

SAR ATR Performance Using a Conditionally Gaussian Model

JOSEPH A. O'SULLIVAN, Senior Member, IEEE

MICHAEL D. DeVORE, Student Member, IEEE

VIKAS KEDIA

Washington University

MICHAEL I. MILLER

Johns Hopkins University

A family of conditionally Gaussian signal models for synthetic aperture radar (SAR) imagery is presented, extending a related class of models developed for high resolution radar range profiles. This signal model is robust with respect to the variations of the complex-valued radar signals due to the coherent combination of returns from scatterers as those scatterers move through relative distances on the order of a wavelength of the transmitted signal (target speckle). The target type and the relative orientations of the sensor, target, and ground plane parameterize the conditionally Gaussian model. Based upon this model, algorithms to jointly estimate both the target type and pose are developed. Performance results for both target pose estimation and target recognition are presented for publicly released data from the MSTAR program.

Manuscript received June 22, 1999; revised March 3 and October 20, 2000; released for publication October 20, 2000.

IEEE Log No. T-AES/37/1/02919.

Refereeing of this contribution was handled by L. M. Kaplan.

This work was supported in part by the U.S. 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.

Authors' addresses: J. A. O'Sullivan, M. D. DeVore, V. Kedia, Department of Electrical Engineering, Box 1127, Washington University, One Brookings Dr., St. Louis, MO 63130; M. I. Miller, Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218.

0018-9251/01/\$10.00 © 2001 IEEE

I. INTRODUCTION

A general automatic target recognition (ATR) problem includes detecting and identifying targets and estimating their parameters in a given scene using one or more sensors. Assuming that a target is detected and its position is known in the synthetic aperture radar (SAR) context implies that a target chip can be extracted from the SAR image of a scene. SAR images are highly sensitive to target orientation. Reasons for large variability in SAR imagery include shadowing effects, interaction of the signature with the environment (clutter), projection of a three dimensional scene onto a slant plane, and variation of complex-valued radar signals due to the coherent combination of returns from scatterers as they move through relative distances on the order of a wavelength of the transmitted signal (target speckle). A signal model for SAR imagery must incorporate these and other important effects and must be robust with respect to the model assumptions in order to be valid for a variety of imaging scenarios.

Mossing, et al. [11] present an overview of ATR algorithms and theory, and they make specific suggestions for the design of experiments with the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset. Performing ATR using SAR data can be quite challenging [21], and it is important to understand the tradeoffs between performance, cost, and operating conditions of ATR systems. Performance of ATR algorithms should be assessed under a range of operating conditions [11]. Several approaches to ATR from SAR images have been proposed and applied to the MSTAR dataset including a least squared error approach based on pixel decibel values [12, 20], least squared error based on pixel quarter power values [26], scattering center approaches [1, 25], and neural network based approaches [24].

We present a family of conditionally Gaussian signal models that capture the variability in SAR imagery and which extend a related class of models for high range resolution radar data which is motivated by the physics of radar imagery [5]. From this signal model, algorithms to jointly estimate the type and pose of a target in a SAR image are derived. These algorithms are then applied to publicly released SAR data from the MSTAR program. Performance of the Hilbert-Schmidt (minimum mean-squared error) estimator for target orientation and of the ATR algorithm are presented in a four class scenario under standard operating conditions and a range of extended operating conditions including target positional uncertainty, configuration variations, and class version variations, using guidelines described [11, 22]. The tradeoffs between performance and cost in terms of database complexity are discussed and experiments

which demonstrate them are presented in a manner similar to that in [21]. Finally, preliminary results under standard operating conditions are presented for a ten class scenario.

Section II introduces a family of conditionally Gaussian models for SAR data. Section III details orientation estimation through the Hilbert–Schmidt estimator. Section IV details target recognition. Section V contains the performance results using MSTAR data on a four class problem. Section VI extends those results to a ten class problem. Conclusions follow in Section VII.

II. MODELING SYNTHETIC APERTURE RADAR DATA

A. Deterministic and Stochastic Models

We have studied in detail the use of both deterministic and conditionally Gaussian models for high range resolution (HRR) radar data. These results, discussed in [17, 18], show that the deterministic model works very well when all possible HRR range profiles are described by the database. That is, in the situation where the database is complete, the deterministic model is ideal. However, the deterministic model is not very robust. For example, as shown in [18], if a deterministic model is trained with a set of range profiles which does not intersect the set of range profiles used for testing, then significant performance degradation results. The conditionally Gaussian model, on the other hand, is robust in the sense that range profiles not used in the training data may still be modeled well. For this reason, we only consider the conditionally Gaussian model for extension to the SAR scenario.

B. Conditionally Gaussian Model

A SAR image is obtained in the form of an $m \times n$ array of pixels which can be treated as a column vector of length mn . We consider a stochastic signal model under which, given the target type and orientation, the signal part of the SAR image $\mathbf{s}(\Theta, a)$ forms a complex Gaussian random process. Using Θ to denote target pose, $\Theta \in \mathbf{SE}(2)$ for ground-based targets, and a to denote target type, the data are assumed to be

$$\mathbf{r} = \mathbf{s}(\Theta, a) + \mathbf{w} \quad (1)$$

where, given $\{\Theta, a\}$, the signal \mathbf{s} is a complex Gaussian random vector with conditional mean $\mu(\Theta, a)$ and covariance matrix $\mathbf{K}(\Theta, a)$, and \mathbf{w} is complex Gaussian noise with mean $\mathbf{0}$ and covariance $N_0 \mathbf{I}$. Thus,

$$E\{\mathbf{r} \mid \Theta, a\} = \mu(\Theta, a) \quad (2)$$

$$E\{[\mathbf{r} - \mu][\mathbf{r} - \mu]^\dagger \mid \Theta, a\} = \mathbf{K}(\Theta, a) + N_0 \mathbf{I} \quad (3)$$

where \dagger denotes complex conjugate transpose. Under this model, the log-likelihood of \mathbf{r} given $\{\Theta, a\}$ is

$$l(\mathbf{r} \mid \Theta, a) = -\ln|\mathbf{K}(\Theta, a) + N_0 \mathbf{I}| - [\mathbf{r} - \mu(\Theta, a)]^\dagger \times (\mathbf{K}(\Theta, a) + N_0 \mathbf{I})^{-1} [\mathbf{r} - \mu(\Theta, a)]. \quad (4)$$

