

Analysis of the Impact of Non-Quadratic Optimization-based SAR Imaging on Feature Enhancement and ATR Performance

Müjdat Çetin, William C. Karl, and David A. Castañon

This work was supported in part by the Air Force Office of Scientific Research under Grants F49620-00-0362 and F49620-96-1-0028, the Office of Naval Research under Grant N00014-00-1-0089, the Army Research Office under Grant DAAG55-97-1-0013, and the National Institutes of Health under Grant NINDS 1 R01 NS34189.

M. Çetin is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA (e-mail: mcetin@mit.edu).

W. C. Karl and D. A. Castañon are with the Multi-Dimensional Signal Processing Laboratory, Boston University, Boston, MA 02215, USA (e-mails: wckarl@bu.edu, dac@bu.edu).

Portions of preliminary versions of this work have appeared as conference papers in [1] and [2].

Abstract

We present an evaluation of the impact of a recently proposed synthetic aperture radar (SAR) imaging technique on feature enhancement and automatic target recognition (ATR) performance. This image formation technique is based on non-quadratic optimization, and the images it produces appear to exhibit enhanced features. In the first part of this paper, we quantify such feature enhancement through a number of criteria. The findings of our analysis indicate that the new feature-enhanced SAR image formation method provides images with higher resolution of scatterers, and better separability of different regions as compared to conventional SAR images. In the second part of this paper, we provide an ATR-based evaluation. We run recognition experiments using conventional and feature-enhanced SAR images of military targets, with three different classifiers. The first classifier is template-based. The second classifier makes a decision through a likelihood test, based on Gaussian models for reflectivities. The third classifier is based on extracted locations of the dominant target scatterers. The experimental results demonstrate that the new feature-enhanced SAR imaging method can improve the recognition performance, especially in scenarios involving reduced data quality or quantity.

Keywords

Synthetic aperture radar, automatic target recognition, image reconstruction, non-quadratic regularization, feature-enhanced imaging, superresolution, classification, feature extraction

I. INTRODUCTION

Automatic target recognition (ATR) methods based on synthetic aperture radar (SAR) imagery depend implicitly or explicitly on certain features of the objects in the reconstructed images [3–6]. Robust feature extraction or exploitation can be difficult when based on SAR images formed by conventional methods, such as the polar format algorithm [7]. One challenge is that the resolution of the formed images is limited by the SAR system bandwidth, which complicates point scatterer localization. In addition, conventional SAR images suffer from speckle [8, 9], which complicates region segmentation for shape-based recognition.

Recently, we have developed a new technique for SAR image formation based on non-quadratic regularization [10]. This method produces images which appear to enhance *point-based features* (e.g. scatterer locations), and *region-based features* (e.g. object shapes, boundaries). Such features have been explicitly used for ATR [5, 11–16]. In particular, both scatterer locations and object shapes have played a central role in feature-based classification algorithms developed recently as part of the Moving and Stationary Target

Acquisition and Recognition (MSTAR) program [5,15]. Similarly, the work in [11] contains a physics-based scattering center model for SAR ATR, which has later been used for ATR performance prediction as well [14]. Scatterer locations and amplitudes have also been used for recognition of occluded objects in SAR images [12,13]. The work in [16] uses region-based features, in particular extracted object boundaries, for ATR. This body of work suggests that the features which (qualitatively) appear to be enhanced by the SAR imaging technique of [10] play an important role in ATR. Therefore it is of interest to analyze the images produced by this technique in terms of such feature enhancement, and also in terms of the effect of their use in ATR systems.

In this paper, we present a quantitative evaluation of the imagery produced by the *feature-enhanced* SAR imaging technique of [10]. Our analysis consists of two parts. The first part is a feature-based evaluation, where the goal is to determine whether the images produced by the new method lead to increased accuracy or robustness in feature extraction. The second part of our analysis is aimed at quantifying the impact of feature-enhanced imaging on ATR performance.

For the first part of our analysis, we propose a number of criteria for feature-based evaluation of the reconstructed SAR images. Most of these criteria have been used in the literature before for evaluating other SAR imaging techniques [17,18], and a number of them are new. In particular, the criteria we use regarding point-based features are target-to-clutter ratio [17], main-lobe width [17], peak matching accuracy, and average associated peak distance. The criteria we use for region-based features are speckle suppression [17], segmentation accuracy, and statistical separability of different regions in the scene from one another [18]. We run experiments on the MSTAR public target data set [19] to compare feature-enhanced SAR images to conventional images in terms of these quantitative measures. The results of this study show that emphasizing point-based features through the new image formation method yields images with higher resolution and better dominant scatterer localization than conventional images. Our experiments indicate that the method is able to produce accurate superresolution reconstructions from considerably reduced amounts of data. Emphasizing region-based features on the other hand, results in enhanced anomaly and speckle suppression in homogeneous regions, which results in

easier-to-segment images.

We would like to clarify the meaning of “superresolution” in the context of our study. Current superresolution techniques, including the one evaluated in this paper, make strong assumptions about the nature of the targets of interest (i.e. that these objects are well-represented by a collection of discrete point scatterers). These techniques do not provide a higher imaging resolution uniformly over all scenes. Therefore, such techniques might yield a “higher resolution” than conventional images only in the sense of more accurate representation of point-based features. Our evaluation is based on this perspective.

In the second part of our analysis, we present an evaluation of the impact of feature-enhanced SAR imaging on ATR performance. There has been some recent work performing ATR-based evaluations of a number of new SAR image formation techniques. In particular, the high-definition imaging (HDI) technique of [17] has been evaluated in [20]; and a number techniques (including HDI, and those described in [21]) have been evaluated in [22], through the use of a template-based classifier [23,24]. In both of these studies, the new SAR imaging methods have been found to improve the classification performance, as compared to conventional imaging. In our work, we use the MSTAR public target data set, and three different classifiers, to make decisions about the target type given conventional or feature-enhanced images, and compare the results for these two types of images. The first classifier is based on template-matching, which is a conventional approach for SAR ATR [23, 24]. The second classifier is based on conditionally Gaussian models for reflectivities, and performs classification through a likelihood test [25]. The third classifier is a feature-based one, and performs classification by measuring how well the dominant scatterer locations extracted from images formed from reduced-resolution data match the true scatterer locations for each hypothesized target. We present the recognition results in the form of classifier confusion matrices. These experiments demonstrate that feature-enhanced SAR imaging can offer higher probability of correct classification than conventional imaging.

In Section II, we provide a brief overview of the feature-enhanced SAR imaging technique we evaluate in this paper. Section III contains the descriptions of the feature-based criteria to be used in the first part of our analysis. In Section IV, we describe the classifiers we use

in the second part of our analysis for ATR-based evaluation. Sections V and VI contain the experimental results corresponding to the first and second parts of our analysis respectively. Finally, we conclude in Section VII.

II. SAR IMAGING BASED ON NON-QUADRATIC OPTIMIZATION

In this section, we provide a brief overview of feature-enhanced SAR image formation based on non-quadratic optimization. The details of the technique can be found in [10]. Let us start from the following assumed discrete model for the SAR observation process:

$$\mathbf{g} = \mathbf{T}\mathbf{f} + \mathbf{w} \quad (1)$$

where \mathbf{g} denotes the SAR observations (phase histories or range profiles), \mathbf{f} is the unknown sampled reflectivity image, \mathbf{w} is additive measurement noise, all column-stacked as vectors, and \mathbf{T} is a complex-valued SAR observation matrix. If phase history data are used, \mathbf{T} is a 2-D Fourier-type operator [26], and if range profile data are used, \mathbf{T} is a complex-valued projection-type operator. In this framework, the objective of SAR image reconstruction is to obtain an estimate of \mathbf{f} based on the data \mathbf{g} . The conventional SAR polar format image formation algorithm [7] can roughly be interpreted in this framework as the application of the adjoint to the data: $\hat{\mathbf{f}}_{\text{CONV}} = \mathbf{T}^H \mathbf{g}$.

In contrast, feature-enhanced SAR image reconstruction is achieved by solving an optimization problem of the following form:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} [\|\mathbf{g} - \mathbf{T}\mathbf{f}\|_2^2 + \lambda_1 \|\mathbf{f}\|_p^p + \lambda_2 \|\mathbf{D}|\mathbf{f}|\|_p^p] \quad (2)$$

where $\|\cdot\|_p$ denotes the ℓ_p -norm ($p \leq 1$), \mathbf{D} is a 2-D derivative operator, $|\mathbf{f}|$ denotes the vector of magnitudes of the complex-valued vector \mathbf{f} , and λ_1, λ_2 are scalar parameters. The first term in the objective function of (2) is a data fidelity term. The second and third terms incorporate prior information regarding both the behavior of the field \mathbf{f} , and the nature of the features of interest in the resulting reconstructions. The optimization problem in (2) can be solved by using an efficient iterative algorithm proposed in [10], based on half-quadratic regularization [27]. Use of non-quadratic constraints, such as the second and third terms in (2), has recently become popular, e.g. in image restoration [28], due to the ability of these constraints to prevent suppression of useful features in the image.

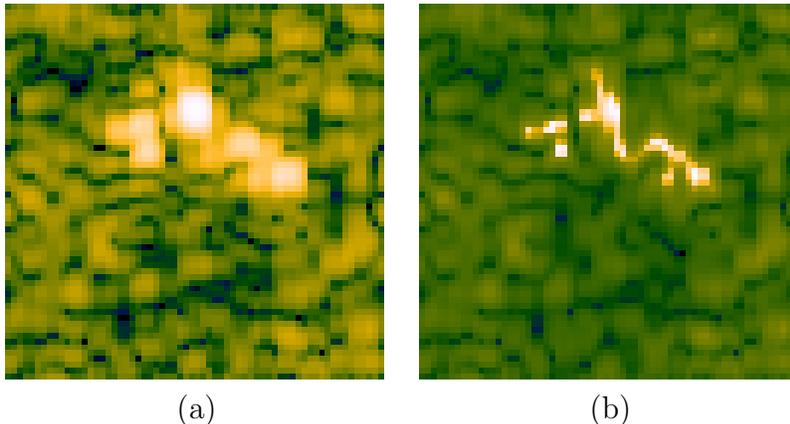


Fig. 1. (a) Conventional SAR image. (b) The corresponding point-enhanced, superresolution SAR image. Taken from [10].

The objective function in (2) extends the use of such constraints to the complex-valued SAR image reconstruction problem.

Each of the last two terms in (2) is aimed at enhancing a particular type of feature that is of importance for SAR ATR. In particular, the term $\|\mathbf{f}\|_p^p$ is an energy-type constraint on the solution, and aims to suppress artifacts and increase the resolvability of *point* scatterers. The $\|\mathbf{D}\mathbf{f}\|_p^p$ term on the other hand, aims to reduce variability in homogeneous *regions*. The relative magnitudes of λ_1 and λ_2 determine the emphasis on such *point-based* versus *region-based* features. Therefore, this framework lets us reconstruct images with two different flavors: using a relatively large λ_1 yields *point-enhanced* imagery, and using a relatively large λ_2 yields *region-enhanced* imagery. For point-enhanced imaging, when the observation model \mathbf{T} maps a high-resolution field to lower-resolution data, we call this *superresolution* imaging, since in this case we attempt to reconstruct the field at a finer pixel-spacing than Rayleigh resolution. When this is not the case, we use the term *non-superresolution* imaging. We refer to any image produced by this technique as a *feature-enhanced* image. Figures 1 and 2 contain samples of point-enhanced and region-enhanced images.

III. FEATURE-BASED CRITERIA FOR EVALUATION OF IMAGE QUALITY

In this section, we propose quantitative criteria for evaluating the images produced by the SAR image formation technique developed in [10], and outlined in Section II.

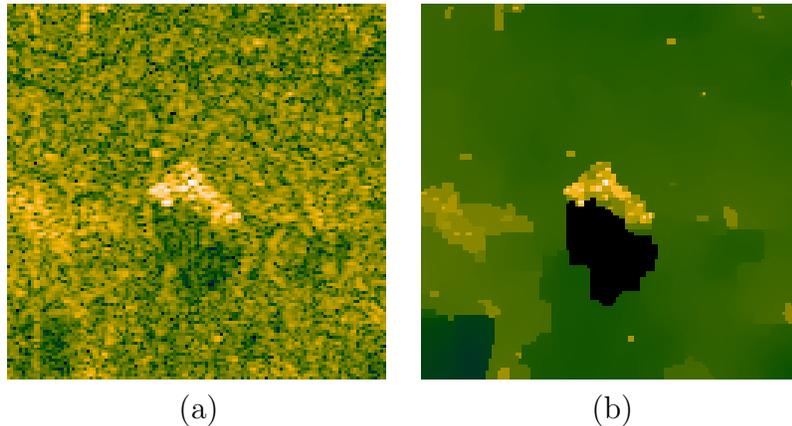


Fig. 2. (a) Conventional SAR image. (b) The corresponding region-enhanced SAR image. Taken from [10].

