

Performance-Complexity Tradeoffs for Several Approaches to ATR from SAR Images

Joseph A. O'Sullivan^a and Michael D. DeVore^b

Electronic Systems and Signals Research Laboratory,
Department of Electrical Engineering, Washington University, St. Louis, MO 63130
^ajao@ee.wustl.edu, ^bmdd2@cis.wustl.edu

ABSTRACT

The performance of an automatic target recognition (ATR) system for synthetic aperture radar (SAR) images is generally dependent upon a set of parameters which captures the assumptions and approximations made in the implementation of the system. This set of parameters implicitly or explicitly determines a level of database complexity for the system. A comprehensive analysis of the empirical tradeoffs between ATR performance and database complexity is presented for variations of several algorithms including a likelihood approach under a conditionally Gaussian model for pixel distribution, a mean squared error classifier on pixel dB values, and a mean squared error classifier on pixel quarter power values. These algorithms are applied under a common framework to identical training and testing sets of SAR images for a wide range of system parameters. Their performance is characterized both in terms of the percentage of correctly classified test images and the average squared Hilbert-Schmidt distance between the estimated and true target orientations across all test images. Performance boundary curves are presented and compared, and algorithm performance is detailed at key complexity values. For the range of complexity considered, it is shown that in terms of target orientation estimation the likelihood based approach under a conditionally Gaussian model yields superior performance for any given database complexity than any of the other approaches tested. It is also shown that some variant of each of the approaches tested delivers superior target classification performance over some range of complexity.

Keywords: synthetic aperture radar, automatic target recognition, MSTAR, performance-complexity

1. INTRODUCTION

Recent literature on the subject of automatic target recognition (ATR) contains many descriptions of systems which perform pose estimation and/or target classification from synthetic aperture radar (SAR) images of unknown targets. These systems are generally dependent upon some set of parameters which govern implementation-specific approximations and assumptions. These parameters include, for example, resolution in the slant plane, resolution in target pose with respect to the radar platform, parameters governing target chip extraction, and level of sophistication of the underlying model. Each combination of parameters which can be selected for an implementation simultaneously determines a benefit in terms of system performance and a cost in terms of system complexity. Knowledge of the relationship between performance and complexity makes possible cost-benefit analyses when selecting approaches and parameterizations for implementation. As discussed by Ross, et al.,⁸ it is important to understand the relationships between operating condition, cost, and performance.

In this paper, we look at variants of three families of approaches to ATR from SAR images. These families include the conditionally Gaussian model which O'Sullivan, et al.⁶ apply to ATR from SAR imagery, a least squared error approach similar to that used by Owirka and Novak⁷ and Novak, et al.³ that uses pixel magnitudes in the decibel scale, and a least squared error approach similar to that reported by Worrell, et al.⁹ that uses pixel quarter-power values. Algorithms for jointly estimating the pose and class of an unknown target from a SAR image are derived for each. The implementations of these algorithms are parameterized by the resolution in azimuth, the range of azimuth over which training data is selected for the i th resolution window, and the size of SAR image chip employed. The results of applying these algorithms to a ten class ATR problem using

data from the Moving and Stationary Target Acquisition and Recognition (MSTAR) program are presented for a wide range of parameterizations. The parameterizations chosen result in a range from low to high complexities of the databases required to support the algorithms and yield a range of performance results. We then present a characterization of the various algorithms in terms of the best performance achievable at any given database complexity. This characterization allows for a direct comparison of the behavior of different approaches under identical conditions and provides a means of quantifying the benefit provided by supporting increased complexity for any given algorithm.

Section 2 contains details of the variants to the three approaches examined in this paper. Section 3 describes the parameterization of the algorithms and introduces the measures of performance and complexity used. In Section 4, the results of applying the characterization method to each of the algorithms are presented. Conclusions follow in Section 5.

2. SAR DATA MODELS AND ALGORITHMS

2.1. Conditionally Gaussian Approach

The conditionally Gaussian model for SAR imagery is based upon a model for individual pixels in the images which holds that every pixel in a SAR image is well described as a complex Gaussian distributed random variable, conditioned on both target type, a , and pose, Θ , relative to the radar platform. Further, each pixel in the image is modeled as independent from other pixels in the image and having a complex mean that is identically equal to zero. Thus, the log-likelihood of any received SAR image \mathbf{r} is dependent upon only the variance of the individual pixels as a function of target type and pose and it can be written as a summation over all pixels i as

$$l(\mathbf{r}|\Theta, a) = - \sum_i \left[\ln(\sigma_i^2(\Theta, a)) + \frac{|r_i|^2}{\sigma_i^2(\Theta, a)} \right]. \quad (1)$$

Given a set of training images, $\mathcal{I}_{k,l}$, of target type a_l formed with the target at pose Θ_k , the variance function from (1) can be estimated for that combination of parameters as

$$\hat{\sigma}_i^2(\Theta_k, a_l) = \frac{1}{|\mathcal{I}_{k,l}|} \sum_{\mathbf{r} \in \mathcal{I}_{k,l}} |r_i|^2. \quad (2)$$