We assume that the noise \mathbf{w} is independent from pixel to pixel and that it is independent of the signal \mathbf{s} . Given the target type, orientation, and variances of the pixels in an image, the data are independent from pixel to pixel. Further, our results from HRR data [8] and initial experiments with SAR data indicate that the data are well modeled as having zero mean, $\mu(\Theta, a) = \mathbf{0}$. These result in a diagonal covariance matrix \mathbf{K} and a log-likelihood function $l(\mathbf{r} \mid \Theta, a)$ which involves no matrix computations but is simply a summation over all pixels,

$$l(\mathbf{r} \mid \Theta, a) = \sum_i \left[-\ln(K_{i,i}(\Theta, a) + N_0) - \frac{|r_i|^2}{K_{i,i}(\Theta, a) + N_0} \right]. \quad (5)$$

This is a direct extension of the conditionally Gaussian model that has been used successfully in modeling high resolution radar range profiles [5–7].

C. MSTAR Dataset

The SAR image data used for the results we present were collected as part of the MSTAR¹ program which was supported under DARPA funding. The MSTAR SAR data were collected using the Sandia National Laboratories Twin Otter SAR sensor payload operating at X-band.

The image set contains data for a number of targets imaged at two or more depression angles. For each target and depression angle, the set contains SAR image data for two to three hundred values of target aspect pose. We have divided these images into two nonoverlapping sets, one of which is used for model training and the other of which is used for assessing the performance of orientation estimation and target recognition algorithms.

The SAR images in the dataset have 1 ft by 1 ft resolution. The target, along with its shadow, lies in the central portion of the chip and occupies a small portion of it. The rest of the chip consists of background clutter which is independent of the target, hence reducing the effective signal-to-noise (SNR). The results we present include an assessment of how the performance of orientation estimation and target recognition algorithms vary as a function of the target chip size. Smaller image chips were constructed by extracting a square region from the center of the MSTAR images.

¹The MSTAR dataset can be requested through the Sensor Data Management System (SDMS) web page of Wright Laboratory at the URL <http://www.mbvlab.wpafb.af.mil/public/sdms/>.

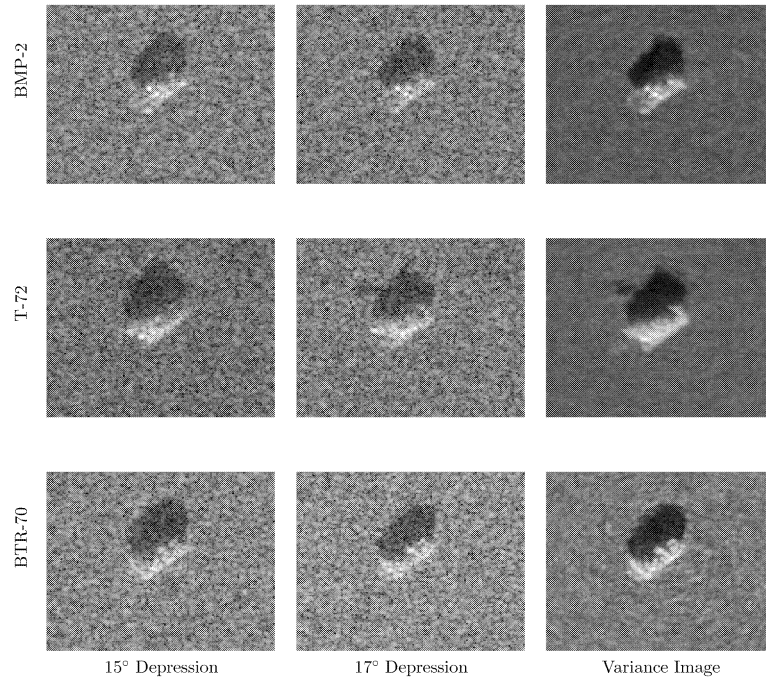


Fig. 1. Sample images for three targets in MSTAR dataset. For each target, left-most image is from sensor with 15° depression angle, middle image is from sensor with 17° depression angle, and right-most image is variance image estimated from SAR images lying in a 10° interval in azimuth.

D. Model Training: Maximum Likelihood Estimation

From the expression for the log-likelihood function (5), the only unknown values are the variances of each pixel i , $\sigma_i^2(\Theta, a) = K_{i,i}(\Theta, a) + N_0$. Model training, then, is a matter of estimating these variances. The set of all estimated values constitutes a database which can be used when performing orientation estimation and target recognition on a SAR image.