A. Evaluation Criteria for Point-Enhanced Images

The criteria we describe in this section for evaluating point-based features are target-to-clutter ratio, mainlobe width, peak matching accuracy, and average associated peak distance. The first two of these criteria have been used for evaluating SAR imaging methods before [17]. The last two criteria are more powerful, and are aimed at measuring how accurately dominant scatterer features are preserved in the face of reduced amounts of data, hence these criteria are particularly useful for testing the superresolution properties of an image formation method.

A.1 Target-to-Clutter Ratio

As a measure of accentuation of the target pixels with respect to the background, we use the target-to-clutter ratio in dB, defined as [17]:

$$\text{Target-to-clutter ratio} = 20 \log_{10} \left(\frac{\max_{(i,j) \in \mathcal{T}} (|\hat{\mathbf{f}}_{ij}|)}{\frac{1}{N_C} \sum_{(i,j) \in \mathcal{C}} |\hat{\mathbf{f}}_{ij}|} \right) \quad (3)$$

where the pair (i, j) denotes the pixel indices, $\hat{\mathbf{f}}$ is the reconstructed image, $\hat{\mathbf{f}}_{ij}$ is the reconstructed reflectivity at location (i, j) , \mathcal{T} denotes the target region, \mathcal{C} denotes a clutter patch in the image, and N_C denotes the number of pixels in the clutter patch.¹

¹Throughout the paper we use both a double index notation, $\hat{\mathbf{f}}_{ij}$, and a single index notation, $\hat{\mathbf{f}}_i$, to refer to the elements of $\hat{\mathbf{f}}$. The two conventions correspond to referring to the elements of a matrix, and to the elements of a

A.2 Mainlobe Width

As one of the measures of the effective resolution of an image, we use the 3-dB mainlobe width of the strong scatterers. To obtain an estimate of the mainlobe width, we concentrate on the target region. In each row and column in the target region of the reconstructed image, we find the first point near the maximum where the reflectivity magnitude is more than 3 dB below the maximum value. We then obtain a better estimate of the 3-dB distance by means of a linear interpolation between pixels. Finally, we average the distances obtained from each row and column in the target region to find an overall estimate of the 3-dB lobe width for a particular image.

A.3 Peak Matching Accuracy

Locations of dominant point scatterers extracted from a target image are important characteristics for recognition [11, 12]. Loss of resolution manifests itself by merging and moving such characteristic points, and this makes the accurate localization of these points in the scene more difficult. Thus, we evaluate the superresolution properties of a reconstruction method by measuring how well the dominant scatterers are preserved when we use reduced-resolution data to form the image.

The procedure we use for measuring peak matching accuracy is as follows. The first step is to extract the locations of N_p brightest scatterers from a reconstructed image. For this purpose, we first find all the peaks in the scene. The peaks are taken to be the points where the discrete spatial derivatives of the reflectivity magnitude in both the x and the y directions change sign from positive to negative. Once the peaks are found, we order them based on their magnitudes, and pick the largest N_p of them. We then evaluate how well the coordinates of these N_p peaks match those of the N_p “reference” locations of the scatterers. The ground truth for these reference positions is usually not available. One way to obtain the reference peak locations is through the use of 3-D CAD models of the vehicles together with an electromagnetic signature prediction tool such as XPATCH [29] to simulate the expected reference scene. A second option is to use conventional SAR images obtained from higher resolution data as the reference scenes, and extract the vector obtained by column-stacking the matrix respectively.

“true” peak locations from such images. Once we have two sets of N_p peak locations, we need to determine how many of the extracted locations “match” a reference location. The simplest matching criterion would be to count the exact position matches between the two sets of locations. However, this criterion may be too restrictive. In order to have a more flexible criterion, we use the feature matching method used in [30]. This method allows a match declaration between two peaks, if the estimated peak location is within a radius r of the “reference” peak location. Hence it is more powerful than counting only the exact matches, with r used as a variable parameter ($r = 0$ corresponds to counting the exact matches). A one-to-one association of the peaks is made such that the sum of the squared distances between the locations of the “reference” peaks and the corresponding matched peaks from the image is minimized. We can then count the number of matched peaks out of N_p , to see how well the peaks are preserved.

We apply the above procedure to the conventional images and the feature-enhanced images reconstructed from the same reduced-resolution data. We compute the number of matched peaks for a large set of such conventional and feature-enhanced images, and compare the average number of matched peaks, parameterized by r .

A.4 Average Associated Peak Distance

Another criterion based on peak locations that we use is the average distance between the two sets of N_p matched peak coordinates. To compute this measure, we relax the matching radius r of Section III-A.3, so that each of the N_p peaks from the reconstructed image is matched to one of the “reference” peaks. We then find the average of the distances between these associated peaks.

B. Evaluation Criteria for Region-Enhanced Images

In this section, we describe the criteria for evaluating region-based features, namely speckle suppression, segmentation accuracy, and statistical separability of distinct regions from one another. The first and third of these criteria have been used for the evaluation of SAR imaging methods, in [17] and [18], respectively. We use the accuracy of the segmentations (obtained through the use of a simple segmenter) as a measure of how easy-to-segment the region-enhanced SAR images are.

B.1 Speckle Suppression

Speckle complicates intensity-based region description in conventional SAR images. One criterion that has been used for measuring speckle amplitude is the standard deviation of a clutter patch in the dB-valued SAR images [17]. We use this measure to quantitatively compare region-enhanced reconstructions with conventional images in terms of speckle suppression.

B.2 Segmentation Accuracy

It is of interest to obtain accurate segmentations of SAR images for effective use of region-based shape features in target recognition. Recently there has been much interest in the development of segmentation algorithms for conventional SAR images. Region-enhanced images appear to provide easier-to-segment regions as compared to conventional SAR images. We demonstrate this property by segmenting region-enhanced reconstructions to target, shadow and background regions by simple adaptive thresholding, where the thresholds for a particular image under consideration depend on the statistics of that image. To determine the thresholds, we find the sample mean μ and the standard deviation σ of the dB-valued pixel magnitudes in the image. Then, we apply the following decision rule at each pixel:

$$\begin{aligned} 20 \log_{10}(|\hat{\mathbf{f}}_{ij}|) < \mu - c_1\sigma &\implies (i, j) \in \mathcal{S} \\ \mu - c_1\sigma \leq 20 \log_{10}(|\hat{\mathbf{f}}_{ij}|) < \mu + c_2\sigma &\implies (i, j) \in \mathcal{B} \\ \mu + c_2\sigma \leq 20 \log_{10}(|\hat{\mathbf{f}}_{ij}|) &\implies (i, j) \in \mathcal{T} \end{aligned}$$

where $\mathcal{T}, \mathcal{S}, \mathcal{B}$ denote the target, shadow and background regions respectively and c_1, c_2 are two constants that are fixed beforehand. Hence, this is really a “histogram-based” threshold. From a statistical standpoint, it would make more sense to develop a decision metric based on the statistics of particular regions. However, our objective here is not to develop the best decision metric, but rather to show that we can obtain reasonable segmentations of the region-enhanced images even by simple suboptimal processing.

The above procedure produces a segmented image. Next, we need to evaluate the accuracy of this segmentation. For such evaluation, we ideally need the ground truth. Ground truth for this problem consists of the assignment of each location in the scene to

one of the three sets $\mathcal{T}, \mathcal{S}, \mathcal{B}$. Such ground truth information is not readily available in general, as is the case with the MSTAR data set. However, for this particular data, the Air Force Research Laboratory (AFRL) has conducted a study which has produced a set of segmentations performed by humans [31]. In our work, we use these human segmentations as the ground truth. The measure of accuracy we use is the percentage of pixels in the automatically segmented image, which are classified correctly (i.e. assigned to the same set as the “ground truth” assignment of the pixel).

B.3 Statistical Separability of Regions

Segmentation accuracy provides a good flavor of the degree of separability of different regions from one another in the SAR image. However it does not provide a statistical measure for the similarity of different regions, $\mathcal{T}, \mathcal{S}, \mathcal{B}$. To obtain such a measure, we first assume that all the (dB-valued) reflectivity magnitudes in a particular region of the reconstructed SAR target image are drawn from the same Gaussian distribution. We also assume that these region-based probability density functions are independent of the target type. We then estimate the mean and variance of such a density for each region, $\mathcal{T}, \mathcal{S}, \mathcal{B}$, using a large set of SAR images. Note that the actual distribution of those elements may not be perfectly modeled by a Gaussian, but such a simple model is sufficient for our objective of measuring region separability.

As a measure of the similarity of two Gaussian densities corresponding to classes (regions) i and j , we use the Bhattacharyya distance d_{ij} [18, 32]:

$$d_{ij} = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{|\frac{\Sigma_i + \Sigma_j}{2}|}{|\Sigma_i|^{1/2} |\Sigma_j|^{1/2}} \right) \quad (4)$$

where μ_i, μ_j are the mean vectors and Σ_i, Σ_j are the covariance matrices for class i and class j . Note that, in our case the means and covariances are just scalars. The distance d_{ij} is a common measure of the separability of classes characterized by multi-dimensional Gaussian distributions and gives an upper bound on the Bayesian error for classification of two classes. We compute Bhattacharyya distances for each pair of different regions, $\mathcal{T}, \mathcal{S}, \mathcal{B}$, and compare such distances for conventional and region-enhanced images. A larger Bhattacharyya distance means that the regions are easier to separate from each other.

IV. CLASSIFIERS USED FOR ATR

In this section, we describe the three classifiers we use for evaluating the impact of feature-enhanced SAR imaging on the success of ATR. The first of these is a template-based classifier, which has been used in a series of ATR systems developed by Lincoln Laboratory [22–24]. The second classifier achieves a decision through a likelihood test on the observed target images based on Gaussian models for the reflectivities parameterized by the target class [25]. The third one, a point-feature-based classifier, is based on measuring how well features (dominant scatter locations) extracted from images reconstructed from low resolution data match those of the hypothesized target class.

A. Template-based Classifier

The idea in template-based classification is to first measure how well a given test image matches reference images, called templates, which represent the hypotheses in the problem, and then declare the test image to be from the class giving the best matching score. A common metric used for the degree of match is the mean-squared error (MSE) between the test image, and the template.²

For template-based classification, we need templates of the targets at a reasonable number of different orientations. Let $\mathbf{t}(\vartheta, a)$ be such a template (stacked as a vector) for vehicle type a , at orientation $\vartheta \in [0^\circ, 360^\circ)$, and let $\hat{\mathbf{f}}_{\text{dB}}$ be a normalized (to have unit ℓ_2 -norm) test image (in dB) stacked as a vector. Then the template-based minimum MSE classification can be done as follows:

$$\hat{a}_{\text{MSE}} = \arg \min_a \left(\min_{\vartheta} \|\hat{\mathbf{f}}_{\text{dB}} - \mathbf{t}(\vartheta, a)\|_2^2 \right). \quad (5)$$

The classifier in (5) needs a stored set of templates, which can be obtained from training data. In practice, we usually do not have many training images at exactly the same orientation. Then, the template for vehicle type a , at orientation ϑ_k can be constructed by averaging the training images that have an orientation close to ϑ_k , as follows:

²Note that MSE-based classifiers are not very robust to target signature variability. There are more sophisticated classifiers which are more successful in this respect. However, for our analysis, it is sufficient to use any reasonable classifier to compare the performance of conventional and feature-enhanced images given that particular classifier.

$$\mathbf{t}(\vartheta_k, a) = \frac{1}{N_k} \sum_{\vartheta \in W_k} \hat{\mathbf{f}}_{\text{dB}}^{r_k}(\vartheta, a) \quad (6)$$

where W_k denotes the range of angular orientations that are used in training for the orientation ϑ_k , N_k is the number of available training images that have an orientation within this range, and $\hat{\mathbf{f}}_{\text{dB}}^{r_k}(\vartheta, a)$ is a training image which has an original orientation of ϑ , and which is rotated $\vartheta_k - \vartheta$ degrees, so that it is aligned with other training images used for the computation of $\mathbf{t}(\vartheta_k, a)$. Figures 3, 4, and 5 show such templates for T72 tanks from the MSTAR data set, at 17° depression angle³ and various orientations for the conventional, point-enhanced (non-superresolution), and region-enhanced images. The orientation ϑ for each image used in the computation of these templates is obtained from the corresponding MSTAR file header. In the construction of these templates, we have used $\vartheta_k = 5(k - 1)$ degrees, and $W_k = [\vartheta_k - 5, \vartheta_k + 5)$, where $k \in \{1, 2, \dots, 72\}$.