Under this model, ATR can be performed by jointly estimating the target pose via maximum *a posteriori* (MAP) estimation and determining a target type via the generalized likelihood ratio test (GLRT). An optimal classifier would be constructed using actual pixel variance functions. Assuming, however, that only the estimated variances are available we substitute the estimated variances and perform joint pose estimation and classification according to

$$[\hat{\Theta}, \hat{a}] = \operatorname{argmax}_{\Theta, a} - \sum_i \left[\ln(\hat{\sigma}_i^2(\Theta, a)) + \frac{|r_i|^2}{\hat{\sigma}_i^2(\Theta, a)} \right]. \quad (3)$$

2.2. Log-Magnitude Approach

The log-magnitude model for SAR imagery is based upon a model for individual pixels under which pixel squared-magnitude is taken to follow a log-normal distribution so that the decibel representation follows a Gaussian distribution conditioned on target type, a , and pose, Θ . Further, the pixels are modeled as independent from one to another, possessing a noise-induced variance that has a constant value across all pixels, target types, and poses. Under this model, given the decibel representation of a received SAR image \mathbf{r} , $\mathbf{r}_{\text{dB}} = 20 \log |\mathbf{r}|$ where $|\cdot|$ denotes element-by-element magnitude, and given the conditional mean vector $\boldsymbol{\mu}_{\text{dB}}(\Theta, a)$, the log-likelihood function is

$$l(\mathbf{r}|\Theta, a) = - \|\mathbf{r}_{\text{dB}} - \boldsymbol{\mu}_{\text{dB}}(\Theta, a)\|^2. \quad (4)$$

Given a set of training images, $\mathcal{I}_{k,l}$, of target type a_l formed with the target at pose Θ_k , the conditional mean vector can be estimated for that combination of parameters as

$$\hat{\boldsymbol{\mu}}_{\text{dB}}(\Theta_k, a_l) = \frac{1}{|\mathcal{I}_{k,l}|} \sum_{\mathbf{r} \in \mathcal{I}_{k,l}} 20 \log |\mathbf{r}|. \quad (5)$$

Joint pose estimation and target classification can be performed by maximizing the log-likelihood in (4) or, alternatively, by minimizing the squared distance between the received SAR image in decibel representation and the conditional mean. Since it is assumed that the conditional mean is not known, it is replaced with the estimate from (5), and decisions are made according to

$$\left[\hat{\Theta}, \hat{a} \right] = \underset{\Theta, a}{\operatorname{argmin}} \left\| \mathbf{r}_{\text{dB}} - \hat{\boldsymbol{\mu}}_{\text{dB}}(\Theta, a) \right\|^2. \quad (6)$$

Noting that performance of the radar transmitter and receiver may vary from platform to platform and even between successive uses of the same platform, normalization can be introduced to (6). One possible approach is to normalize the squared magnitude vector $|\mathbf{r}|^2$ by the geometric mean of its components. In the decibel domain, this is equivalent to subtracting the arithmetic mean of the components and decisions are made according to

$$\left[\hat{\Theta}, \hat{a} \right] = \underset{\Theta, a}{\operatorname{argmin}} \left\| (\mathbf{r}_{\text{dB}} - \bar{\mathbf{r}}_{\text{dB}}) - (\hat{\boldsymbol{\mu}}_{\text{dB}}(\Theta, a) - \bar{\hat{\boldsymbol{\mu}}}_{\text{dB}}(\Theta, a)) \right\|^2, \quad (7)$$

where $\bar{\mathbf{x}}$ denotes the arithmetic mean of the elements in vector \mathbf{x} . An alternative approach to normalization is to ensure that in the decibel domain all vectors have unit arithmetic mean and to make decisions according to

$$\left[\hat{\Theta}, \hat{a} \right] = \underset{\Theta, a}{\operatorname{argmin}} \left\| \frac{\mathbf{r}_{\text{dB}}}{\bar{\mathbf{r}}_{\text{dB}}} - \frac{\hat{\boldsymbol{\mu}}_{\text{dB}}(\Theta, a)}{\bar{\hat{\boldsymbol{\mu}}}_{\text{dB}}(\Theta, a)} \right\|^2. \quad (8)$$

2.3. Quarter-Power Approach

In the quarter-power approach to ATR from SAR imagery, the square root of pixel magnitude is approximated as a Gaussian distributed random variable conditioned on target type, a , and pose Θ . The pixels are modeled as independent from one another and having a noise-induced variance which is constant across all pixels, target types, and poses. The quarter-power representation of a received SAR image, $\mathbf{r}_{\text{QP}} = |\mathbf{r}|^{1/2}$ where $|\cdot|^{1/2}$ denotes the pixelwise square root of magnitude, has components which are all non negative. However, to the extent that the components can be approximated as conditionally Gaussian with conditional mean $\boldsymbol{\mu}_{\text{QP}}(\Theta, a)$, an approximate log-likelihood function is

$$l(\mathbf{r}|\Theta, a) = - \left\| \mathbf{r}_{\text{QP}} - \boldsymbol{\mu}_{\text{QP}}(\Theta, a) \right\|^2. \quad (9)$$

Given a set of training images, $\mathcal{I}_{k,l}$, of target type a_l formed with the target at pose Θ_k , the conditional mean vector can be estimated for that combination of parameters as

$$\hat{\boldsymbol{\mu}}_{\text{QP}}(\Theta_k, a_l) = \frac{1}{|\mathcal{I}_{k,l}|} \sum_{\mathbf{r} \in \mathcal{I}_{k,l}} |\mathbf{r}|^{1/2}. \quad (10)$$