We divide the database into N_w windows each trained over an interval in azimuth of width d radians. We assume that the training data has been accurately registered so that azimuth θ is the only pose parameter we need consider. Hence, the k th window is centered at $\theta_k = 2\pi k/N_w$ and consists of variances estimated from SAR images in the testing set in which the target azimuth lies in the interval $W_k = [2\pi k/N_w - d/2, 2\pi k/N_w + d/2]$. The variance function of the i th pixel is approximated as piecewise constant in θ and its value over the k th window is estimated as

$$\hat{\sigma}_i^2(\theta_k, a_l) = \frac{1}{N_k} \sum_{\theta \in W_k} |r_i(\theta, a_l)|^2, \quad 1 \leq k \leq N_w, \quad 1 \leq l \leq t \quad (6)$$

where N_k is the number of training images with azimuth in the interval W_k and t is the number of targets in the database. This estimate is unbiased for the case where the mean is known to be zero. Sample SAR images from two different depression angles and from the trained variance database are shown for three targets in Fig. 1. For these variance images, the database has 72 windows each trained over an interval of width 10°.

Fig. 2 shows a concatenation of all 72 variance images which were trained for the T-72 target. The top left image represents the variance in the pixel values of training images which have an azimuth angle in the interval 0° to 10°. These variances are taken as the value of the piecewise constant variance function over the first window, which spans the 5° interval from 2.5° to 7.5°. Each successive image row-wise, represents a target rotation of 5° clockwise with respect to the radar platform. The second image, then, represents pixel variances of training images which have an azimuth angle in the interval 5° to 15° and these are taken as the value of the variance function over the second window, which spans the interval from 7.5° to 12.5°, and so on.

III. POSE ESTIMATION

A. Hilbert–Schmidt Estimator

In a typical ATR scenario, a target is placed in a 3-dimensional space and, at any fixed time, has an orientation in $\mathbf{SO}(3)$, the special orthogonal group of dimension 3. In this paper we study ground based targets assumed to be on a flat surface. Such targets have their orientations in $\mathbf{SO}(2)$. An element in the group $\mathbf{SO}(n)$ can be represented by an orthogonal $n \times n$ matrix with determinant one. For matrices $A, B \in \mathbf{SO}(n)$, the squared Hilbert–Schmidt distance between them is

$$d_{\text{HS}}^2(A, B) = \|A - B\|_{\text{HS}}^2 = \text{tr}[(A - B)^T (A - B)] \quad (7)$$

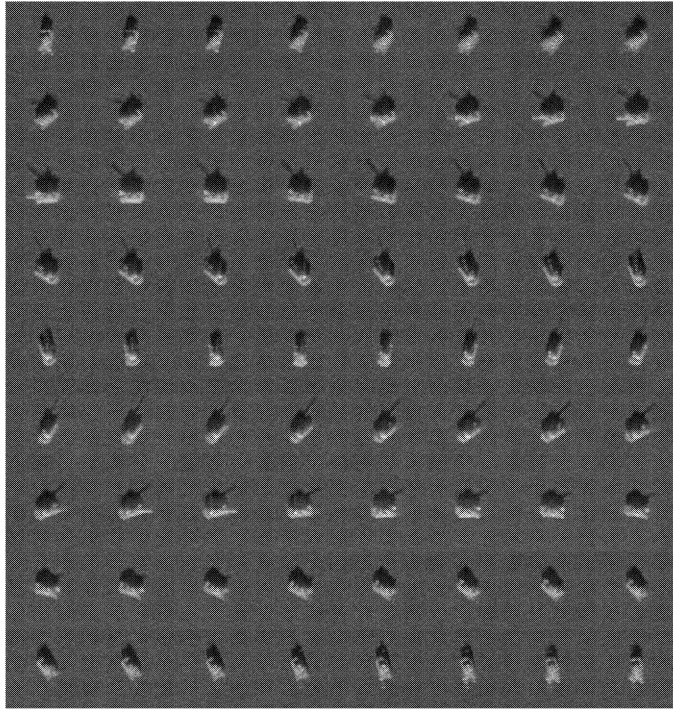


Fig. 2. All 72 variance images trained for T-72 target. Top left image represents variance in SAR image of T-72 tank at approximately 5° azimuth. Each successive image represents image variance when target has an additional rotation of 5° clockwise with respect to radar platform. Final image, on bottom right, is image variance of target with 0° azimuth angle, pointing directly at radar platform.

where $\text{tr}[\cdot]$ is the matrix trace operation and $(\cdot)^T$ denotes transpose. Since $\text{tr}[A^T A] = n$ for all $A \in \mathbf{SO}(n)$,

$$d_{\text{HS}}^2(A, B) = 2n - 2\text{tr}[A^T B] = 2n - 2\text{tr}[B^T A]. \quad (8)$$

For ground-based targets, a 2×2 rotation matrix $\mathbf{O}(\theta) \in \mathbf{SO}(2)$ defines the orientation of a target relative to a high resolution radar, where θ is the azimuth of the target relative to the radar. Then,

$$\mathbf{O}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (9)$$

and for any two azimuth angles, θ_1 and θ_2 ,

$$d_{\text{HS}}^2(\mathbf{O}(\theta_1), \mathbf{O}(\theta_2)) = 4 - 4\cos(\theta_1 - \theta_2). \quad (10)$$

Given a true orientation for the target \mathbf{O}_{true} and any estimate $\hat{\mathbf{O}}$, the squared Hilbert–Schmidt distance $d_{\text{HS}}^2(\mathbf{O}_{\text{true}}, \hat{\mathbf{O}})$ gives a measure of the squared error in the estimate. Let \mathbf{r} denote an observed SAR image of a target of type a from which an estimate $\hat{\mathbf{O}}(\mathbf{r})$ is to be computed and let $p(\mathbf{O} | \mathbf{r}, a)$ denote the probability density function for \mathbf{O} given \mathbf{r} and the target type a . Then the Hilbert–Schmidt estimator for orientation is defined as an estimator that, given \mathbf{r} , minimizes the conditional mean of the squared estimation error as defined by the squared Hilbert–Schmidt distance [4, 23]. Since $p(\mathbf{O} | \mathbf{r}, a) \propto p(\mathbf{r} | \mathbf{O}, a)p(\mathbf{O} | a)$, the

