = L_{n,j}/L_{m,i}$, and thus the expected GLLR decreases when the sub-aperture becomes too small.

For Eq. (7) in which $\hat{A} = q_{0,0}$, the expected value of the GLLR is

$$\begin{aligned} \mathbb{E}[\ell_{n,j} \mid H_{m,i}] &= \mathbb{E}\left[\frac{1}{L_{n,j}} |q_{n,j}|^2 - \frac{1}{L_{n,j}} |q_{0,0} - q_{n,j}|^2 - |q_{0,0}|^2 \mid H_{m,i}\right] \\ &= 2\left(\frac{\alpha - 1}{L_{n,j}}\right) |A|^2 + \left(\frac{1}{L_{n,j}} - 1\right) (|A|^2 - 2\sigma^2). \end{aligned} \quad (17)$$

Again, we see the intuitive behavior described above. In particular, when $S_{n,j}$ contains $S_{m,i}$, the overlap $\alpha = 1$ and the GLLR increases with decreasing $L_{n,j}$. When the sub-aperture becomes too small, i.e. $S_{n,j} \subset S_{m,i}$, the overlap is $\alpha = L_{n,j}/L_{m,i}$, and the GLLR can be written as

$$\mathbb{E}[\ell_{n,j} \mid H_{m,i}] = 2\frac{|A|^2}{L_{m,i}} - \frac{|A|^2 + 2\sigma^2}{L_{n,j}} - |A|^2 + 2\sigma^2$$

which is a decreasing function of $L_{n,j}$ for strong scatterers, i.e. $|A|^2 \gg 2\sigma^2$.

Because the difficult form of our regularized multiple scatterer log-likelihood given in Eq. (15), we have not yet shown the same pattern for this extended test, however intuition leads us to believe it holds here too.

3.4. Boxcar Model Deviations

The hypothesis set defined in Eq. (2) are simplified models to which real scatterers will not exactly correspond. One may question whether deviations from this model may drastically effect our hypothesis tests. For example, if the scattering has a $\text{sinc}(\cdot)$ like dependence in azimuth, then the sidelobes will have a large response for a strong scatter and cannot be well modeled as background noise. To address this issue, we incorporate deviations from the boxcar response into our model. In particular, we start by assuming the underlying scattering pattern has been perturbed by white Gaussian noise. For the single scatterer model, this changes the model response in Eq. (3) to

$$\tilde{b}(m, i)_j = \int_{S_{M,j}} 1_{S_{m,i}}(s) + \nu(s) ds$$

where $\nu(s)$ is a white Gaussian process with spectral density ρ^2 and is independent of the measurement noise $\eta(s)$. Thus, our modeled measurement vector is now a random vector characterized as

$$\tilde{\mathbf{b}}(m, i) \sim \mathcal{N}(\mathbf{b}(m, i), 2\rho^2\Lambda).$$

This results in the new measurement model as

$$\mathbf{q}_M = A\mathbf{b}(m, i) + \mathbf{w}, \text{ with } \mathbf{w} \sim \mathcal{N}(\mathbf{0}, 2(|A|^2\rho^2 + \sigma^2)\Lambda).$$

Thus, we have essentially the same model as in Eq. (5) except that the variance of the noise now depends affinely on the square-magnitude of the underlying scatterer. The effect in Eqs. (6) and (7) is a simple scaling of the GLLR's of *all* the hypotheses. For the multiple scatterer model, the extension is similar and has the same result.

4. BAYES CLASSIFICATION

With a method for anisotropy attribution in hand, we describe in this section a classifier based on Bayesian probability theory. This matcher allows us to evaluate the utility and explore the phenomenology of anisotropy in SAR by incorporating the labeled anisotropy of scatterers into the feature set. A more thorough description of this classifier can be found in the paper by Chiang and Moses.¹

4.1. Classification Problem Statement

The Bayes matching problem is given as follows. At the input to the classifier stage, we are given a set of n feature vectors $Y = [Y_1, Y_2, \dots, Y_n]^T$ extracted from a measurement, and for each candidate hypothesis[‡] $H \in \mathcal{H}$ we are given a set of m predicted feature vectors $X = [X_1, X_2, \dots, X_m]^T$ (where m may vary with H). We wish to find the hypothesis whose posterior likelihood of the observed features, Y , is maximum. From Bayes' rule, we have

$$P(H|Y) = \frac{f(Y|H, n)P(n|H)P(H)}{f(Y|n)P(n)}.$$

Since the denominator does not depend on hypothesis H , the MAP decision is found by maximizing the numerator $f(Y|H, n)P(n|H)P(H)$ over $H \in \mathcal{H}$. In this paper, we assume the priors $P(H)$ and $P(n|H)$ are uniform, so we need only compute $f(Y|H, n)$.