B. Likelihood-based Classifier

We now describe a likelihood-based classifier, proposed in [25] for SAR ATR. This approach starts from a statistical signal model, in which the underlying SAR image $\mathbf{f}(\vartheta, a)$, where ϑ and a are as defined in Section IV-A, is assumed to be a complex Gaussian vector, and the observed image $\hat{\mathbf{f}}$ is given by the following observation model:

$$\hat{\mathbf{f}} = \mathbf{f}(\vartheta, a) + \mathbf{n} \quad (7)$$

where \mathbf{n} is complex Gaussian noise with mean $\mathbf{0}$ and covariance $\sigma^2 \mathbf{I}$. The classifier is based on a generalized likelihood test as follows:

$$\hat{a}_{\text{GLRT}} = \arg \max_a \left(\max_{\vartheta} l(\hat{\mathbf{f}}|\vartheta, a) \right) \quad (8)$$

where $l(\hat{\mathbf{f}}|\vartheta, a)$ denotes the logarithm of the likelihood that the test image is of target type a , at orientation ϑ . Let the signal $\mathbf{f}(\vartheta, a)$ have mean $\mathbf{m}(\vartheta, a)$ and covariance $\mathbf{Q}(\vartheta, a)$. Then the probability density function $p(\hat{\mathbf{f}}|\vartheta, a)$ is Gaussian with the following mean and covariance:

³Depression angle is the angle between the horizontal plane and the line that connects the radar platform to the imaged scene.

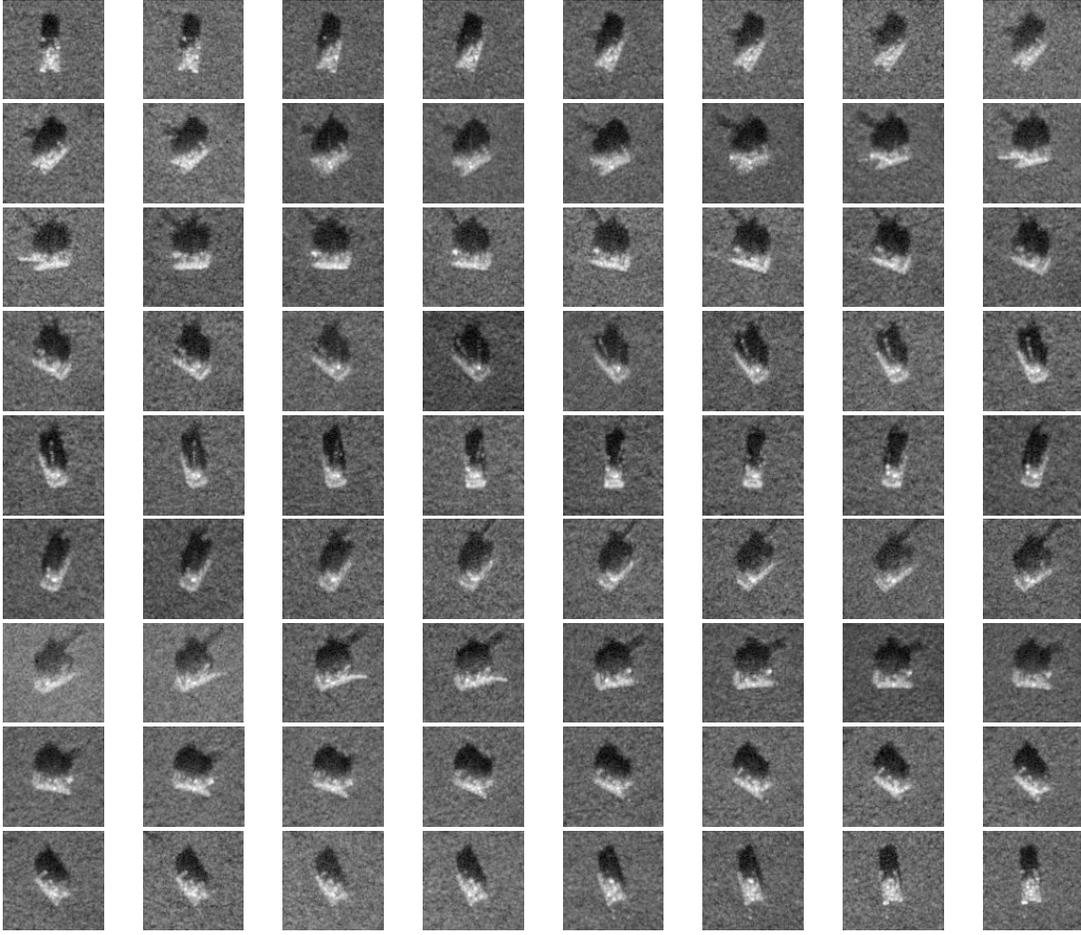


Fig. 3. Conventional templates for the T72 target at 17° depression angle. Each image shows the template for a different aspect angle, starting from 0° on the top left, and covering all 360° with 5° increments, and a 10° window for averaging.

$$E\{\hat{\mathbf{f}}|\vartheta, a\} = \mathbf{m}(\vartheta, a) \quad (9)$$

$$E\{[\hat{\mathbf{f}} - \mathbf{m}][\hat{\mathbf{f}} - \mathbf{m}]^H|\vartheta, a\} = \mathbf{Q}(\vartheta, a) + \sigma^2\mathbf{I}. \quad (10)$$

Let us define $\mathbf{K}(\vartheta, a) \triangleq \mathbf{Q}(\vartheta, a) + \sigma^2\mathbf{I}$. Then the log-likelihood is given by:

$$l(\hat{\mathbf{f}}|\vartheta, a) \propto -\log |\mathbf{K}(\vartheta, a)| - [\hat{\mathbf{f}} - \mathbf{m}(\vartheta, a)]^H (\mathbf{K}(\vartheta, a))^{-1} [\hat{\mathbf{f}} - \mathbf{m}(\vartheta, a)]. \quad (11)$$

We assume that $\mathbf{m}(\vartheta, a) = \mathbf{0}$, $\forall\{\vartheta, a\}$, as in [25]. This is a reasonable assumption due to the random phase nature of SAR reflectivities. The second simplifying assumption made in [25] is that of statistical independence of different pixels in the SAR image, which results in a diagonal $\mathbf{Q}(\vartheta, a)$, hence a diagonal $\mathbf{K}(\vartheta, a)$. This in turn simplifies the computation

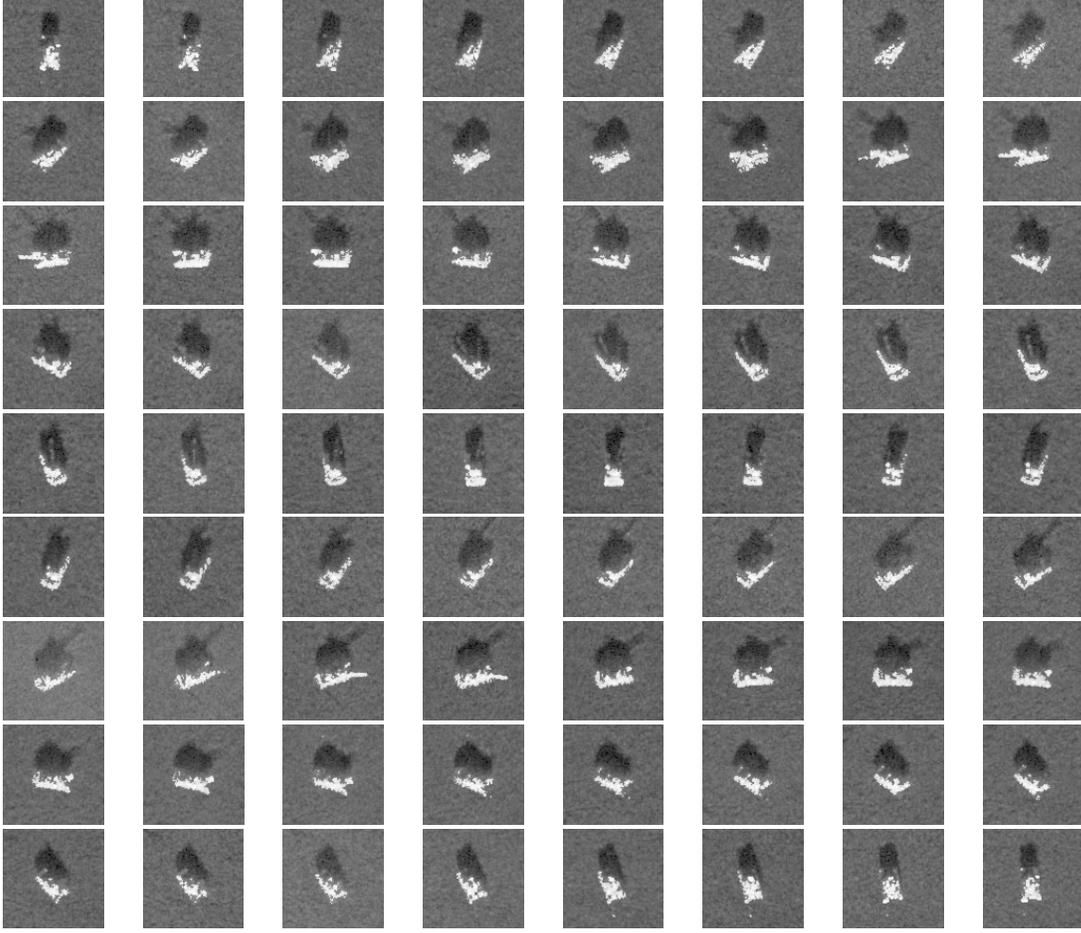


Fig. 4. Point-enhanced templates for the T72 target at 17° depression angle. Each image shows the template for a different aspect angle, starting from 0° on the top left, and covering all 360° with 5° increments, and a 10° window for averaging.

of the likelihood to a simple summation:

$$l(\hat{\mathbf{f}}|\vartheta, a) \propto \sum_i \left[-\log(\mathbf{K}_{i,i}(\vartheta, a)) - \frac{|\hat{\mathbf{f}}_i|^2}{\mathbf{K}_{i,i}(\vartheta, a)} \right] \quad (12)$$

where $(\cdot)_i$ and $(\cdot)_{i,i}$ denote the i -th element of a vector, and the i -th diagonal element of a matrix respectively. The classifier in (8) with the likelihood function in (12) requires the model variances $\mathbf{K}_{i,i}(\vartheta, a)$. These variances can be estimated from training data as follows:

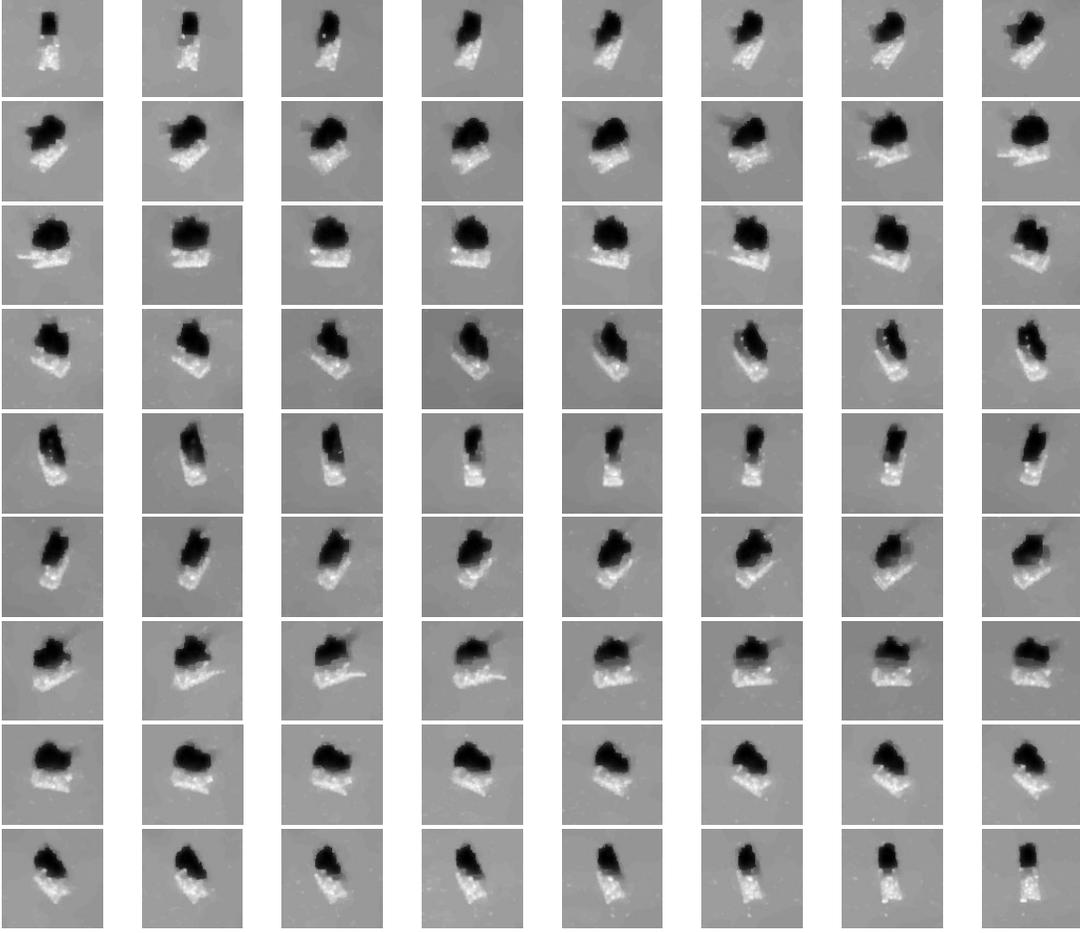


Fig. 5. Region-enhanced templates for the T72 target at 17° depression angle. Each image shows the template for a different aspect angle, starting from 0° on the top left, and covering all 360° with 5° increments, and a 10° window for averaging.