Hilbert–Schmidt estimator can be written as

$$\hat{\mathbf{O}}_{\text{HS}}(\mathbf{r}) = \arg \min_{\mathbf{O} \in \mathbf{SO}(2)} \int_{\mathbf{SO}(2)} \|\mathbf{O}' - \mathbf{O}\|_{\text{HS}}^2 \times p(\mathbf{r} | \mathbf{O}', a) p(\mathbf{O}' | a) \gamma(d\mathbf{O}') \quad (11)$$

where γ is the base measure on $\mathbf{SO}(2)$.

The minimum mean-squared error is lower bounded by the Hilbert–Schmidt bound which is the mean squared error achieved by $\hat{\mathbf{O}}_{\text{HS}}$,

$$\text{HSB} = E_{\mathbf{O}} \{ E_{\mathbf{r}|\mathbf{O}} \{ \|\hat{\mathbf{O}}_{\text{HS}}(\mathbf{r}) - \mathbf{O}\|_{\text{HS}}^2 | \mathbf{O} \} \}. \quad (12)$$

The orientation estimation results presented are in terms of the average squared Hilbert–Schmidt estimation error over all images in the test set. As an aid to interpreting the results, this average squared Hilbert–Schmidt error e_{HS}^2 is also reported as $\cos^{-1}(1 - e_{\text{HS}}^2/4)$ under the notation “equivalent error in degrees.”

B. Pose Estimation Algorithms

The integral in (11) can be written in terms of the target azimuth θ and, after simplifying, the azimuth angle corresponding to the Hilbert–Schmidt estimator for orientation can be expressed as

$$\hat{\theta}_{\text{HS}}(\mathbf{r}) = \tan^{-1} \left[\frac{E[\sin(\theta) | \mathbf{r}, a]}{E[\cos(\theta) | \mathbf{r}, a]} \right] \quad (13)$$

$$= \tan^{-1} \left[\frac{\int \sin(\theta) p(\theta | \mathbf{r}, a) d\theta}{\int \cos(\theta) p(\theta | \mathbf{r}, a) d\theta} \right]. \quad (14)$$

Since

$$p(\theta | \mathbf{r}, a) \propto \int_{\mathbb{R}^2} p(\mathbf{r} | \theta, \mathbf{s}, a) p(\theta | a) p(\mathbf{s} | a) d\mathbf{s} \quad (15)$$

where \mathbf{s} represents object position and \mathbf{s} and θ are assumed to be independent, (14) becomes

$$\hat{\theta}_{\text{HS}}(\mathbf{r}) = \tan^{-1} \left[\frac{\int \sin(\theta) p(\theta | a) \int_{\mathbb{R}^2} p(\mathbf{r} | \theta, \mathbf{s}, a) p(\mathbf{s} | a) d\mathbf{s} d\theta}{\int \cos(\theta) p(\theta | a) \int_{\mathbb{R}^2} p(\mathbf{r} | \theta, \mathbf{s}, a) p(\mathbf{s} | a) d\mathbf{s} d\theta} \right]. \quad (16)$$

As discussed in Section IID, we approximate the covariance function $\sigma^2(\theta, a)$ as piecewise constant in θ . Hence, the probability density function $p(\mathbf{r} | \theta, \mathbf{s}, a)$ is approximated as piecewise constant in θ and \mathbf{s} so the above integrals are reduced to summations over all azimuth windows W_k and all possible object positions \mathbf{s}_j . Assuming a uniform prior on both orientation and position, and evenly spaced windows, the estimate for orientation becomes

$$\hat{\theta}_{\text{HS}}(\mathbf{r}) = \tan^{-1} \left[\frac{\sum_k \sin(\theta_k) \sum_j p(\mathbf{r} | \theta_k, \mathbf{s}_j, a)}{\sum_k \cos(\theta_k) \sum_j p(\mathbf{r} | \theta_k, \mathbf{s}_j, a)} \right]. \quad (17)$$

The minimum mean-squared error estimate of target position, $\hat{\mathbf{s}}_{\text{MMSE}}$, is defined as

$$\hat{\mathbf{s}}_{\text{MMSE}} = \arg \min_{\mathbf{s}} E[\|\mathbf{s} - \mathbf{s}'\|^2 | \mathbf{r}, a] \quad (18)$$

$$= \arg \min_{\mathbf{s}} \int \|\mathbf{s} - \mathbf{s}'\|^2 p(\mathbf{s}' | \mathbf{r}, a) d\mathbf{s}'. \quad (19)$$

Differentiating the above integral with respect to \mathbf{s} , setting the result equal to zero, and noting that

$$p(\mathbf{s} | \mathbf{r}, a) \propto \int p(\mathbf{r} | \theta, \mathbf{s}, a) p(\theta | a) p(\mathbf{s} | a) d\theta \quad (20)$$

yields the result

$$\hat{\mathbf{s}}_{\text{MMSE}}(\mathbf{r}) = \frac{\int \mathbf{s} \int p(\mathbf{r} | \theta, \mathbf{s}, a) p(\theta | a) p(\mathbf{s} | a) d\theta d\mathbf{s}}{\int \int p(\mathbf{r} | \theta, \mathbf{s}, a) p(\theta | a) p(\mathbf{s} | a) d\theta d\mathbf{s}}. \quad (21)$$