We incorporate uncertainty in both the predicted and extracted feature sets, and assume the predict and extract uncertainties are conditionally independent. This gives⁴

$$f(Y|H, n) = \int f(Y|\hat{X}, H, n)f(\hat{X}|H, n)d\hat{X} \quad (18)$$

where $f(\hat{X}|H, n)$ models the feature prediction uncertainty, and $f(Y|\hat{X}, H, n)$ models feature extraction uncertainty.

[‡]The hypotheses in this section correspond to possible vehicle classifications. They are not the anisotropic hypotheses used in Section 3.

The computation of $f(Y|\hat{X}, H, n)$ requires a correspondence between the elements of Y and \hat{X} , or equivalently between Y and X . We consider two correspondence mappings. The first is a probabilistic many-to-many map, in which we assume that

$$f(Y|X) = f(Y|H) = \prod_{j=1}^n f(Y_j|X, H) = \prod_{j=1}^n \left[B f_{FA}(Y_j) + \sum_{i=1}^m D_i(H) f(Y_j|X_i, H) \right] \quad (19)$$

where λ is the average number of false alarms features present, f_{FA} models false alarm probability of a particular feature vector, $P_k(H)$ is the detection probability of the i^{th} predicted feature under hypothesis H , $B = \lambda / [\lambda + \sum_{k=1}^m P_k(H)]$ is the probability that an extracted feature is a false alarm, and $D_i(H) = (1-B)P_i(H) / [\sum_{k=1}^m P_k(H)]$ is the probability that an extracted feature comes from the i^{th} predicted feature.

The second mapping considered is a deterministic one-to-one map, in which the correspondence is assumed to be a deterministic nuisance parameter and the match score is maximized over the correspondence. In this case the likelihood score is given by

$$\begin{aligned} f(Y|\Gamma, H, n) = & \left\{ P(n_{FA} \text{ false alarms}) \prod_{\{j:\Gamma_j=0\}} f_{FA}(Y_j) \right\} \\ & \cdot \left\{ \prod_{\{j:\Gamma_j=i>0\}} P_i(H) \right\} \cdot \left\{ f(Y_j|\Gamma_j=i, H, n) \prod_{\{i:\Gamma_j \neq i, \forall j\}} (1 - P_i(H)) \right\} \end{aligned} \quad (20)$$

where Γ defines the feature correspondences, including the n_{FA} extracted features that correspond to no predicted features (denoted $\{j : \Gamma_j = 0\}$) and the predicted features that correspond to no extracted features (denoted $\{i : \Gamma_j \neq i, \forall j\}$).

For the case that $P(n_{FA} \text{ false alarms})$ obeys an exponential rule $P(n_{FA} \text{ false alarms}) = ce^{-\beta n_{FA}}$ for constants c and β , the search for the correspondence that maximized the above likelihood can be efficiently implemented.⁴

To implement either Eq. (19) or (20), we require a probability model for $f(Y|\hat{X}, \Gamma, H, n)$. We assume that the uncertainties of the X_i are conditionally independent given H , and that the uncertainties of the Y_j are conditionally independent given H , X , and n . This yields

$$f(Y|\Gamma, H, n) = \prod_{j=1}^n f(Y_j|\Gamma, H, n).$$

Each extracted feature Y_j either corresponds to a predicted feature or is a false alarm. If Y_j is a false alarm, we assign $\Gamma_j = 0$, and we model the feature attribute as a random vector with probability density function

$$f(Y_j|\Gamma_j = 0, H, n) = f_{FA}(Y_j).$$

If Y_j corresponds to a predicted feature X_i , we write $\Gamma_j = i$ (for $i > 0$) and compute the feature match score from Eq. (18). In particular, from Eq. (18) it follows that for $i > 0$,

$$f(Y_j|\Gamma_j = i, H, n) = \int f(Y_j|\hat{X}_i, H, n) f(\hat{X}_i|X_i, H) d\hat{X}_i. \quad (21)$$