Joint pose estimation and target classification can be performed by maximizing (9) over all combinations of target and pose for a given SAR image \mathbf{r} . Alternatively, decisions can be made by choosing the combination of parameters that minimize the distance between the quarter-power representation and the conditional mean vector as in

$$\left[\hat{\Theta}, \hat{a} \right] = \underset{\Theta, a}{\operatorname{argmin}} \left\| \mathbf{r}_{\text{QP}} - \boldsymbol{\mu}_{\text{QP}}(\Theta, a) \right\|^2. \quad (11)$$

Normalization can be performed to help reduce the effects of radar transmitter and receiver variation by ensuring that all vectors have unit magnitude, making decisions according to

$$\left[\hat{\Theta}, \hat{a} \right] = \underset{\Theta, a}{\operatorname{argmin}} \left\| \frac{\mathbf{r}_{\text{QP}}}{\|\mathbf{r}_{\text{QP}}\|} - \frac{\boldsymbol{\mu}_{\text{QP}}(\Theta, a)}{\|\boldsymbol{\mu}_{\text{QP}}(\Theta, a)\|} \right\|^2. \quad (12)$$

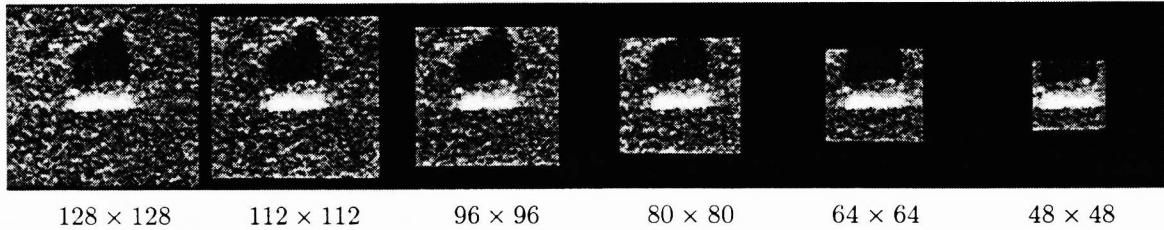


Figure 1. Variance images of the 2S1 facing broadside to the left of the image. Each image represents a different size from 128×128 to 48×48 .

3. ALGORITHM PARAMETERIZATION AND MEASUREMENTS

3.1. Parameters and Database Complexity

Two dimensions to the parameterization of these algorithms stem from Equations (2), (5), and (10). These equations express estimates of distribution properties from a set of training data, $\mathcal{I}_{k,l}$ taken from a target at pose Θ_k . Though target pose is continuous valued, training data are generally available only at a finite number of poses so that in these algorithms target pose is discretized. The number of discrete pose values w used to cover the range of representable poses is thus an implementation parameter of the algorithms. The range of representable poses associated with Θ_k we will call the representation window w_k . For purposes of demonstration, we will assume ground-based targets on level terrain imaged from a SAR platform with a known depression angle so that we may consider only the azimuth angle, $\theta \in [0, 2\pi)$. In the examples of Section 4, we consider a range of pose windows from $w = 1$, where a single template covering all 360° of aspect pose is used to represent a target, to $w = 72$, where each template represents a target over a 5° range of aspect pose. In each case, the azimuth θ_k represents the range of aspect angles $w_k = [2\pi k/w - \pi/w, 2\pi k/w + \pi/w)$ which are centered at $2\pi k/w$ and of width $2\pi/w$.

It is also unlikely that an amount of training data sufficient to accurately estimate the distribution parameters will be available from a single pose. Rather, the images comprising the sets $\mathcal{I}_{k,l}$ will be formed from some range of target pose centered around Θ_k which we will call the training window for Θ_k . The size of this range is independent of w and the regions from which training images are drawn may or may not overlap for neighboring windows in pose. The size of this range is another implementation parameter which must be chosen and in the examples of Section 4 we consider multiples of one, two, three, and four times the representation width $2\pi/w$. Choosing larger training window widths results in algorithms with less capability to represent small changes in target pose. For a fixed amount of training data, however, large training window widths can allow better estimates of the distribution parameters because more data is included in the sample averages. As will be demonstrated, for each of the algorithms of Section 2 there is a performance penalty for training window sizes that are either too large or too small.

A third dimension to the parameterization of these algorithms is the length of the SAR image vector \mathbf{r} . In general the length of this vector is governed by both SAR image resolution and parameters governing the extraction of the target from the rest of the image. In the available MSTAR data, only a single image resolution is represented so for the examples in Section 4 we consider only parameters governing the extraction of target regions. In order to produce extracted images which are reproducible and are not biased toward any particular algorithm, we extract a square region centered in the middle of the SAR image chip. We extract regions of six different image sizes 48×48 , 64×64 , 80×80 , 96×96 , 112×112 , and 128×128 . As can be seen in Figure 1, the smaller image sizes eliminate most of the background clutter near the target but also remove much of the target shadow and, in some cases, some of the pixels at the target extremes. The larger image sizes, on the other hand, include more of the target shadow and imaging artifacts, but they also include more background clutter which is independent of the target and essentially decrease the signal-to-noise ratio of the image.