Again assuming that $p(\mathbf{r} | \theta, \mathbf{s}, a)$ is piecewise constant in θ and \mathbf{s} , a uniform prior on both orientation and position, and evenly spaced windows, the estimate for position becomes

$$\hat{\mathbf{s}}_{\text{MMSE}}(\mathbf{r}) = \frac{\sum_j \mathbf{s}_j \sum_k p(\mathbf{r} | \theta_k, \mathbf{s}_j, a)}{\sum_j \sum_k p(\mathbf{r} | \theta_k, \mathbf{s}_j, a)}. \quad (22)$$

An alternative method of estimating the target pose is to use maximum a posteriori (MAP) estimators. The MAP estimator for θ and \mathbf{s} is defined as the pose which maximizes the posterior density $p(\theta, \mathbf{s} | \mathbf{r}, a)$. Noting that $p(\theta, \mathbf{s} | \mathbf{r}, a) \propto p(\mathbf{r} | \theta, \mathbf{s}, a) p(\theta | a) p(\mathbf{s} | a)$, assuming a uniform prior on orientation, and allowing for piecewise constant variance functions, the MAP

estimate can be written as

$$\begin{bmatrix} \hat{\theta}_{\text{MAP}}(\mathbf{r}) \\ \hat{\mathbf{s}}_{\text{MAP}}(\mathbf{r}) \end{bmatrix} = \arg \max_{\theta_k, \mathbf{s}_j} p(\mathbf{r} | \theta_k, \mathbf{s}_j, a). \quad (23)$$

Performance of the Hilbert–Schmidt and MAP estimates for orientation were nearly identical in all experiments we conducted and so only results for the Hilbert–Schmidt estimator are presented in this paper. Performance of the MMSE and MAP estimates for position were nearly identical, as well.

In practice, the target type a may not be known. In this case, the pose estimate is a function of the target type. The pose of the object can then be estimated based upon the most likely target $\hat{a}(\mathbf{r})$ given the observed SAR image \mathbf{r} as discussed in Section IV.

IV. Target Recognition

A. Recognition Algorithms

ATR, given a SAR image chip \mathbf{r} , can be performed in a Bayesian approach by selecting the target class $\hat{a}(\mathbf{r})$ which maximizes the probability $P(a | \mathbf{r}) \propto p(\mathbf{r} | a) P(a)$. Since we have estimated the parameters of the density $p(\mathbf{r} | \theta, \mathbf{s}, a)$ and the orientation of the target is unknown, we can treat θ and \mathbf{s} as nuisance parameters, integrating over all orientations and positions, selecting a target class according to

$$\hat{a}_{\text{Bayes}}(\mathbf{r}) = \arg \max_a P(a) \int \int p(\mathbf{r} | \theta, \mathbf{s}, a) p(\theta | a) p(\mathbf{s} | a) d\theta d\mathbf{s}. \quad (24)$$

Because we have approximated the parameters of $p(\mathbf{r} | \theta, \mathbf{s}, a)$ as piecewise constant functions in θ and \mathbf{s} , we reduce the above integrals to finite summations. Assuming a uniform prior on target type, orientation, and pose, and assuming uniformly spaced azimuth windows, we select a target class according to

$$\hat{a}_{\text{Bayes}}(\mathbf{r}) = \arg \max_a \sum_k \sum_j p(\mathbf{r} | \theta_k, \mathbf{s}_j, a). \quad (25)$$

An additive threshold can be incorporated to account for relative costs of classification errors and to bias results toward or away from individual targets yielding

$$\hat{a}_{\text{Bayes}}(\mathbf{r}; \tau) = \arg \max_a \log \left[\sum_k \sum_j p(\mathbf{r} | \theta_k, \mathbf{s}_j, a) \right] + \tau_a. \quad (26)$$

An alternative method of selecting a target type corresponding to an observed image stems from the generalized likelihood ratio test (GLRT) in which that target is selected which has the largest maximum likelihood over all poses. That is,

$$\hat{a}_{\text{GLRT}}(\mathbf{r}) = \arg \max_a \max_{k,j} p(\mathbf{r} | \theta_k, \mathbf{s}_j, a). \quad (27)$$

Performance results for this recognition algorithm were nearly identical to the Bayesian approach in all experiments we conducted and so only results for the Bayesian classifier are presented in this paper.

B. Recognition Performance Measures

Performance of the ATR algorithm is characterized in this paper in terms of experimental receiver operating characteristic (ROC) curves and confusion matrices. This performance is measured as the fraction of test images from the MSTAR dataset which are recognized correctly. Performance is measured as a function of target class and SAR image size.

The experimental ROC curves demonstrate the recognition performance tradeoffs in a binary hypothesis test between a pair of targets a_1 and a_2 . Each point on such a curve indicates the fraction of images of type a_1 which are correctly recognized versus the fraction of images of type a_2 which are incorrectly classified as type a_1 . The curve is parameterized by an additive threshold τ and recognition is performed according to

$$\hat{a}(\mathbf{r}) = \begin{cases} a_1 & \text{if } \log \left[\sum_i \sum_j p(\mathbf{r} | \theta_i, \mathbf{s}_j, a_1) \right] + \tau > \log \left[\sum_i \sum_j p(\mathbf{r} | \theta_i, \mathbf{s}_j, a_2) \right] \\ a_2 & \text{otherwise} \end{cases} \quad (28)$$

The confusion matrices demonstrate the recognition performance when all target classes are allowable choices. Denote by M_l the number of images of target type a_l in the set of test images and denote by $n_{l,k}$ the number of images of target type a_l which are classified by the algorithm as type a_k . The element of the confusion matrix at row l and column k is the quantity $n_{l,k}$ and the overall probability of correct classification for target type a_l is estimated as $n_{l,l}/M_l$. Let $M = \sum_l M_l$ denote the total number of test images, then the overall probability of correct classification is estimated as $\sum_l n_{l,l}/M$. Confusion matrices and the corresponding overall probabilities of recognition are functions of the threshold vector τ from (26).