For the special case of Gaussian uncertainties, we have $f(Y_j|\hat{X}_i, H, n) \sim \mathcal{N}(\hat{X}_i, \Sigma_e)$, and $f(\hat{X}_i|X_i, H, n) \sim \mathcal{N}(X_i, \Sigma_p)$, so from Eq. (21) we obtain

$$f(Y_j|\Gamma_j = i, H, n) = f(Y_j|X_i, H, n) \sim \mathcal{N}(X_i, \Sigma_p + \Sigma_e). \quad (22)$$

Similarly, for features whose attributes are discrete-valued, the likelihood is the sum

$$P(Y_j|X_i, H, n) = \sum_{\hat{X}_i} P(Y_j|\hat{X}_i, H, n) P(\hat{X}_i|X_i, H, n). \quad (23)$$

5. RESULTS

Public release MSTAR data is used for the results presented here. These images have a resolution of $0.3m$ in both range and cross-range. The transmitted signal had a bandwidth of $0.591GHz$ and a center frequency of $9.60GHz$.

All of the results in this section are based on the three-level half-overlapping half-aperture pyramid depicted in Figure 1. The multiple scatterer model will be used to characterize anisotropy. The number of neighboring scatterers considered is set by $K = 6$. The value of the regularization parameter on neighboring reflectivities is set to $\gamma = 0.5$. We incorporate a bias in our anisotropy test towards full-aperture scattering. In particular, to be declared anisotropic, an anisotropic likelihood must be at least twice the full-aperture likelihood. The purpose of this higher threshold is to aid in protecting against the effects of neighboring scatterers whose reflectivities may not have been estimated exactly, and will thus induce a modulation across the aperture which can be mistaken as anisotropy.

This paper is based on the idea of detecting anisotropic scattering in SAR imagery. To illustrate the anisotropy assignments made by our model, we show in Figure 4 the results for a BMP2 (serial number c21) at $0^\circ, 20^\circ, 40^\circ, 60^\circ$, and 80° azimuths with a 17° depression. Even though the aperture associated with this data set is relatively small (about 3°), we note that we are still able to detect anisotropic scattering. In particular, it usually appears to be associated with the turret, barrel, or leading edge of the tank. We make particular note of the classifications at the 0° azimuth. Here we see many clutter pixels being classified as anisotropic. The cause of this error is the unmodeled behavior of neighboring anisotropic scatterers. Recall that in our current formulation, we only consider the possibility of neighboring scatterers which are isotropic. However, the front edge of the tank generates a strong anisotropic response which is not accurately captured in our current model. Extending our model to account for such scattering should alleviate problems such as this.

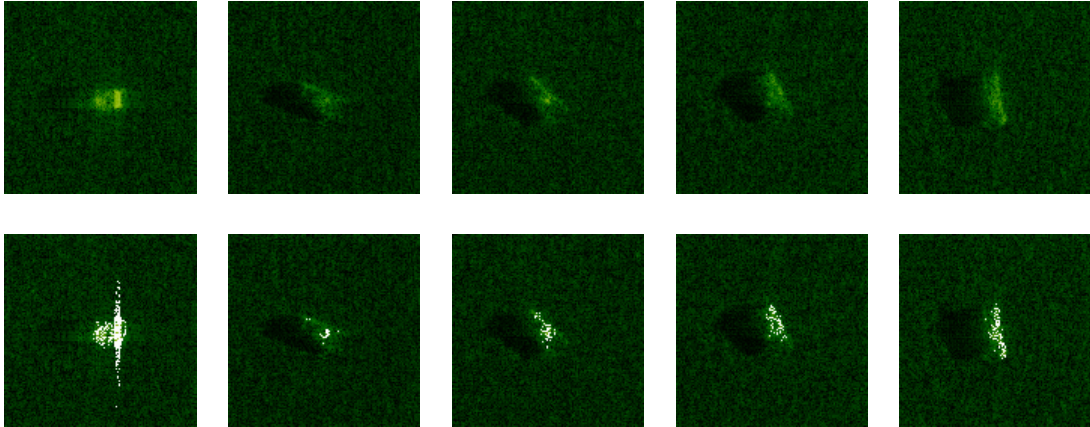


Figure 4. Anisotropy characterization of several instances of a BMP2. Top row: Log-magnitude reflectivity image. Bottom row: Log-magnitude reflectivity image with pixel locations declared to be anisotropic masked out in white.