The choices of parameter values for a particular algorithm implementation govern the size of the database of parameters required to support the algorithm. We define database complexity to be the logarithm of the number of floating point values in the database per target. For each of the methods presented in this paper, the logarithm of the number of floating point operations required to compute the likelihood function for a target is roughly proportional to this quantity. Each of these methods requires one floating point value per image pixel for each representation window.

3.2. Performance Measures

The orientation of an object in the plane is an element of the special orthogonal group $O \in \text{SO}(2)$ and can be represented in matrix form for any azimuth angle θ as $O(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$. The difference between two orientation matrices is a two by two matrix, $O(\theta_1) - O(\theta_2) \in \mathbb{R}^{2 \times 2}$. For a general matrix A , the squared Hilbert-Schmidt norm of A is defined as the sum of squares of its elements, that is $|A|_{HS}^2 = \sum_{i,j} A_{ij}^2$ and when applied to the difference of two orientation matrices serves as a measure of the difference between the two orientations,

$$d_{HS}^2(O(\theta_1), O(\theta_2)) = \|O(\theta_1) - O(\theta_2)\|_{HS}^2 = 4 - 4 \cos(\theta_1 - \theta_2). \quad (13)$$

This function takes on a minimum value of zero, which occurs whenever the two orientations differ by an integer multiple of 2π , and a maximum value of eight, which occurs when they differ by an odd multiple of π . The average of this squared distance over a given test set, e_{HS}^2 , is used to assess the performance of the orientation estimation algorithm in this paper. As an aid to interpretation of the results, the quantity $\cos^{-1}(1 - e_{HS}^2/4)$ is also noted and is labeled as the “equivalent error in degrees.”

Given the average squared Hilbert-Schmidt error between estimated and actual orientations and the percentage of correctly classified test images over a test set for a wide range of parameterizations, an approximation of the best achievable performance as a function of database complexity can be obtained. This can be accomplished for either performance measure by treating the performance-complexity pairs corresponding to each combination of parameters as samples from the region of performance-complexity pairs that are achievable by the algorithm. The convex outer boundary of this set of points is an approximation of the achievable region boundary and represents the best achievable performance for any specified database complexity.

4. EXPERIMENTAL RESULTS

We have applied the algorithms of Section 2 to ten targets from the publicly released data from the MSTAR program as detailed in Table 1. SAR images of the targets listed in the table that were acquired with a 17° depression angle were used to estimate the unknown model parameters, and images taken from 15° were used to assess algorithm performance. Notable features of the dataset include the facts that the training and testing images were taken under similar conditions with a 2° change in depression angle and that two of the targets, the BMP-2 and T-72, incorporate several different vehicles of differing configurations into a single trained model. Here we present initial results from these experiments, a more complete description is presented by DeVore and O’Sullivan.¹ In the following section, we present target classification results for the conditionally Gaussian approach and the best performing variant of both the log-magnitude and quarter power approaches using an algorithm parameterization that was among the best performing for all three approaches. Following that, in Section 4.2, we present the full performance-complexity results in terms of both orientation estimation and target classification for one variant of each of the approaches to ATR and summarize the results for all variants.

4.1. Confusion Matrices

Tables 2, 3, and 4 show examples of confusion matrices for the conditionally Gaussian, geometric mean normalized log-magnitude, and normalized quarter power approaches to ATR. For all approaches, the image size selected for the tests were chosen to maximize the overall performance and, for the log-magnitude and quarter power approaches, the normalized variants yielded the best overall performance. In all three cases, 45 representation

Target	Train			Test		
	Vehicles	Images	Depression	Vehicles	Images	Depression
2S1	b01	299	17°	b01	274	15°
BMP-2	9563, 9566, c21	697	17°	9563, 9566, c21	587	15°
BRDM-2	E-71	298	17°	E-71	263	15°
BTR-60	k10yt7532	256	17°	k10yt7532	195	15°
BTR-70	c71	233	17°	c71	196	15°
D7	92v13015	299	17°	92v13015	274	15°
T62	A51	299	17°	A51	273	15°
T-72	132, 812, s7	691	17°	132, 812, s7	582	15°
ZIL131	E12	299	17°	E12	274	15°
ZSU 23 4	d08	299	17°	d08	274	15°

Table 1. Subset of the MSTAR publicly released data which was used for the performance assessment of all approaches presented. In all, 3670 images were used for model parameter estimation and 3192 images were used for testing.

	2S1	BMP 2	BRDM2	BTR 60	BTR 70	D7	T62	T72	ZIL131	ZSU 23 4	
2S1	259	1	0	0	0	0	7	7	0	0	94.53%
BMP 2	1	575	0	3	0	0	0	8	0	0	97.96%
BRDM2	4	1	231	4	0	9	4	2	7	1	87.83%
BTR 60	0	0	0	193	0	0	0	0	0	2	98.97%
BTR 70	2	3	0	0	189	0	0	2	0	0	96.43%
D7	2	0	0	0	0	271	1	0	0	0	98.91%
T62	0	0	0	0	0	0	259	10	4	0	94.87%
T72	0	0	0	0	0	0	1	581	0	0	99.83%
ZIL131	0	0	0	0	0	0	2	0	272	0	99.27%
ZSU 23 4	0	0	0	0	0	1	0	0	0	273	99.64%

Table 2. Confusion matrix for the conditionally Gaussian approach with 45 representation windows and with model parameters estimated from a 16° interval. Images of size 80 × 80 were used and the overall percentage of correctly classified images was 97.21%.