$$\mathbf{K}_{ii}(\vartheta_k, a) = \frac{1}{N_k} \sum_{\vartheta \in W_k} \left| \left(\hat{\mathbf{f}}^{rk}(\vartheta, a) \right)_i \right|^2 \quad (13)$$

where $(\hat{\mathbf{f}}^{rk}(\vartheta, a))_i$ denotes the i -th pixel of an aligned training image, and N_k , W_k are as defined in Section IV-A. The operation in (13) produces variance “images”, which are, in spirit, very similar to the template images of Figures 3, 4, and 5. We pick the values of ϑ_k and W_k as in Section IV-A.

C. Point-feature-based Classifier

We now concentrate on the reconstruction problem from reduced-resolution data, and propose a classifier specifically aimed at evaluating the ability of the point-enhanced, super-resolution images to preserve the locations of dominant scatterers in the scene. Similar peak-based classification schemes have previously been proposed in [12] and [15]. Such feature-based classification techniques, as opposed to the pixel-based classification schemes described in Sections IV-A and IV-B, have been an important component of recent research efforts such as those in DARPA’s MSTAR program [15].

The classifier works as follows. Given a test image, the first step is to extract the locations of the N_p brightest scatterers, which can be done as described in Section III-A.3. Once the N_p dominant peaks are extracted, a one-to-one association is established between these peaks and the true peak locations for the hypothesis under test. This association is made such that the sum of the squared distances between the locations of the true peaks and the corresponding matched peaks from the test image is minimized. Next, an average of the distances between the associated peak pairs is computed. Note that this distance is exactly the feature-based evaluation criterion we described in Section III-A.4. This gives a single number that provides an indication of the similarity of the test image and the hypothesized class: a smaller number tends to indicate a closer correspondence. The class assignment is achieved by finding the vehicle type which yields the minimum associated peak distance over all orientations:

$$\hat{a}_{\text{PEAK}} = \arg \min_a \left(\min_{\vartheta} \frac{1}{N_p} \|\mathbf{d}_{\text{PEAK}}(\hat{\mathbf{f}}, \mathbf{f}(\vartheta, a))\|_1 \right) \quad (14)$$

where $\mathbf{d}_{\text{PEAK}}(\cdot, \cdot)$ is a vector of Euclidean distances between associated peak pairs extracted from the observed test image $\hat{\mathbf{f}}$, and the reference image $\mathbf{f}(\vartheta, a)$. Note that the ℓ_1 -norm in (14) is just equivalent to a summation of the distances between individual peak pairs. The classifier in (14) requires the “true” peak locations for each target type, at a reasonable number of orientations ϑ , extracted from the reference scenes $\mathbf{f}(\vartheta, a)$. We obtain these reference peak locations as described in Section III-A.3.

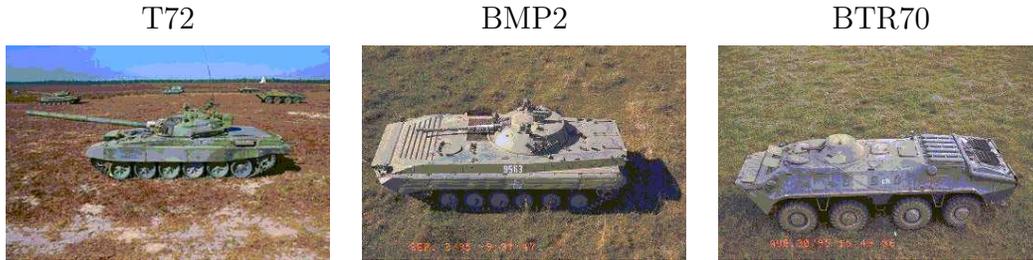


Fig. 6. Sample pictures of the three types of vehicles whose SAR images are used in the experiments. Taken from [15].

V. EXPERIMENTAL RESULTS FOR FEATURE ENHANCEMENT

In this section, we present the results of our experiments comparing feature-enhanced SAR images with conventional images in terms of the criteria described in Section III.

A. Experimental Setup

The MSTAR public target data set [19] provides SAR images of various military vehicles. We use images of T72 tanks, BMP2 tanks, and BTR70 armored personnel carriers from this data set. Sample pictures of these three types of vehicles are shown in Figure 6. We use 72 images for each vehicle type, all at 17° depression angle, and evenly spaced (approximately 5°) in azimuth (aspect angle) to cover 360° . Figure 7 shows the magnitude of a sample MSTAR image in dB for each target type.

We obtain the phase history data from the 128×128 complex-valued MSTAR images, by undoing the final steps of MSTAR image formation, as described in, e.g. [33, 34]. In particular, we first take the 2-D Discrete Fourier Transform (DFT) of the images, then we remove the zero-padding to obtain 100×100 phase history samples. Next we remove the windowing applied. From the MSTAR file headers, we know that a 35 dB Taylor window has been used, and we assume that the order of the Taylor window used is $\bar{n} = 4$. We divide the phase history samples by a 2-D Taylor window we thus obtain.

When applying feature-enhanced reconstruction, we choose the feature accentuation parameters λ_1 and λ_2 in (2) based on subjective qualitative assessment of one image from the entire data set, and use those values for all the images for that particular experiment. In the following sections, we provide the precise values of the parameters used in each experiment. In order for the parameter values used at different experiments to be relatable,

T72 (azimuth=100.77°) BMP2 (azimuth=175.19°) BTR70 (azimuth=94.00°)

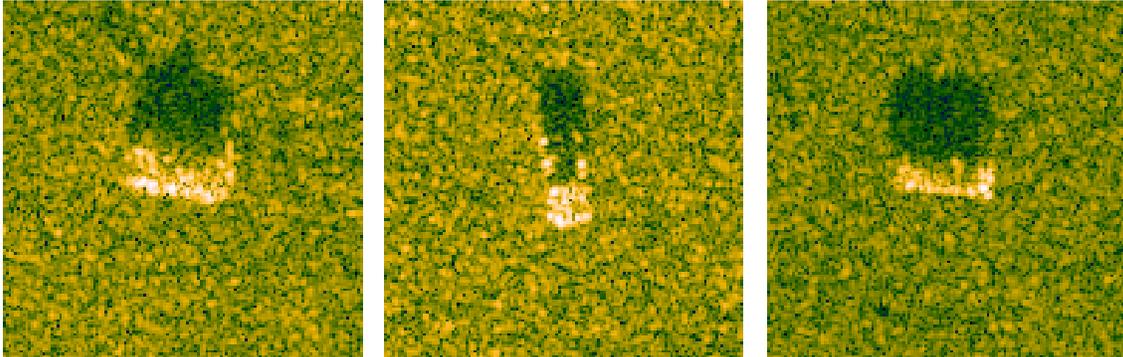


Fig. 7. 128×128 sample images from the MSTAR public target data set.

we normalize \mathbf{T} such that the magnitude of the diagonal elements of $\mathbf{T}^H \mathbf{T}$ is unity in each experiment. We should note that we have not optimized any of the parameter values in terms of the evaluation metrics we use, neither have we studied the sensitivity of our evaluation results to the choice of these parameters.

B. Point-Enhanced Imaging from Full-Resolution Data

In this section, we report the results of our experiments using 100×100 phase history samples to form point-enhanced images. The resolution supported by such data is 0.3 m. We form “critically-sampled” (i.e. 100×100) feature-enhanced images. In order to have conventional SAR images of identical dimensions for comparison, we form 100×100 Taylor-windowed Fourier images. Samples of such reconstructions are shown in the top row of Figure 8. Naturally these are very similar to their oversampled versions in Figure 7.

We form point-enhanced images with $p = 0.8$, $\lambda_1 = 2$, and $\lambda_2 = 0$ in (2). We do not apply any windowing to the data before processing, since feature-enhanced imaging is able to suppress sidelobes considerably even with rectangular weighting. However, if desired, the method can be used with windowing, as it was demonstrated in [34]. The bottom row in Figure 8 shows the reconstructions obtained. The dominant scatterers appear to be accentuated as compared to the conventional images in the top row.

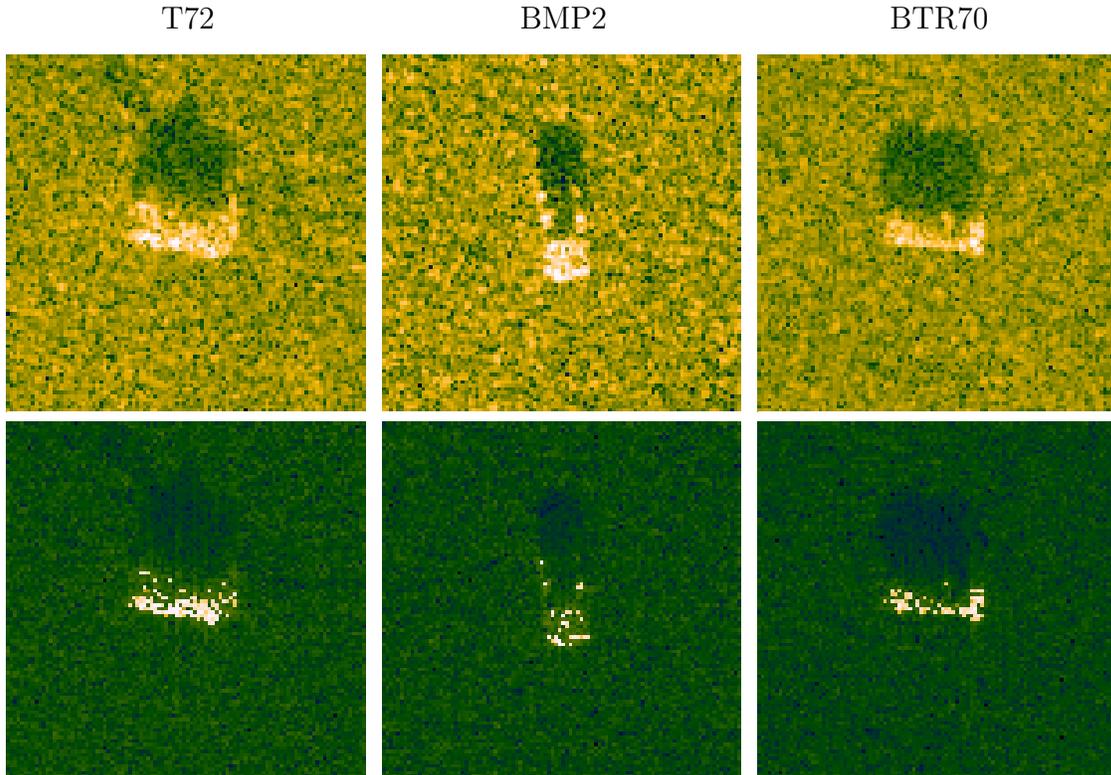


Fig. 8. 100×100 sample images reconstructed from full-resolution (100×100) data. Top: conventional. Bottom: proposed point-enhanced, with $p = 0.8$, $\lambda_1 = 2$, $\lambda_2 = 0$.