The images in Figure 4 show that scatterers are being classified as anisotropic, however it does not convey how useful that information is in characterizing targets or explaining phenomenology. To address these issues, we consider empirical confusion matrices of anisotropy based upon the MSTAR data set. In particular, we consider the data set composed of the following vehicles: 2S1 (b01), BMP2 (c21), BRDM2 (E-71), D7 (92v13015), T72 (132), ZIL131 (E12), and ZSU23-4 (d08) where truth is taken to be the empirical results from 17° depression data and test data is taken from the 15° depression data at the same azimuth as the truth. For each pair of training and testing images at the same azimuth, a set of peaks are extracted from each image and a correspondence match based on relative location is performed and taken as truth. The empirical confusion matrices are then computed from the anisotropy attributions of these peaks. For the remainder of this section, the use the term “confusion matrix” will refer to one computed in this fashion.

The confusion matrix for this data set is given in Table 1. One noticeable property from the confusion matrix is that regardless of the conditioned training anisotropy, the full-aperture hypothesis is the most likely testing anisotropy. We give the following reasons for this. First, recall that we bias our anisotropy decision towards the

full-aperture hypothesis which partially accounts for this. The bias in the confusion matrix may also be attributed to incorrect correspondences. It is widely believed that anisotropic scatterers are less stable than isotropic scatterers and therefore are not always extracted as peaks.⁵ If an anisotropic training scatterer is not extracted in the testing data and there is a nearby isotropic scatterer, then our correspondence will incorrectly match the two. We also note that there is the unmodeled dependence on depression.

Another prominent aspect of this confusion matrix is the apparent independence of testing anisotropy and training anisotropy which would support the argument that anisotropy is not a stable feature. Under the model that anisotropy is caused by irresolvable interfering scatterers, this is reasonable especially when one considers that the depression angle has been changed. However, this result counters intuition gained from the canonical scatterer model. To further evaluate the source and stability of anisotropic phenomena, we partition the data set into those images at a near-cardinal azimuth ($\pm 2.5^\circ$ of a cardinal angle) and those at off-cardinal azimuths. The motivation being that for near-cardinal azimuths, we expect the influence of canonical scatterers to be most pronounced because of the natural rectangular shape of vehicles. Thus, near-cardinal azimuths should exhibit canonical anisotropic scattering associated with flat plates and other large simple scatterers oriented orthogonal to the impinging radar signal.

The confusion matrices for the near-cardinal and off-cardinal azimuths are given in Table 2. Here, we see striking differences. There is still a tendency to favor the full-aperture hypothesis, which we explain by the same reasoning as for the confusion matrix in Table 1. One significant difference in the near-cardinal confusion matrix from the other two is that there is a noticeable presence along the diagonal, signaling that anisotropy is more stable at these near-cardinal angles as expected for canonical scatterers. The off-cardinal confusion matrix however shows that training anisotropy is independent of testing anisotropy. This leads us to believe that there are at least two fundamental sources of anisotropy. The first is canonical scattering which dominates at cardinal azimuths and not much at other azimuths. The second is an unstable source of anisotropy which is more commonly exhibited at off-cardinal azimuths. A likely candidate for this unstable anisotropy is the scintillation produced by irresolvable interfering scatterers. Anisotropy arising from such interference is highly variable and changes unpredictably with depression, which may account for lack of correlation in the anisotropy classifications between the 15° and 17° data in the off-cardinal confusion matrix.

Table 1. Anisotropy confusion matrix for 2S1, BMP2, BRDM2, D7, T72, ZIL131, and ZSU23-4.

Training \ Testing	full-aperture	half-aperture	quarter-aperture
full-aperture	0.88	0.09	0.03
half-aperture	0.83	0.13	0.15
quarter-aperture	0.81	0.12	0.07

Table 2. Anisotropy confusion matrix for vehicles at near-cardinal and off-cardinal angles.