	2S1	BMP 2	BRDM2	BTR 60	BTR 70	D7	T62	T72	ZIL131	ZSU 23 4	
2S1	260	2	2	1	0	0	5	4	0	0	94.89%
BMP 2	1	577	0	0	0	0	0	9	0	0	98.30%
BRDM2	5	11	236	5	0	0	0	2	3	1	89.73%
BTR 60	0	6	3	178	1	2	3	0	0	2	91.28%
BTR 70	2	4	0	0	189	0	0	1	0	0	96.43%
D7	0	0	0	0	0	272	0	1	1	0	99.27%
T62	1	0	0	0	0	0	247	22	3	0	90.48%
T72	0	11	0	0	2	0	0	569	0	0	97.78%
ZIL131	0	1	0	0	0	0	0	2	270	1	98.54%
ZSU 23 4	0	0	0	0	0	2	1	2	2	267	97.45%

Table 3. Confusion matrix for the geometric mean normalized log-magnitude approach with 45 representation windows and with model parameters estimated from a 16° interval. Images of size 48 × 48 were used and the overall percentage of correctly classified images was 96.02%.

	2S1	BMP 2	BRDM2	BTR 60	BTR 70	D7	T62	T72	ZIL131	ZSU 23 4	
2S1	265	5	1	0	0	0	1	2	0	0	96.72%
BMP 2	2	575	1	0	0	0	0	5	2	2	97.96%
BRDM2	1	8	248	2	0	0	0	3	0	1	94.30%
BTR 60	0	1	0	193	1	0	0	0	0	0	98.97%
BTR 70	0	2	0	0	194	0	0	0	0	0	98.98%
D7	0	0	0	0	0	272	0	1	1	0	99.27%
T62	0	0	0	0	0	0	256	15	0	2	93.77%
T72	0	5	0	0	1	0	1	574	0	1	98.63%
ZIL131	1	0	0	0	0	0	0	0	272	1	99.27%
ZSU 23 4	0	0	0	0	0	1	0	0	0	273	99.64%

Table 4. Confusion matrix for the normalized quarter-power approach with 45 representation windows and with model parameters estimated from a 16° interval. Images of size 128×128 were used and the overall percentage of correctly classified images was 97.81%.

windows were used and model parameters were trained over an interval of width 16° . This combination was among the top performers for all methods.

From these tables, it can be seen that all three algorithms performed well on the given SAR data. Also common among the three algorithms is that models of the BMP-2 and T-72, though incorporating images of several vehicles of varying configuration, were able to sufficiently capture the essence of the target so as to yield recognition rates of nearly 98% and higher. The targets 2S1, BRDM-2, and T62 were consistently low performers. The most common error was the classification of T62 images as the target T-72.

4.2. Performance and Complexity

The algorithms of Section 2 were applied to the dataset of Table 1 at 40 combinations of number of representation windows and training interval width and at six image sizes to determine the outer boundary of the performance-complexity region for each. Examples of the performance-complexity pairs for all 240 parameterizations are shown in Figures 2, 3, and 4 for the conditionally Gaussian, geometric mean normalized log-magnitude, and normalized quarter-power approaches, respectively. A legend for these plots is shown in Figure 5. Each plot consists of short curves each of which represent one of the 40 combinations of windows and training intervals and each curve has six point markers which correspond to the image sizes from 48×48 through 128×128 . The solid curve which runs along the left and bottom edge of each graph represents the convex outer boundary of the collection of performance-complexity pairs. The top-left point on this curve is determined by the lowest complexity parameterization and the bottom-right point corresponds to the parameterization such that increasing complexity does not yield an increased performance.

On the plots detailing orientation estimation, the approximately horizontal curves representing performance as a function of image size for all three approaches indicate that orientation estimation performance is, in general, less sensitive to image size than to the number of representation windows and training interval width. All three exhibit a clustering of curves at the bottom edge of the graph indicating that several different parameterizations yield nearly identical orientation estimation performance and that for database complexities above 4.5, only minor improvements in performance can be achieved by simply varying the algorithm parameterization. The plots detailing the recognition error rate of the algorithms show that the conditionally Gaussian approach is moderately sensitive to changes in image size with the lowest error rate for most combinations of representation windows and training interval widths coming at the middle image sizes. On the other hand, it can be seen that the geometric mean normalized log-magnitude and normalized quarter-power approaches are quite sensitive to changes in image size. In all three approaches, for database complexities less than around 4.5, the algorithms are more sensitive to changes in the representation window parameters than to image size, but for larger complexities they tend to be more sensitive to changes in image size.