V. PERFORMANCE IN THE FOUR CLASS PROBLEM

In this section, we look at the orientation estimation and target recognition performance computed in a series of tests in which each test image was classified as one of four targets for which models had been trained. The tests were chosen to support performance analysis along several different dimensions encompassing a variety of operating conditions and model complexity [11].

Each test involves the variation of a single operational parameter from its nominal value while holding all other parameters fixed. The baseline test was performed under standard operating conditions (SOC) in which test images are independent of the training data but were formed under similar conditions. The SOC tests were performed on six different image sizes from 48×48 to 128×128 with models approximated as piecewise constant over 72 orientation windows and with each window trained from images coming from a 10° interval in azimuth. Target models were trained using 17° depression angle images and, for two of the targets, images from multiple vehicles of different serial numbers were used to train the models. Testing was performed using 15° depression angle images of the same vehicles used in model training. The test images were assumed to be accurately registered with the training data and the estimation and classification algorithms considered only azimuth uncertainty. To analyze performance as a function of database complexity under SOC, the number of piecewise constant orientation windows and the width

of the azimuth intervals from which training data was drawn for each window was varied across a wide range of values.

The remaining tests were conducted under extended operating conditions (EOC), in which a single operational parameter was significantly different between the images used for training and those used for testing. Tests under the registration uncertainty EOC employ the same training and testing data as the SOC tests, however each test image was randomly shifted by 0, 1, or 2 pixel widths in the horizontal and vertical directions to simulate the case where image registration is not performed with a high degree of accuracy. Only images of size 80×80 were used and the estimation and classification routines were extended to include positional uncertainty. Tests under the configuration variation EOC employ test images of vehicles which differ in serial number and have minor configuration changes from those used for model training. There were two such tests, one performed with target models trained from a very rich set of images and another performed with models trained from a more limited set of images. Finally, tests under the version variation EOC employ test images of vehicles which differ in serial number and have significant version changes from those used for

TABLE I
MSTAR Images Comprising Training Set Number 1

Model	Training Set #1			
	Vehicle	Serial No.	Depression	Images
BMP-2	#1	9563	17°	233
	#2	9566		231
	#3	c21		233
BTR-70	#1	c71	17°	233
BRDM-2	#1	E-71	17°	298
T-72	#1	132	17°	232
	#2	812		231
	#3	s7		228

Note: The four models trained from this data were used in analysis of performance under SOC, performance under the registration uncertainty EOC, and performance under varying database complexities.

model training. Again, there were two such tests, one performed with target models trained from a very rich set of images and another performed with models trained from a more limited set of images.

Table I shows training set number 1 which consists of the MSTAR data used to train the four models when conducting performance analyses for the SOC, for analyzing performance under variations of database complexity, and for analyzing performance under the registration uncertainty EOC. For the BMP-2 and T-72 models, images from three different vehicles were combined to train each of the models. Table II shows training sets 2 and 3 which consist of the MSTAR data used to train the four models when conducting performance analyses for both the configuration and version variation EOCs. These training sets were selected so that no vehicles have images which are used for both training and testing.

Table III shows the subset of the MSTAR data which was used for testing in the SOC, database complexity, and registration uncertainty EOC cases. Note that for two of the targets, the BMP-2 and the T-72, images from multiple vehicles are collected under a single heading. Table IV shows the subsets of the MSTAR data which were used for testing in the configuration variation and version variation EOC cases. In these tests, different vehicles of the same target class are reported on individually and are not aggregated as in the SOC testing set.

TABLE II
MSTAR Images Comprising Training Sets 2 and 3

Model	Training Set #2				Training Set #3			
	Vehicle	Serial No.	Depression	Images	Vehicle	Serial No.	Depression	Images
BMP-2	#3	c21	15° and 17°	429	#3	c21	17°	233
BTR-70	#1	c71	15° and 17°	429	#1	c71	17°	233
BRDM-2	#1	E-71	15° and 17°	561	#1	E-71	17°	298
T-72	#1	132	15° and 17°	428	#1	132	17°	232

Note: The four models trained from this data were used in analysis of performance under both the configuration variation and version variation EOC.

TABLE III
MSTAR Images Comprising SOC Testing Set

Target	SOC Testing Set			
	Vehicle	Serial No.	Depression	Images
BMP-2	#1	9563	15°	195
	#2	9566		196
	#3	c21		196
BTR-70	#1	c71	15°	196
BRDM-2	#1	E-71	15°	263
T-72	#1	132	15°	196
	#2	812		195
	#3	s7		191

Note: Testing done on these images was used in analysis of performance under SOC, performance under the registration uncertainty EOC, and performance under varying database complexities.