Training \ Testing	Near-cardinal			Off-cardinal		
	full-ap.	half-ap.	quarter-ap.	full-ap.	half-ap.	quarter-ap.
full-ap.	0.82	0.11	0.07	0.89	0.08	0.03
half-ap.	0.72	0.26	0.02	0.83	0.12	0.05
quarter-ap.	0.61	0.09	0.30	0.83	0.12	0.05

To explore how anisotropy attribution might help the recognition problem, we use anisotropy as a feature in the matcher described in Section 4. The confusion matrices in Table 2 are used to characterize uncertainty in anisotropy. Our experiments involve detection of the BMP2 and T72 from a test set composed of BMP2's, T72's, T62's, and BTR70's. In particular, for the predict data, we use peak extractions from the BMP2 (c21) and T72 (132) at a 17° depression. The extract data consists of the following vehicles at a 15° depression angle: BMP2 (9563 and 9566), T72 (812 and s7), BTR70 (c71), and T62 (a51). For the BMP2 detection statistic, we compare the likelihood ratio of the test vehicle under the BMP2 model to the T72 model, where we are treating the T72 as our model for "non-BMP2" vehicles. Similarly, we take the reciprocal for the detection statistic of the T72 letting BMP2's serve as the model for "non-T72" scatterers. We recognize that modeling the "other" class with a single vehicle is simplistic and crude,

however it is not our current goal to build a full classifier, but to setup a framework where we can study anisotropic phenomena. The T62 and BTR70 are used in the testing set because they are well known to be difficult confusers. The resulting ROC's are displayed in Figure 5 for three different sets of scatterer features:

- (F1) location
- (F2) location and anisotropy
- (F3) location and anisotropy (while restricting predict scatterers to be full-aperture).

Each feature set uses the top 10 amplitude scatterers, except the third set which uses the top 10 amplitude scatterers which are declared to be full-aperture in the predict stage. The motivation behind the conditioning in (F3) is that if anisotropic scatterers are unstable, then they are unlikely to match in the extract data, so we remove them from consideration.

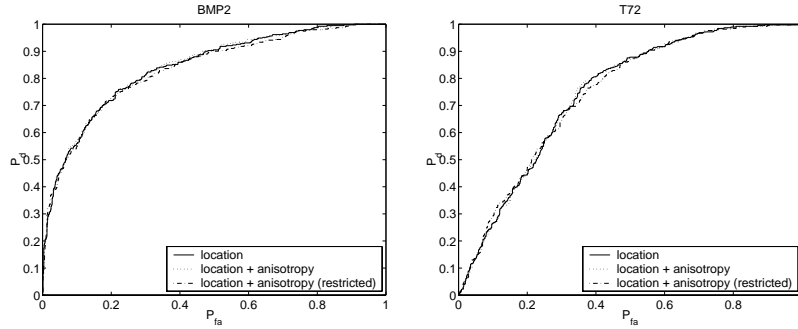


Figure 5. ROC curves for the BMP2 (left) and T72 (right) using features (F1)-(F3).

At first glance of the ROC's in Figure 5, it appears that all the tests perform equally well and the anisotropy is not useful as an attribution. However, these vehicles contain quite complex scattering phenomena and many of the anisotropy declarations may be due to volumetric interference between irresolvable scatterers which would change unpredictably with depression. With this in mind, we examine the ROC's for the test vehicles at near-cardinal ($\pm 2.5^\circ$) azimuths. These ROC's are shown in Figure 6. Although there is less statistical significance in these numbers due to the relatively small number of test vehicles at these orientations, there does appear to be a separation between each of the tests. The feature set using both location and anisotropy appears to perform the best which is what we would expect since the scattering at near-cardinal azimuths is heavily influenced by canonical scatterers. The worst performer of the three feature sets is (F3) which uses location and anisotropy, but only considers predicted scatterers which are full-aperture. Thus, this test is discarding the anisotropic scatterers it observed on the model, which is valuable information since the anisotropy exhibited from these canonical scatterers should be stable.