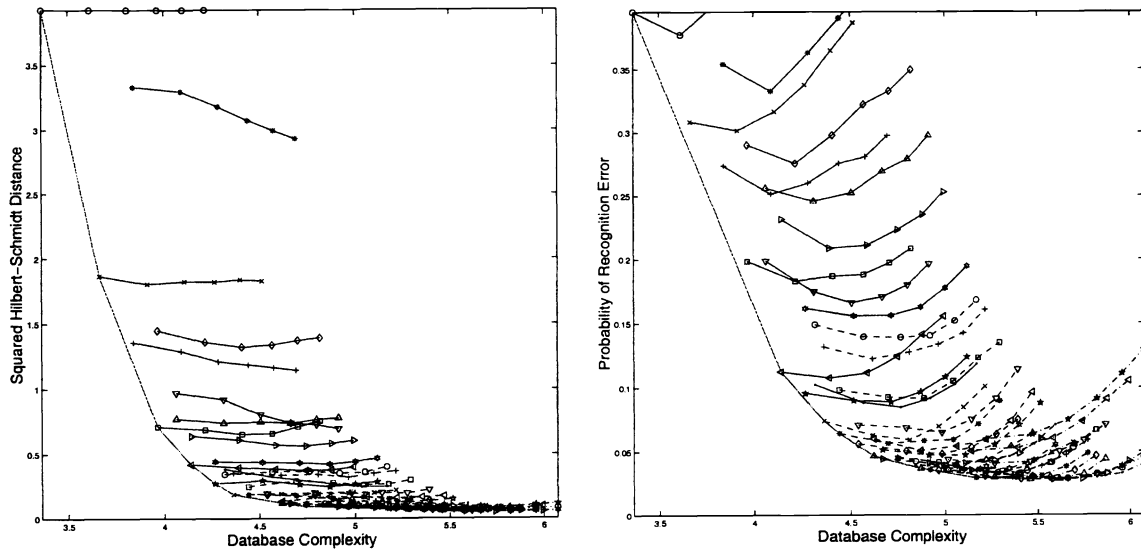


Figure 2. Performance-complexity pairs for the conditionally Gaussian approach. The left pane shows performance in terms of average orientation estimation error and the right pane shows percentage of classification errors. Performance is shown at six image sizes for each combination of number of representation windows and training interval width as indicated in the legend shown in Figure 5. Database complexity is defined as the logarithm of the number of floating point values per target model.

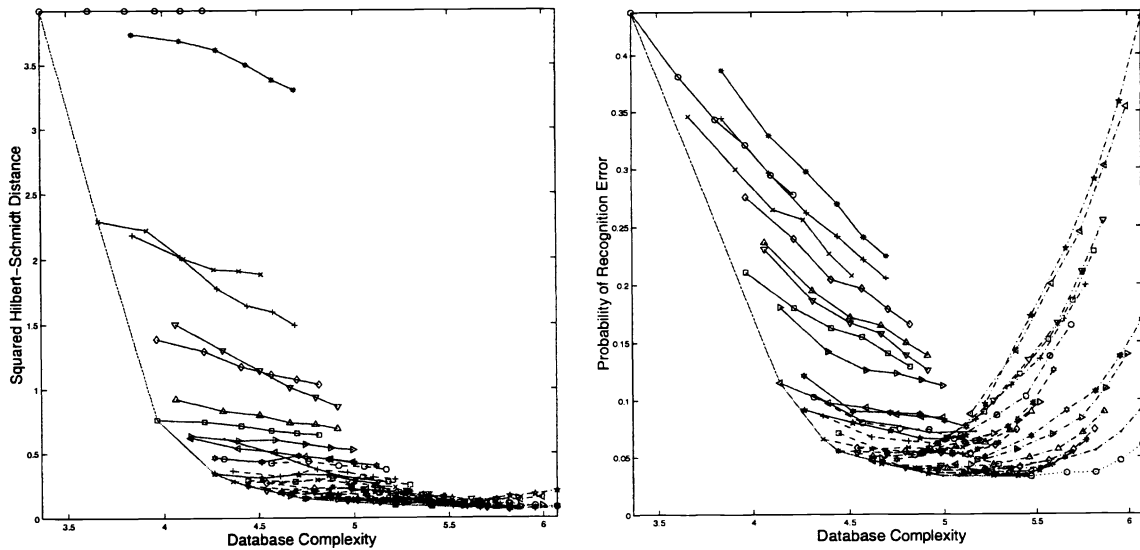


Figure 3. Performance-complexity pairs for the geometric mean normalized log-magnitude approach. The left pane shows performance in terms of average orientation estimation error and the right pane shows percentage of classification errors. Performance is shown at six image sizes for each combination of number of representation windows and training interval width as indicated in the legend shown in Figure 5. Database complexity is defined as the logarithm of the number of floating point values per target model.