B.1 Target-to-Clutter Ratio

We quantify the enhancement of the target pixels in the full-resolution data reconstructions by means of their target-to-clutter ratio. We compute the target-to-clutter ratio as defined in (3), by using the bottom 20 rows (2000 pixels) of the reconstructed images as the clutter region. This region is big enough to give a reliable estimate of the mean reflectivity magnitude, and is safe to use, since target and shadow appear to be located outside this region for the entire data set. Table I shows the average target-to-clutter ratio achieved by the conventional and the point-enhanced reconstructions over the 72 images for each target type. These results indicate a clear improvement of the target-to-clutter ratio by point-enhanced imaging.

Average Target-to-Clutter Ratio	T72	BMP2	BTR70
Conventional	31.88 dB	28.92 dB	26.92 dB
Proposed	88.28 dB	85.38 dB	82.62 dB

TABLE I

Average target-to-clutter ratios of images reconstructed from full-resolution data.

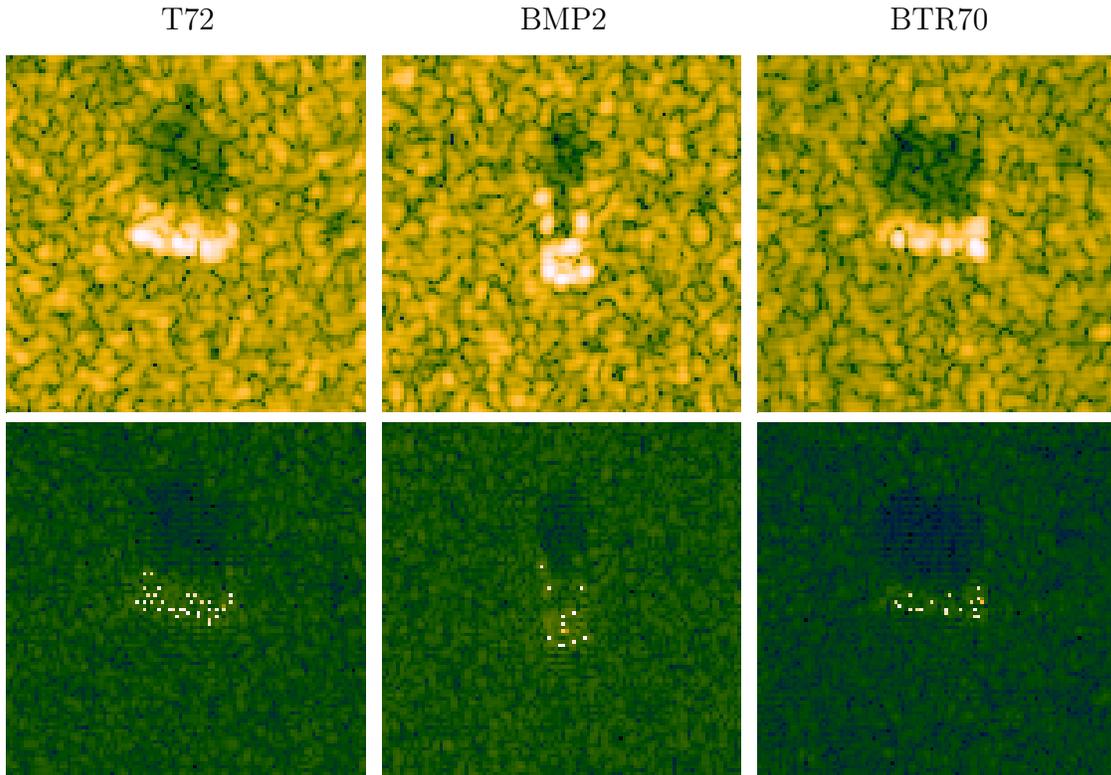


Fig. 9. 100×100 sample images reconstructed from reduced-resolution (50×50) data. Top: conventional. Bottom: proposed point-enhanced, with $p = 0.8$, $\lambda_1 = 4$, $\lambda_2 = 0$.

C. Point-Enhanced Superresolution Imaging from Reduced-Resolution Data

In this section, we report the results of experiments on two sets of data: the actual MSTAR data, and the synthetic point scatterer scenes constructed using the MSTAR images. The reason for using synthetic examples is to demonstrate the superresolution properties of point-enhanced imaging in a situation where the ground truth is exactly known.

For experiments on actual MSTAR data, we form images from a 50×50 subset of the 100×100 phase history samples previously used. This results in a 2-to-1 resolution loss in the range and cross-range directions. Hence, the resolution supported by such reduced data is 0.6 m. All the images we present in this section are composed of 100×100 pixels. The top row in Figure 9 shows Taylor weighted Fourier images reconstructed from the reduced-resolution, 50×50 data. The resolution loss in these images is evident when compared to their high-resolution counterparts in Figure 8. We form point-enhanced images with $p = 0.8$, $\lambda_1 = 4$, and $\lambda_2 = 0$ in (2), samples of which are shown in the bottom row of Figure 9.

We now consider the synthetic examples. To generate synthetic scenes, we find the 20 peaks with the largest magnitude in each of the 72 100×100 Taylor-windowed T72 Fourier images, and form a synthetic scene by placing simulated point-scatterers at the locations of these peaks, with the original complex reflectivities, and zeros in the rest of the scene. Hence these synthetic scenes do not contain any clutter. Our goal in using such simple scenes is to observe the consequence of resolution loss on point scatterers. The impact of the presence of clutter in real scenes could be observed through our experiments with actual MSTAR data. An example contour plot of the magnitude of such a synthetic scene is shown in the left third of Figure 10. We then generate simulated phase histories from this scene. The reconstructed conventional Taylor-windowed image from 50×50 phase history samples is shown in the top row, middle column of Figure 10. The loss of resolution is easy to observe. The corresponding point-enhanced image ($p = 0.8$, $\lambda_1 = 2$, and $\lambda_2 = 0$) is shown in the bottom row, middle column of the same figure. We can visually observe that most of the scatterers that were merged by the conventional reconstruction are now resolved. The images in the rightmost column demonstrate similar results for the 25×25 data case. Although the amount of data we use here is only one sixteenth of the data required for full resolution, point-enhanced imaging appears to be able to localize most of the scatterers.

In the remainder of this section, we quantitatively demonstrate the resolution improvement achieved by the real and synthetic images presented here. We do not present the target-to-clutter ratio results, since they are very similar to the full-resolution target-to-

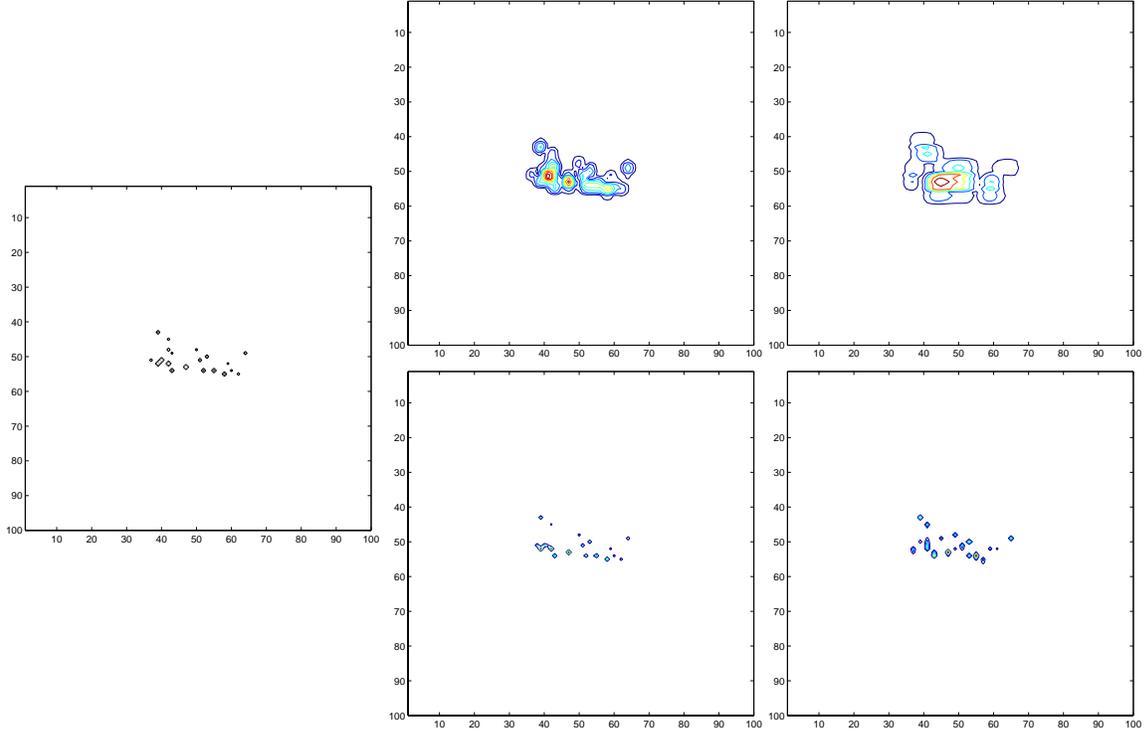


Fig. 10. Synthetic T72 image reconstruction example from reduced-resolution data. Left: ground truth. Middle: results from 50×50 data. Right: results from 25×25 data. Top row: conventional. Bottom row: proposed point-enhanced, with $p = 0.8$, $\lambda_1 = 2$, and $\lambda_2 = 0$.

Average Mainlobe Width	T72	BMP2	BTR70
Conventional	0.596 m	0.621 m	0.612 m
Proposed	0.131 m	0.129 m	0.125 m

TABLE II

Average mainlobe widths of MSTAR images reconstructed from reduced-resolution (50×50) data.

clutter ratio results of Table I.

C.1 Mainlobe Width

We compute the average 3-dB mainlobe width as described in Section III-A.2. The results in Table II for 50×50 data reconstructions of MSTAR images show that feature-enhanced imaging is able to reduce the mainlobe width considerably. To put these numbers in perspective, note that the resolution supported by the data is 0.6 m in this experiment.

C.2 Peak Matching Accuracy

We now evaluate the preservation of the locations of the dominant peaks in reduced-resolution data situations, using both real and synthetic data. For the real data examples, we use the locations of the 20 peaks extracted from the Taylor-windowed images reconstructed from full-resolution data, as the “reference” locations. Figure 11 provides a visual flavor of the peak matching experiment for the conventional and point-enhanced reconstructions from 50×50 data. One clear observation we can make out of these images is that, since conventional image formation causes peaks to merge, some of the peaks in the target area are lost, and peaks outside this area may become dominant. Figure 12 contains a zoomed-in version of the peak extraction results for one of the targets in Figure 11. We now compare the peak matching accuracy of point-enhanced and conventional images by using the criterion described in Section III-A.3. In Figure 13, we plot the average number of peak matches for each reconstruction method as a function of the radius r within which a match declaration is allowed. The standard deviation of this estimate of the mean is very small, hence we do not show error bars on these plots. The peak matching accuracy of point-enhanced images appear to be higher than that of the conventional images. Note that our analysis is based on finding peaks all around the scene. Alternatively, the search for peaks can be done in a pre-determined target region only. In that case we wouldn’t expect the conclusion of our analysis to change considerably, but the number of peak matches for conventional images would reach the number for point-enhanced images at a smaller value of the radius r , due to forced associations within the target region, possibly with spurious peaks.

We now report the results of similar experiments for the synthetic T72 scenes. Figure 14 shows the extracted peak locations on sample reconstructions, where the improved accuracy provided by point-enhanced imaging can be visually observed. The peak matching accuracy results from the entire data, shown in Figure 15, quantitatively verify the superresolution properties of point-enhanced imaging.