A. Pose Estimation Performance

Fig. 3 shows plots of the average orientation estimation performance as a function of image size under SOC, the configuration variation EOC, and the version variation EOC. These plots show that on average, good estimates of the orientation can be achieved under SOC, with the best results occurring at the middle image sizes 80×80 and 96×96 . The plots under EOC also show that performance suffers dramatically when estimating the orientation of vehicles which are significantly different from the training data, with the best results occurring at small image sizes 48×48 and 64×64 . Comparing the configuration variation results when training data was selected from both 15° and 17° depression angles with the results when only 15° training data is used, it can be seen that the increased quantity of training data did not noticeably improve orientation estimation performance at small image sizes. At large image sizes, however, the additional training data did improve the performance. Comparing the results of version variant tests given large and small training sets, however, shows that a noticeable increase is obtained at all image sizes when the larger training sets are used. The presence of registration uncertainty in the test images did not degrade the orientation estimation performance. In fact the extended

TABLE IV
MSTAR Images Comprising Configuration Variation and Version Variation EOC Testing Sets

Configuration Variant Testing Set				Version Variant Testing Set			
Target	Serial No.	Depression	Images	Target	Serial No.	Depression	Images
T-72 #2	812	15° and 17°	426	T-72 #3	s7	15° and 17°	419
T-72 #4	A04	15° and 17°	573	T-72 #8	A32	15° and 17°	572
T-72 #5	A05	15° and 17°	573	T-72 #9	A62	15° and 17°	573
T-72 #6	A07	15° and 17°	573	T-72 #10	A63	15° and 17°	573
T-72 #7	A10	15° and 17°	567	T-72 #11	A64	15° and 17°	573
BMP-2 #1	9563	15° and 17°	428				
BMP-2 #2	9566	15° and 17°	427				

Note: Testing done on these images was used in analysis of performance under both the configuration variation and version variation EOC.

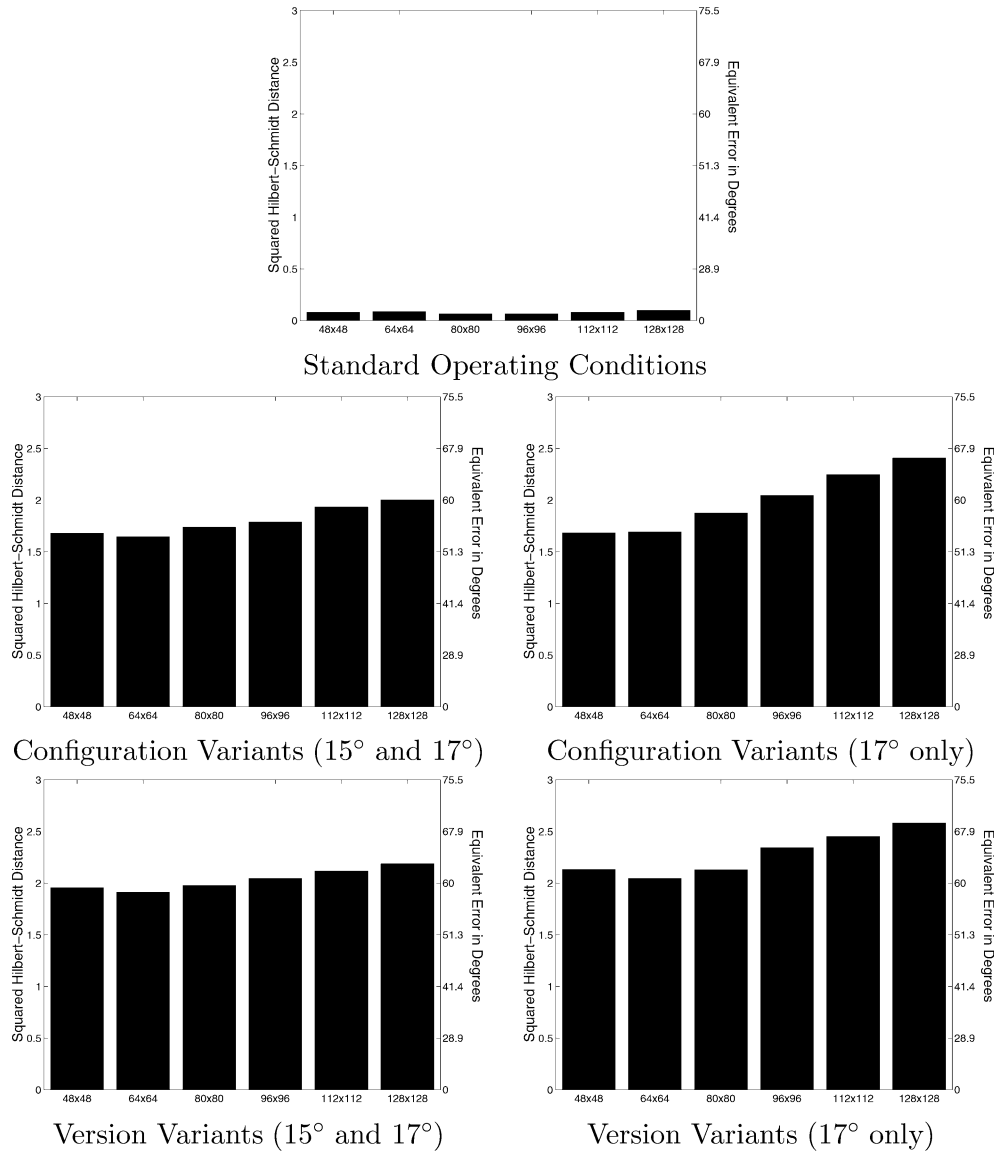


Fig. 3. Average over all test images of Hilbert-Schmidt distance between actual and estimated orientations as function of target chip size. From top to bottom right the graphs represent estimation error under SOC, configuration variants when both 15° and 17° depression data was used for training, configuration variants when only 17° depression data was used for training, version variants when both 15° and 17° depression data were used for training, and version variants when only 17° depression data was used for training, respectively.