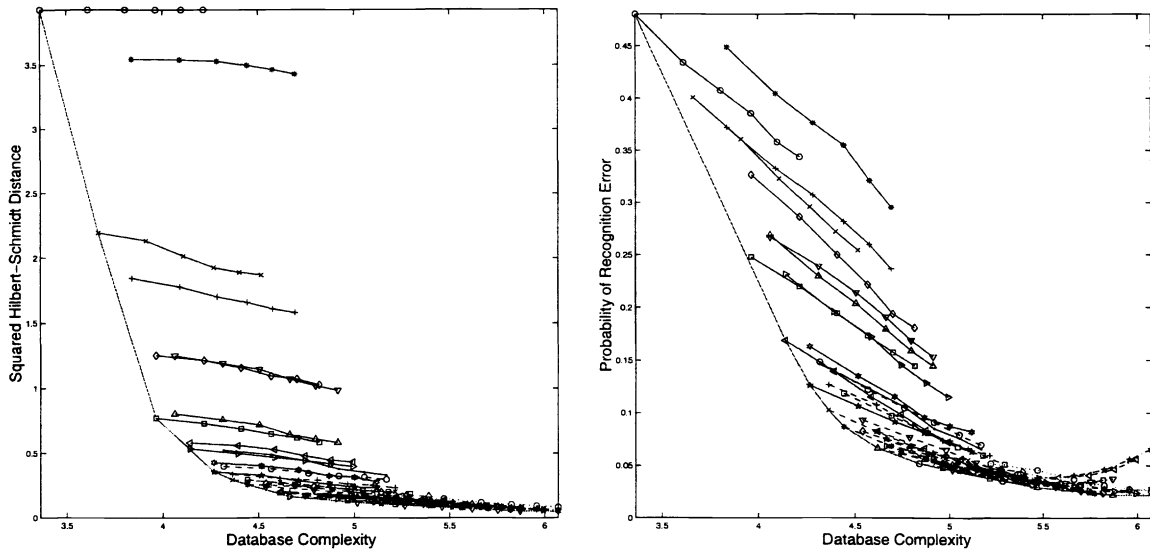


Figure 4. Performance-complexity pairs for the normalized quarter power approach. The left pane shows performance in terms of average orientation estimation error and the right pane shows percentage of classification errors. Performance is shown at six image sizes for each combination of number of representation windows and training interval width as indicated in the legend shown in Figure 5. Database complexity is defined as the logarithm of the number of floating point values per target model.

$\circ-\circ$ $w = 1, d = 360^\circ$	$\triangleleft-\triangleleft$ $w = 6, d = 60^\circ$	$*-*$ $w = 12, d = 30^\circ$	$\star-\star$ $w = 24, d = 15^\circ$	$\nabla-\nabla$ $w = 45, d = 8^\circ$
$\times-\times$ $w = 2, d = 180^\circ$	$\triangleright-\triangleright$ $w = 6, d = 120^\circ$	$\square-\square$ $w = 12, d = 60^\circ$	$\leftarrow-\leftarrow$ $w = 24, d = 30^\circ$	$\triangle-\triangle$ $w = 45, d = 16^\circ$
$+ - +$ $w = 3, d = 120^\circ$	$\star-\star$ $w = 8, d = 45^\circ$	$\diamond-\diamond$ $w = 15, d = 24^\circ$	$\circ-\circ$ $w = 30, d = 12^\circ$	$\triangleleft-\triangleleft$ $w = 60, d = 6^\circ$
$*-*$ $w = 3, d = 240^\circ$	$\star-\star$ $w = 8, d = 90^\circ$	$\nabla-\nabla$ $w = 15, d = 48^\circ$	$\times-\times$ $w = 30, d = 24^\circ$	$\triangleright-\triangleright$ $w = 60, d = 12^\circ$
$\square-\square$ $w = 4, d = 90^\circ$	$\leftarrow-\leftarrow$ $w = 9, d = 40^\circ$	$\triangle-\triangle$ $w = 18, d = 20^\circ$	$+ - +$ $w = 36, d = 10^\circ$	$\star-\star$ $w = 72, d = 5^\circ$
$\diamond-\diamond$ $w = 4, d = 180^\circ$	$\circ-\circ$ $w = 9, d = 80^\circ$	$\triangleleft-\triangleleft$ $w = 18, d = 40^\circ$	$*-*$ $w = 36, d = 20^\circ$	$\star-\star$ $w = 72, d = 10^\circ$
$\nabla-\nabla$ $w = 5, d = 72^\circ$	$\times-\times$ $w = 10, d = 36^\circ$	$\triangleright-\triangleright$ $w = 20, d = 18^\circ$	$\square-\square$ $w = 40, d = 9^\circ$	$\leftarrow-\leftarrow$ $w = 72, d = 15^\circ$
$\triangle-\triangle$ $w = 5, d = 144^\circ$	$+ - +$ $w = 10, d = 72^\circ$	$\star-\star$ $w = 20, d = 36^\circ$	$\diamond-\diamond$ $w = 40, d = 18^\circ$	$\circ-\circ$ $w = 72, d = 20^\circ$

Figure 5. Legend for each of the performance-complexity detail plots. Each pair of point markers and line styles correspond to a combination of number of representation windows, w , and training interval width, d .

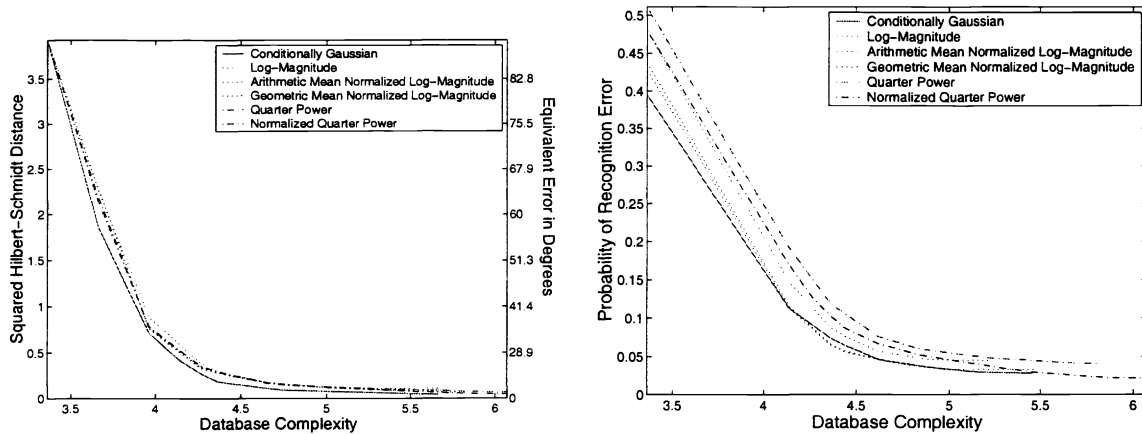


Figure 6. Best achievable performance as a function of complexity for all of the algorithms. The left pane details the average error in orientation estimates and the right pane details the classification error rate.