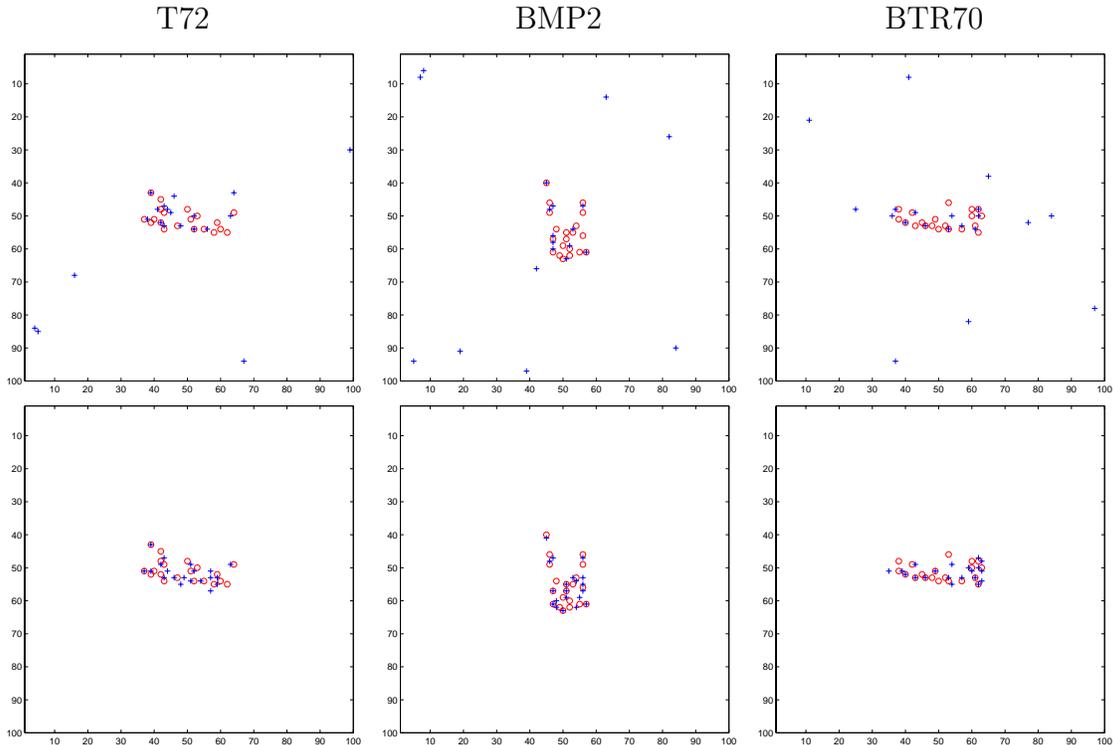


Fig. 11. Sample peak extraction results for MSTAR images reconstructed from 50×50 data. Circles indicate “reference” dominant scatterer locations extracted from full-resolution conventional images. Plus signs indicate peaks extracted from the reconstructed images. Top: conventional. Bottom: proposed point-enhanced, with $p = 0.8$.

C.3 Average Associated Peak Distance

We now compute the average distance between the 20 pairs of reference and extracted peaks, as described in Section III-A.4. Tables III and IV illustrate the average associated peak distances for the real and synthetic scene experiments respectively. These results indicate that point-enhanced imaging provides a clear reduction in the distances between the peak locations extracted from limited-data reconstructions and the true peak locations.

Note that this criterion computes distances for the case where all peaks in the reconstruction are associated with a reference peak. One might generalize this criterion to compute and plot the average associated peak distances as a function of the number of matched peaks, by varying the radius within which a match declaration is allowed. This generalized criterion would in some sense be the dual version of the analysis in Section V-C.2, and we would expect a similar performance difference between point-enhanced and

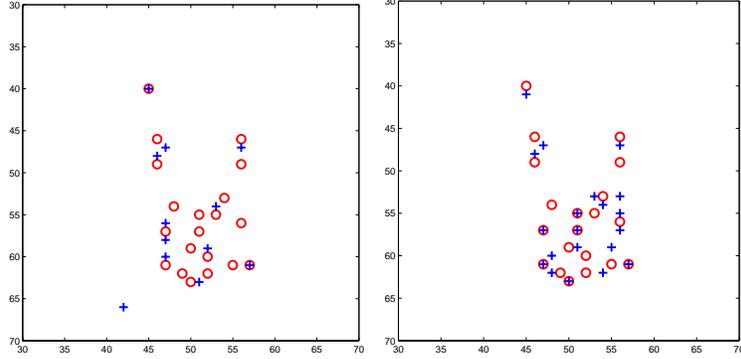


Fig. 12. Zoomed-in versions of the BMP2 peak extraction results from Figure 11. Left: conventional. Right: proposed point-enhanced, with $p = 0.8$.

Average Associated Peak Distance	T72	BMP2	BTR70
Conventional	3.13 m	4.32 m	4.30 m
Proposed	0.82 m	1.06 m	1.25 m

TABLE III

Average associated peak distances in MSTAR images reconstructed from reduced-resolution (50×50) data.

conventional images.

D. Region-Enhanced Imaging

We now compare region-enhanced images with conventional ones in terms of the criteria described in Section III-B. Here, we form 128×128 images from 128×128 Taylor-windowed phase history samples. In feature-enhanced image reconstruction, we use $p = 1$, and set $\lambda_2 > \lambda_1$ in (2) to enhance regions, in particular we choose $\lambda_1 = 0.1$ and $\lambda_2 = 0.3$. The second row of Figure 16 shows sample region-enhanced reconstructions. In contrast to the conventional images in the top row, these reconstructions reduce variability in homogeneous regions, while preserving discontinuities at region boundaries.

D.1 Speckle Suppression

We quantify the speckle amplitude in images as described in Section III-B.1, by using the bottom 20 rows (2560 pixels) of the reconstructed images as the clutter region. The results

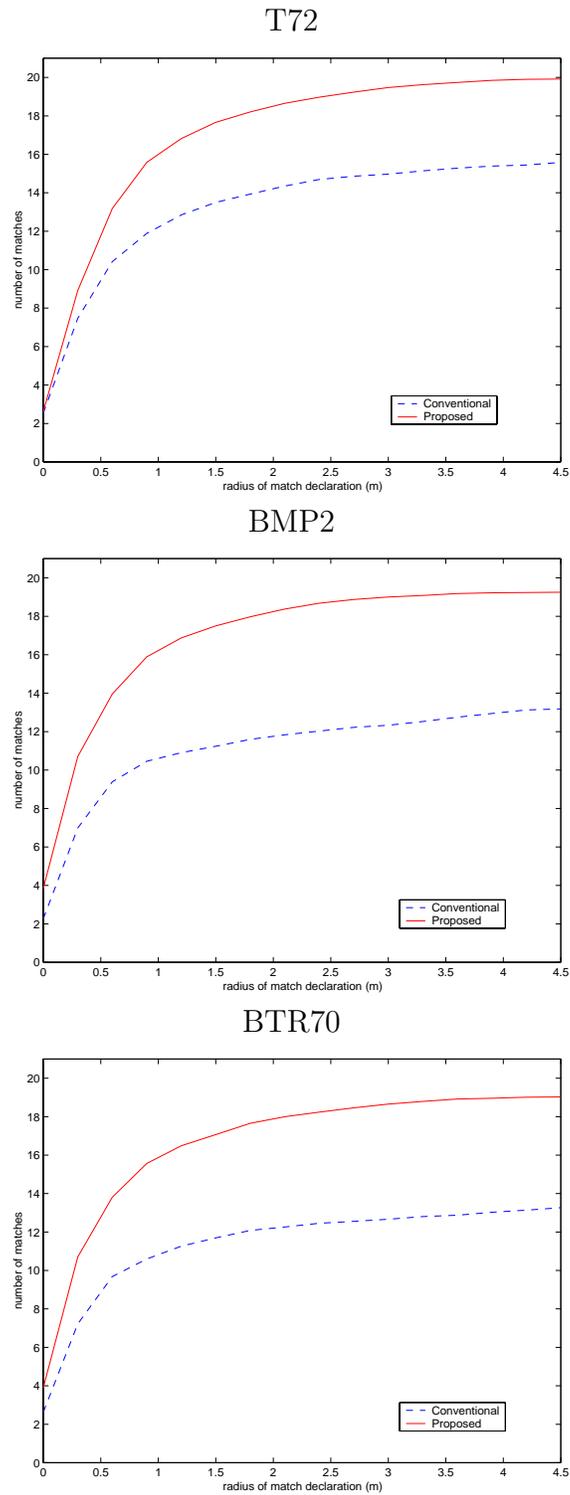


Fig. 13. Average number of peak matches in MSTAR images reconstructed from reduced-resolution (50×50) data as a function of the radius of match declaration r .

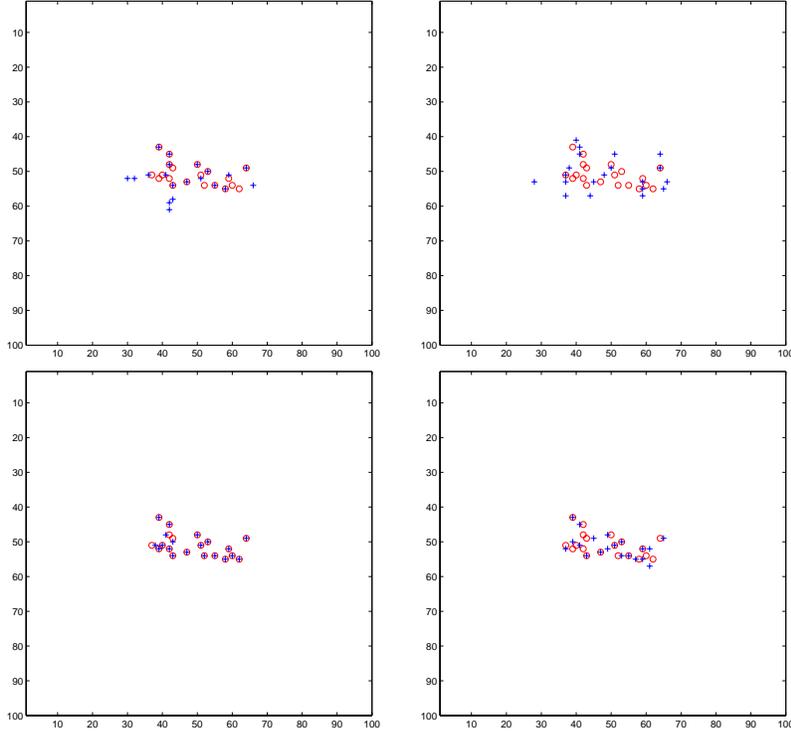


Fig. 14. Sample peak extraction results for the synthetic T72 scenes. Circles indicate the scatterer locations in the synthetic scene. Plus signs indicate peaks extracted from the reconstructed images. Left: 50×50 data. Right: 25×25 data. Top: conventional. Bottom: proposed point-enhanced, with $p = 0.8$.

Average Associated Peak Distance (synthetic T72)	50×50 data	25×25 data
Conventional	1.22 m	1.35 m
Proposed	0.07 m	0.61 m

TABLE IV

Average associated peak distances in the synthetic T72 reconstructions.

in Table V illustrate the improved speckle suppression achieved by the region-enhanced reconstructions.

D.2 Segmentation Accuracy

We now demonstrate that region-enhanced images simplify segmentation of the images into target, shadow and background regions. In our evaluation, we use the AFRL human segmentations of the MSTAR images [31] as the ground truth. Samples of such human

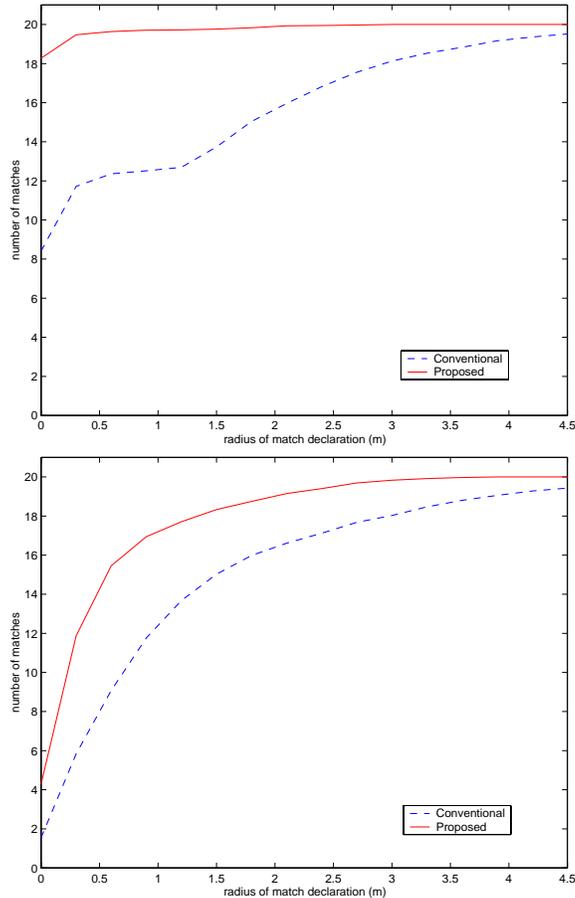


Fig. 15. Average number of peak matches for the synthetic T72 scenes as a function of the radius of match declaration r . Top: 50×50 data. Bottom: 25×25 data.

segmentations are shown in the third row of Figure 16. Naturally, these segmentations are themselves not perfect.

We segment the region-enhanced images by simple adaptive thresholding, as described in Section III-B.2, using $c_1 = 1.2$ and $c_2 = 2.5$. Sample results of such processing are shown in the bottom row of Figure 16. These results suggest that segmentation is considerably simplified by region-enhanced imaging. If such thresholding-based segmentation were applied to conventional images, the result would be dominated by fluctuations in homogeneous regions, as shown in the fourth row of Figure 16. For the conventional images we have used $c_1 = c_2 = 1$, which appears to be a good choice based on visual assessment.