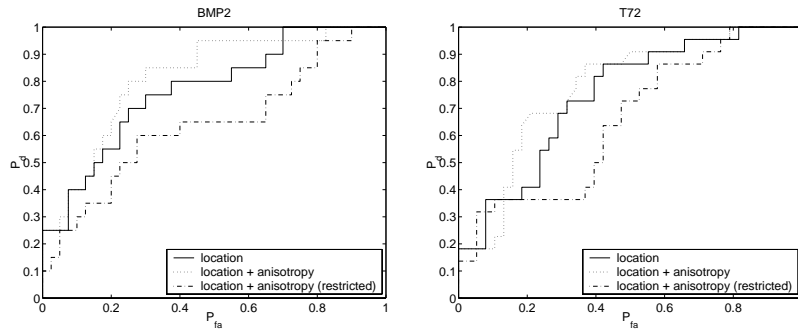


Figure 6. ROC curves for the BMP2 (left) and T72 (right) at near-cardinal angles using features (F1)-(F3).

6. SUMMARY AND DISCUSSION

We have proposed a general characterization of anisotropy based on a sub-aperture pyramid. The sub-aperture pyramid generates a tree of multi-resolution images at a variety of cross-range versus azimuthal resolution trade-offs allowing for the detection of anisotropic phenomena. With each sub-aperture in the pyramid, we associate a hypothesis that the azimuthal scattering is confined to and uniform over that sub-aperture. This then leads to a sequence of hypothesis tests to classify the anisotropy for a pixel which can be approximated with an efficient pruning algorithm due to the tree-structure over the sub-apertures.

This characterization of anisotropy allows us to explore the underlying phenomenology of anisotropic scattering. In particular, our results show that while apparent at all orientations, there seems to be markedly different sources for anisotropy. At near-cardinal azimuths, anisotropic scatterers are stable as we would expect under canonical scattering models. However, at off-cardinal angles, anisotropy is erratic and difficult to predict from a different depression. This suggests that there is a different source of anisotropy at these intermediate azimuths. A likely candidate for this anisotropy is the scintillating scattering produced by volumetric scattering. Such a group of irresolvable scatterers would exhibit anisotropic behavior due to their interference and would be unpredictable with changes in depression like we observed in our experiments.

As demonstrated by the ROC curves, these different sources of anisotropy will need to be addressed separately in order to fully utilize the information contained in each. The classification approach here used here is useful at near-cardinal angles where anisotropy is stable, but not at off-cardinal angles. The sub-aperture models used here are in their elementary stages and as they develop, they should further aid in the studying of anisotropic scattering in SAR. Even though they are motivated by canonical scattering, they detect anisotropic behavior regardless of the source and can be used to study the phenomenon in general.

ACKNOWLEDGMENTS

The work by Kim, Fisher, and Willsky was sponsored by Air Force Office of Scientific Research grant FF49620-96-1-0028 (subcontract BU GC123919NGD), Office of Naval Research grant N00014-9I-J-1004, and DARPA under Wright Patterson grant F33615-97-1014. The work by Dogan, and Moses was sponsored in part by DARPA and the US Air Force Research Laboratory under grant F33615-979191020.

REFERENCES

1. H. Chiang and R. Moses, "ATR Performance Prediction Using Attributed Scattering Features," in *Proc. of the SPIE, Algorithms for SAR Imagery*, vol. 3721, pp. 785–796, Apr. 1999.
2. A. Kim, J. Fisher, A. Willsky, and P. Viola, "Nonparametric Estimation of Aspect Dependence for ATR," in *Proc. of the SPIE, Algorithms for SAR Imagery*, vol. 3721, pp. 332–342, Apr. 1999.
3. R. Chaney, A. Willsky, and L. Novak, "Coherent aspect-dependent SAR image formation," in *Proc. of the SPIE, Algorithms for SAR Imagery*, vol. 2230, pp. 256–274, Apr. 1994.
4. H. Chiang, R. Moses, and L. Potter, "Model-based bayesian feature matching with application to synthetic aperture radar target recognition," *to appear in Pattern Recognition*, 2000.
5. W. Smith, T. Irons, J. Riordan, and S. Sayre, "Peak stability derived from phase history in synthetic aperture radar," in *Proc. of the SPIE, Algorithms for SAR Imagery*, vol. 3721, pp. 450–461, Apr. 1999.
6. G. Benitz, "High-Definition Vector Imaging for Synthetic Aperture Radar," in *31st Asilomar Conference on Signals, Systems, & Computers*, vol. 2, pp. 1204–1209, Nov. 1997.
7. L. Potter and R. Moses, "Attributed Scattering Centers for SAR ATR," *IEEE Trans. on Image Proc.* **5**, pp. 79–91, Jan. 1997.