The outer boundaries from all of the algorithms are shown together in Figure 6. In terms of orientation estimation performance, the conditionally Gaussian approach yielded a slightly lower average error at each complexity than any of the other approaches. The other approaches remain relatively undifferentiated at all complexities. The plots of probability of classification error show that at database complexities up to around 4.2 and over the the range from 5 to 5.5, the conditionally Gaussian approach delivered a lower error rate than the other approaches. For the range from 4.2 to 4.6, the two normalized variants of the log-magnitude approach yielded an error rate that was slightly lower and for the range beyond 5.5, the normalized variant of the quarter power approach had a lower error rate. In both the log-magnitude and quarter power approaches, normalization lowered the error rate significantly. The two forms of normalization in the log-magnitude approach performed nearly identically except at very low complexities where the arithmetic mean normalized variant had a slightly lower error rate.

5. CONCLUSIONS

We have introduced an approach to allow direct comparisons between alternative algorithms and algorithm variants for ATR from SAR imagery. This method allows a quantitative assessment of the performance benefit attained by increasing model complexity and points to algorithm parameterizations which may be preferred for the data at hand. We have demonstrated this method on six algorithms using actual SAR data to form a ten class test. The training and testing data were taken under similar conditions and used in an identical manner in all algorithms. This method can also be used to compare the performance of algorithms when faced with test data that differs from the training data in some significant way so as to assess algorithm robustness and sensitivity to operating conditions. We have demonstrated two measures of performance, alternate measures of complexity can be easily substituted.

For the algorithms presented, it has been shown that the conditionally Gaussian approach yields consistently better target orientation estimation than any of the other methods and that for all of the approaches, the quality of orientation estimates is not highly dependent upon the size of image used. While for the log-magnitude and quarter power approaches image normalization did not noticeably affect orientation estimation performance, it did significantly improve target classification performance. For complexities below around 4.2, the conditionally Gaussian approach delivered a classification performance superior to even the normalized variants of the log-magnitude and quarter power approaches. For the complexity range of 4.2 to 4.6, and above 5.5, the normalized log-magnitude and normalized quarter power, respectively yielded superior performance. More complete numerical results are presented by DeVore and O’Sullivan.¹

ACKNOWLEDGMENTS

This work was supported in part by the US Army Research Office grant DAAH04-95-1-0494, by the Office of Naval Research grant N00014-98-1-06-06, and by the Boeing McDonnell Foundation. A special word of thanks goes to L. Novak of Lincoln Laboratory to M. Bryant of Wright Laboratory who generously shared information concerning the log-magnitude and quarter-power type approaches, respectively.

REFERENCES

1. M. D. DeVore and J. A. O'Sullivan. A performance complexity study of several approaches to automatic target recognition from synthetic aperture radar images. *IEEE Transactions on Aerospace and Electronic Systems*. Submitted for publication.
2. M. I. Miller, U. Grenander, J. A. O'Sullivan, and D. L. Snyder. Automatic target recognition organized via jump-diffusion algorithms. *IEEE Transactions on Image Processing*, 6(1):157–174, Jan. 1997.
3. L. M. Novak, G. J. Owirka, and A. L. Weaver. Automatic target recognition using enhanced resolution SAR data. *IEEE Transactions on Aerospace and Electronic Systems*, 35(1):157–175, Jan. 1999.
4. J. A. O'Sullivan, R. E. Blahut, and D. L. Snyder. Information-theoretic image formation. *IEEE Transactions on Information Theory*, 44(6):2094–2123, Oct. 1998.
5. J. A. O'Sullivan and M. D. DeVore. Performance analysis of ATR from SAR imagery. In *Proceedings of the 33rd Annual Conference on Information Sciences and Systems*, Apr. 1999.
6. J. A. O'Sullivan, M. D. DeVore, V. Kedia, and M. I. Miller. Automatic target recognition performance for SAR imagery using a conditionally Gaussian model. *IEEE Transactions on Aerospace and Electronic Systems*, 1999. Submitted for Publication.
7. G. J. Owirka and L. M. Novak. A new SAR ATR algorithm suite. In *Algorithms for Synthetic Aperture Radar Imagery, Proc. of SPIE*, volume 2230, pages 336–343, 1994.
8. T. D. Ross, J. J. Bradley, L. J. Hudson, and M. P. O'Connor. SAR ATR - so what's the problem? an MSTAR perspective. In E. G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery VI, Proc. of SPIE*, volume 3721, pages 566–573, 1999.
9. S. W. Worrell, S. Parker, and M. L. Bryant. Class separability assessments and MSE algorithm robustness. In E. G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery IV, Proc. of SPIE*, volume 3070, pages 294–304, 1997.