In Table VI, we present the average percentage of accurately classified pixels in adaptive thresholding-based segmentations of conventional and region-enhanced images using the

Average Speckle Amplitude	T72	BMP2	BTR70
Conventional	5.919 dB	5.921 dB	5.898 dB
Proposed	2.261 dB	2.283 dB	2.269 dB

TABLE V

Average speckle amplitude in the dB-valued conventional and region-enhanced images.

Average Segmentation Accuracy	T72	BMP2	BTR70
Conventional	77.49 %	76.11 %	76.79 %
Proposed	96.85 %	97.28 %	97.64 %

TABLE VI

Average segmentation accuracy based on adaptive thresholding for conventional and region-enhanced images, measured as the percentage of correctly classified pixels.

entire data set. As expected, region-enhanced images have significantly higher accuracy. Note that we are in no way suggesting that adaptive thresholding is the state-of-the-art for segmenting conventional SAR images. In fact, conventional images can be more accurately segmented using advanced segmentation tools. Our point here is that even such a simple segmentation approach results in quite reasonable results for region-enhanced images.

We should note that the major error for region-enhanced images is due to the gap between the target and shadow regions in the segmentations, shown in the bottom row of Figure 16. This is a systematic error and may be improved upon by incorporation of additional information during segmentation. Our error analysis in Table VI has the limitation that the human segmentations, which we use as the truth, are really not perfect. We should also note that the feature-enhancement parameters λ_1 , λ_2 , and the thresholding parameters c_1 , c_2 have not been optimized for best segmentation performance, but rather picked based on visual assessment of one image, and applied to the processing of the entire data set.

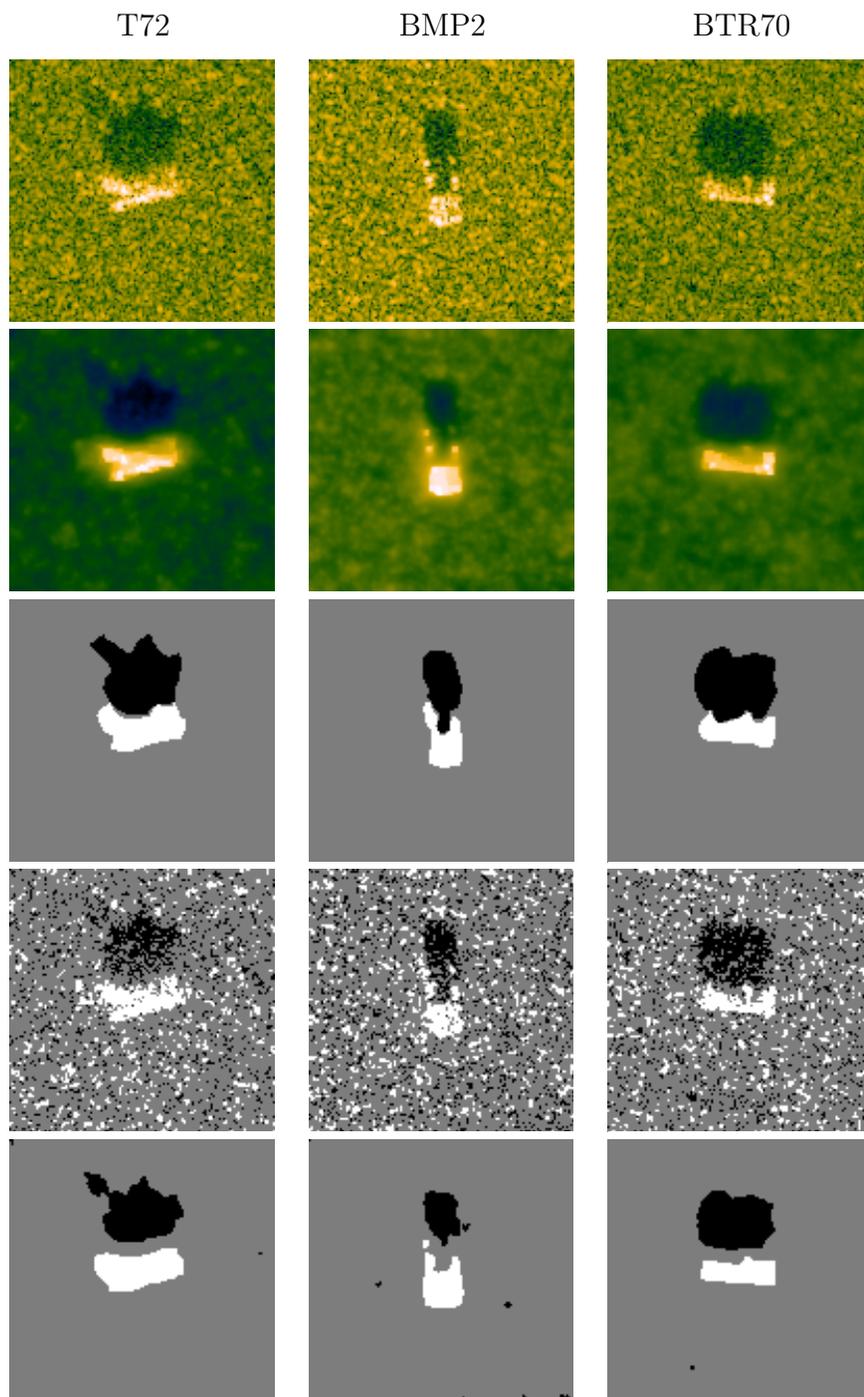


Fig. 16. Region-enhanced image formation and segmentation. Top row: MSTAR images. Second row: region-enhanced reconstructions with $p = 1$, $\lambda_1 = 0.1$ and $\lambda_2 = 0.3$. Third row: human segmentations. Fourth row: threshold-based segmentations of conventional imagery. Bottom row: threshold-based segmentations of region-enhanced imagery.

Bhattacharyya Distances	Targ.-Backg.	Targ.-Shad.	Backg.-Shad.
Conventional	0.28	0.76	0.25
Region-Enhanced	1.48	1.81	0.45

TABLE VII

Bhattacharyya distances between Gaussian densities characterizing different regions in SAR images.

D.3 Statistical Separability of Regions

We now provide a measure of the separability of different regions from each other in conventional and region-enhanced images by using Bhattacharyya distances as described in Section III-B.3. To this end, we treat the human segmented images as the truth, and extract region labels from these segmentations for each location in the scene.

Figure 17 illustrates the estimated Gaussian densities for the target, shadow and background regions of the conventional and region-enhanced images. By visual inspection, we can claim that the densities for the region-enhanced images are farther from each other than those for the conventional ones. We quantify this claim by computing the Bhattacharyya distances between region pairs. Table VII contains the results of this computation, which demonstrate that the Bhattacharyya distances between all region pairs are larger for the region-enhanced images than the conventional images.

VI. EXPERIMENTAL RESULTS FOR ATR PERFORMANCE

A. *Experimental Setup*

We now evaluate the performance of the classifiers described in Section IV, given MSTAR images produced by feature-enhanced versus conventional imaging. In our recognition experiments, we use the T72, BMP2 and BTR70 targets. For all three classifiers, our training set is composed of images at 17° depression angle, and our test set is composed of an independent set of images at 15° depression angle. The numbers of each type of target images in each of these sets are shown in Table VIII. In training and testing with all three classifiers, we extract and use a near-central portion of the SAR image which contains the target. In all cases, we normalize the images so that they have the same

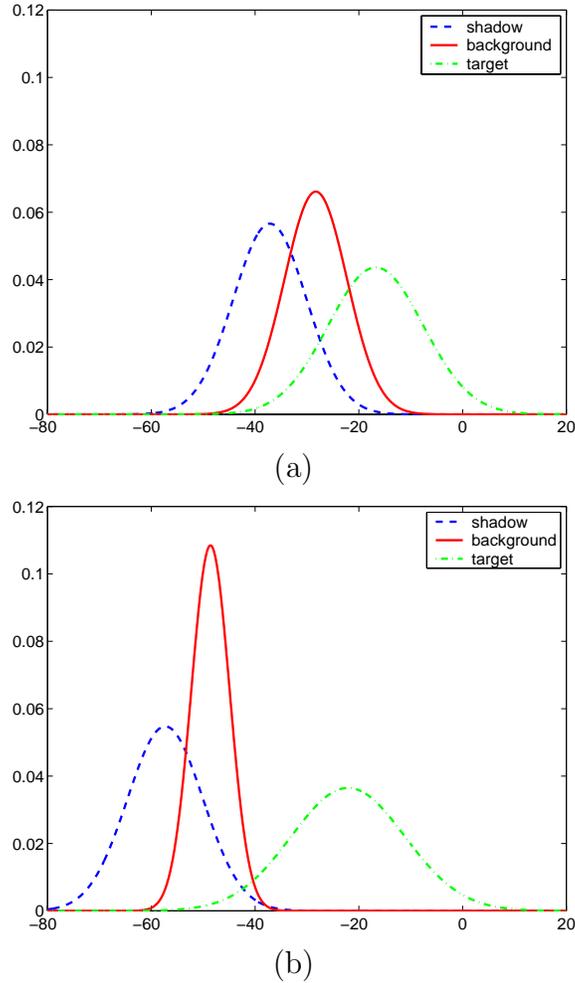


Fig. 17. Gaussian probability density functions for regions in SAR images. (a) Conventional images. (b) Region-enhanced images.

ℓ_2 -norm, before using them in the classifiers.

We present the results of our evaluations in the form of classifier confusion matrices, which show the number of correct and incorrect classifications achieved on test inputs of each type. A single number characterizing the classifier’s ability to recognize test inputs can be obtained through the probability of correct classification, P_{cc} , which is defined as the fraction of all target test inputs that were correctly classified.

B. Template-based Classification Results

We now present the recognition performance of the conventional, point-enhanced (non-superresolution) and region-enhanced images, when they are used as inputs to the template-

Target	Training set		Test set	
	Depression	No. of images	Depression	No. of images
T72	17°	232	15°	196
BMP2	17°	233	15°	195
BTR70	17°	233	15°	196

TABLE VIII

Composition of the MSTAR data set used in recognition experiments.

based classifier, described in Section IV-A.

We have initially conducted recognition experiments with images reconstructed from high-SNR data. By high-SNR data, we mean the SAR data obtained by undoing the image formation steps for the images in the MSTAR data set, without any additional measurement noise. We have observed that for this high-SNR, 3-target classification problem both conventional and feature-enhanced images result in a high recognition rate, and the performance difference is not significant. Next, we investigate the recognition performance of these images in the face of degraded data, which may provide a better representation of a practical situation. To this end, we have included additive complex Gaussian noise in the projectional SAR data. Treating the original data before the addition of this noise as the clean signal, the SNR of the corrupted data we use is -4 dB. This represents a very noisy scenario.

We now present the recognition results for this low-SNR case. Table IX shows the confusion matrices for the classification of conventional and feature-enhanced images. The conventional images result in an overall probability of correct classification of 69.85%. Point-enhanced and region-enhanced imaging increase this rate to 88.93% and 96.76% respectively. These results suggest that noise causes a much smaller degradation in recognition performance for feature-enhanced images, as compared to conventional images. Hence feature-enhanced imaging is more robust to limitations in data quality.

	T72	BMP2	BTR70
T72	159	30	7
BMP2	53	120	22
BTR70	33	32	131

(a) Conventional images. $P_{cc} = 69.85\%$

	T72	BMP2	BTR70
T72	176	13	7
BMP2	8	173	14
BTR70	3	20	173

(b) Point-enhanced images. $P_{cc} = 88.93\%$

	T72	BMP2	BTR70
T72	184	11	1
BMP2	3	191	1
BTR70	1	2	193

(c) Region-enhanced images. $P_{cc} = 96.76\%$

TABLE IX

Confusion matrices summarizing the template-based classification results. The entry in row i , column j shows the number of images from vehicle type i classified as vehicle j .

C. Likelihood-based Classification Results

We now present the results of running the low-SNR data used in Section VI-B through the likelihood-based classifier described in Section IV-B. The results in Table X indicate that the classifier has a correct classification rate of 87.05% with conventional images, 94.38% with point-enhanced images, and 99.15% with region-enhanced images.

D. Point-feature-based Classification Results

We now report the results of recognition experiments using the classifier of Section IV-C. In order to investigate the preservation of point features in the face of reduced-resolution

	T72	BMP2	BTR70
T72	170	18	8
BMP2	7	161	27
BTR70	4	12	180

(a) Conventional images. $P_{cc} = 87.05\%$

	T72	BMP2	BTR70
T72	190	5	1
BMP2	10	178	7
BTR70	8	2	186

(b) Point-enhanced images. $P_{cc} = 94.38\%$

	T72	BMP2	BTR70
T72	193	3	0
BMP2	1	194	0
BTR70	1	0	195

(c) Region-enhanced images. $P_{cc} = 99.15\%$

TABLE X

Confusion matrices summarizing the likelihood-based classification results. The entry in row i , column j shows the number of images from vehicle type i classified as vehicle j .

data, we use 50×50 phase history data with a resolution of 0.6 m .⁴ The data used for these low resolution experiments do not contain any noise in addition to what is already present in the MSTAR data. Hence, we consider a high-SNR, but limited-data scenario here. Since the classifier we consider here explicitly uses point-based features, the question of interest is whether enhancing point-based features during the imaging process improves the recognition performance. For that reason, our evaluation in this section involves only point-enhanced imagery and not region-enhanced imagery.

We obtain the true peak locations for all targets at all the orientations available, using conventionally formed SAR images at 0.3 m resolution, and 17° depression angle. We

⁴Note that the data used in Sections VI-B and VI-C had a resolution of 0.3 m .

	T72	BMP2	BTR70
T72	38	84	74
BMP2	14	114	67
BTR70	4	81	111

(a) Conventional images. $P_{cc} = 44.80\%$

	T72	BMP2	BTR70
T72	162	23	11
BMP2	28	147	20
BTR70	10	17	169

(b) Point-enhanced, superresolution images. $P_{cc} = 81.43\%$

TABLE XI

Confusion matrices summarizing the point-feature-based classification results, from reduced-resolution data. The entry in row i , column j shows the number of images from vehicle type i classified as vehicle j .

then run the classifier on conventional and point-enhanced, superresolution images at 15° depression angle, reconstructed from data supporting a resolution of 0.6 m. The number of dominant scatterers we consider for each image is $N_p = 20$.

The classification results presented in Table XI demonstrate a clear improvement in recognition performance through the use of point-enhanced, superresolution images as compared to the case of conventional images. The correct classification rates for the conventional and point-enhanced images are 44.80%, and 81.43%, respectively.

E. Further Analysis and Extensions

We have presented some recognition experiments to compare the performance of feature-enhanced and conventional images. Although this analysis has provided a flavor of the impact of feature-enhanced imaging, more extensive experimentation on a more general setting would be an important contribution for a number of reasons. First, the recognition problem posed in this paper contained three classes. This is too small to represent a practical setting where the decisions usually involve a larger number of vehicle types.

Furthermore, our experiments did not involve *confuser* vehicles. Confuser vehicles are test inputs, for which no training is done, and the ATR system should ideally respond as “unknown” to data containing such vehicles. This is an important practical issue, and a more realistic recognition evaluation should involve confuser vehicles. In addition, one might include SAR images of natural clutter and urban scenes in the experiments. Finally, although the classifiers that we considered were mostly based on real recognition systems in use, there are more advanced, feature-based systems under development, such as that associated with the MSTAR program, as described in [15, 35]. Evaluation of the performance of feature-enhanced SAR images directly in such a feature-based recognition system would be an important extension of the work presented in this paper.

Another line of future work might be to experimentally compare feature-enhanced imaging to other advanced SAR imaging techniques, such as those proposed in [17, 21], which have also been found to result in better ATR performance [20, 22] than conventional images.

VII. CONCLUSIONS

We have presented an analysis demonstrating the impact of a recently proposed, non-quadratic optimization-based SAR image formation technique on feature enhancement and ATR performance.

We have demonstrated the feature enhancement properties through a variety of quantitative criteria. The results indicate that the images produced by the new image formation method exhibit superresolution and improved localization accuracy for dominant scatterers, and improved separability for different regions in the scene. It might be useful to compare our results to those in similar analyses carried out for other enhanced SAR image formation techniques (e.g. [17, 18, 33]).

We have conducted an evaluation of how the new feature-enhanced SAR imaging method affects ATR performance. We have defined a 3-class target recognition problem to compare the recognition performance of feature-enhanced images and conventional images. We have constructed two pixel-based classifiers, and one feature-based classifier. With high-SNR data we have not observed a significant performance difference between the conventional and the feature-enhanced images. With reduced-quality data however, feature-enhanced

High-resolution, low-SNR experiment			
	Conventional	Point-Enhanced	Region-Enhanced
Template-based classifier	69.85 %	88.93 %	96.76 %
Likelihood-based classifier	87.05 %	94.38 %	99.15 %

Reduced-resolution, high-SNR experiment		
	Conventional	Point-Enhanced, Superresolution
Point-feature-based classifier	44.80 %	81.43 %

TABLE XII

Overall summary of the classification experiments, in terms of the probability of correct classification, P_{cc} .

images resulted in higher recognition rates. We have also conducted reduced-resolution data experiments, and used the feature-based classifier to test the performance of feature-enhanced, superresolution images. We have observed recognition improvements as compared to the conventional images in this case, as well. Table XII summarizes the recognition results obtained in this paper in terms of the probability of correct classification.

VIII. ACKNOWLEDGMENTS

We would like to thank Gil Ettinger for discussions about the work in this paper and related future experiments; Zhengrong Ying for letting us use his feature matching code; and Robert A. Weisenseel for his help with the human segmented data.

REFERENCES

- [1] M. Çetin, W. C. Karl, and D. A. Castañon, "Evaluation of a regularized SAR imaging technique based on recognition-oriented features," in *Algorithms for Synthetic Aperture Radar Imagery VII*, E. G. Zelnio, Ed., Orlando, FL, USA, Apr. 2000, vol. 4053 of *Proc. SPIE*, pp. 40–51.
- [2] M. Çetin, W. C. Karl, and D. A. Castañon, "Analysis of the impact of feature-enhanced SAR imaging on ATR performance," in *Algorithms for Synthetic Aperture Radar Imagery IX*, E. G. Zelnio, Ed., Orlando, FL, USA, Apr. 2002, vol. 4727 of *Proc. SPIE*, pp. 134–145.
- [3] B. Bhanu, "Automatic target recognition: state of the art survey," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 4, pp. 364–379, July 1986.
- [4] J. A. Ratches, C. P. Walters, and R. G. Buser, "Aided and automatic target recognition based upon sensory

- inputs from image forming systems,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 9, pp. 1004–1019, Sept. 1997.
- [5] R. Hummel, “Model-based ATR using synthetic aperture radar,” in *IEEE International Radar Conf.*, 2000, pp. 856–861.
- [6] L. M. Novak, “State-of-the-art of SAR automatic target recognition,” in *IEEE International Radar Conf.*, 2000, pp. 836–843.
- [7] J. Walker, “Range-Doppler imaging of rotating objects,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-16, no. 1, pp. 23–52, Jan. 1980.
- [8] D. C. Munson Jr. and J. L. C. Sanz, “Image reconstruction from frequency-offset Fourier data,” *Proc. IEEE*, vol. 72, no. 6, pp. 661–669, June 1984.
- [9] J. W. Goodman, “Statistical properties of laser speckle patterns,” in *Laser Speckle and Related Phenomena*, J. C. Dainty, Ed., pp. 9–75. Springer-Verlag, Berlin, 1975.
- [10] M. Çetin and W. C. Karl, “Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization,” *IEEE Trans. Image Processing*, vol. 10, no. 4, pp. 623–631, Apr. 2001.
- [11] L. C. Potter and R. L. Moses, “Attributed scattering centers for SAR ATR,” *IEEE Trans. Image Processing*, vol. 6, no. 1, pp. 79–91, Jan. 1997.
- [12] G. Jones, III and B. Bhanu, “Recognition of articulated and occluded objects,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, no. 7, pp. 603–613, July 1999.
- [13] G. Jones, III and B. Bhanu, “Recognizing occluded objects in SAR images,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 1, pp. 316–328, Jan. 2001.
- [14] H. Chiang, R. L. Moses, and L. C. Potter, “Classification performance prediction using parametric scattering feature models,” in *Algorithms for Synthetic Aperture Radar Imagery VII*, E. G. Zelnio, Ed., Orlando, FL, USA, Apr. 2000, vol. 4053 of *Proc. SPIE*, pp. 546–557.
- [15] W. W. Irving and G. J. Ettinger, “Classification of targets in synthetic aperture radar imagery via quantized grayscale matching,” in *Algorithms for Synthetic Aperture Radar Imagery VI*, E. G. Zelnio, Ed., Orlando, FL, USA, Apr. 1999, vol. 3721 of *Proc. SPIE*, pp. 320–331.
- [16] C. F. Olson and D. P. Huttenlocher, “Automatic target recognition by matching oriented edge pixels,” *IEEE Trans. Image Processing*, vol. 6, no. 1, pp. 103–113, Jan. 1997.
- [17] G. R. Benitz, “High-definition vector imaging,” *Lincoln Laboratory Journal*, vol. 10, no. 2, pp. 147–170, 1997.
- [18] W. Phillips, S. DeGraaf, and R. Chellappa, “Enhanced segmentation of SAR images using non-Fourier imaging,” in *IEEE International Conference on Image Processing*, Chicago, IL, USA, 1998, vol. 1, pp. 583–586.
- [19] “Air Force Research Laboratory, Model Based Vision Laboratory, Sensor Data Management System MSTAR Web Page: <http://www.mbvlab.wpafb.af.mil/public/sdms/datasets/mstar/>,” .
- [20] L. M. Novak, G. J. Owirka, and A. L. Weaver, “Automatic target recognition using enhanced resolution SAR data,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 1, pp. 157–175, Jan. 1999.
- [21] S. R. Degraaf, “SAR imaging via modern 2-D spectral estimation methods,” *IEEE Trans. Image Processing*, vol. 7, no. 5, pp. 729–761, May 1998.
- [22] G. J. Owirka, S. M. Verbout, and L. M. Novak, “Template-based SAR ATR performance using different image enhancement techniques,” in *Algorithms for Synthetic Aperture Radar Imagery VI*, E. G. Zelnio, Ed., Orlando, FL, USA, Apr. 1999, vol. 3721 of *Proc. SPIE*, pp. 302–319.

- [23] L. M. Novak, S. D. Halversen, G. J. Owirka, and M. Hiett, "Effects of polarization and resolution on the performance of a SAR ATR system," *Lincoln Laboratory Journal*, vol. 8, no. 1, pp. 49–67, 1995.
- [24] L. M. Novak, G. J. Owirka, W. S. Brower, and A. L. Weaver, "The automatic target-recognition system in SAIP," *Lincoln Laboratory Journal*, vol. 10, no. 2, pp. 187–202, 1997.
- [25] J. A. O'Sullivan, M. D. Devore, V. Kedia, and M. I. Miller, "SAR ATR performance using a conditionally Gaussian model," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 1, pp. 91–108, Jan. 2001.
- [26] C. V. Jakowatz Jr., D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-mode Synthetic Aperture Radar: a Signal Processing Approach*, Kluwer Academic Publishers, Norwell, MA, 1996.
- [27] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.
- [28] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Image Processing*, vol. 7, no. 6, pp. 813–824, June 1998.
- [29] M. Hazlett, D. J. Andersh, S. W. Lee, H. Ling, and C. L. Yu, "XPATCH: a high-frequency electromagnetic scattering prediction code using shooting and bouncing rays," in *Targets and Backgrounds: Characterization and Representation*, W. R. Watkins and D. Clement, Eds., Orlando, FL, USA, Apr. 1995, vol. 2469 of *Proc. SPIE*, pp. 266–275.
- [30] Z. Ying and D. A. Castañón, "Statistical model for occluded object recognition," in *IEEE International Conference on Information, Intelligence and Systems*, 1999, pp. 324–327.
- [31] G. J. Power and R. A. Weisenseel, "ATR subsystem performance measures using manual segmentation of SAR target chips," in *Algorithms for Synthetic Aperture Radar Imagery VI*, E. G. Zelnio, Ed., Apr. 1999, vol. 3721 of *Proc. SPIE*, pp. 685–692.
- [32] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, New York, 1990.
- [33] A. E. Brito, S. H. Chan, and S. D. Cabrera, "SAR image formation using 2-D re-weighted minimum norm extrapolation," in *Algorithms for Synthetic Aperture Radar Imagery VI*, E. G. Zelnio, Ed., 1999, vol. 3721 of *Proc. SPIE*, pp. 78–90.
- [34] M. Çetin, *Feature-Enhanced Synthetic Aperture Radar Imaging*, Ph.D. thesis, Boston University, Feb. 2001.
- [35] K. M. Simonson, "Multinomial pattern matching: A robust technique for target identification," in *Proceedings of the Automatic Target Recognizer Working Group*, Huntsville, AL, 1997.