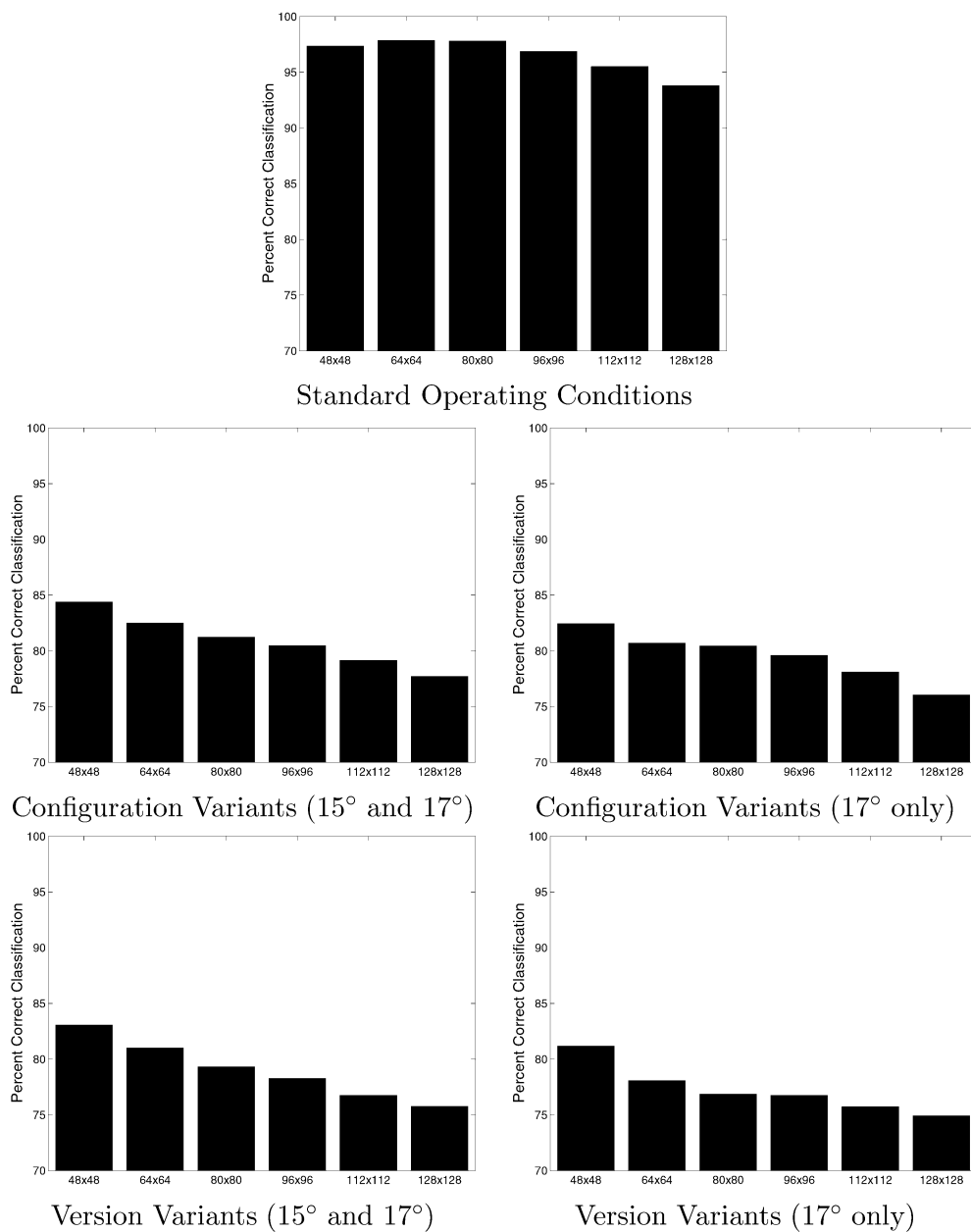


Fig. 4. Average probability of correct recognition as function of target chip size. From top to bottom right the graphs represent the estimation error under SOC, configuration variants when both 15° and 17° depression data were used for training, configuration variants when only 17° depression data was used for training, version variants when both 15° and 17° depression data were used for training, and version variants when only 17° depression data was used for training, respectively.

algorithm, which accommodated unknown target position, yielded an average squared Hilbert–Schmidt error of 0.04, which is an equivalent error of 8.06° , whereas the tests under SOC yielded an average squared Hilbert–Schmidt error of 0.07, suggesting that the images used for the SOC tests may not themselves have been well registered.

B. Four Class Target Recognition

Fig. 4 shows the average probability of correct recognition as a function of target chip size under SOC, configuration variation EOC, and version

variation EOC. These graphs show that under SOC, classification rates of nearly 98% were obtained at the 64×64 and 80×80 image sizes. For tests under both configuration and version variation EOC, the best classification rates were achieved at the smallest image size, 48×48 . As expected, performance under EOC was slightly higher when more training data was available. For all tests, performance dropped by several percentage points when large images, which include significant background clutter, were used.

Tables V–X detail the recognition performance through confusion matrices for 80×80 images under SOC, registration uncertainty EOC, configuration

TABLE V
Recognition Test as Four Class Problem Under SOC for 80×80 Images

	BMP-2	BRDM-2	BTR-70	T-72	Percent
BMP-2	581	0	0	6	98.98%
BRDM-2	6	243	0	14	92.40%
BTR-70	7	0	186	3	94.90%
T-72	0	0	0	582	100%

Note: The row headings indicate the true target class for all images represented in each row. Each row contains the number of images from the test set that were chosen as belonging to the target class indicated in the corresponding column heading. The last column lists the correct recognition rate for each row. The overall percentage of correctly classified test images was 97.79%.

TABLE VI
Recognition Test as Four Class Problem Under Registration Uncertainty EOC for 80×80 Images

	BMP-2	BRDM-2	BTR-70	T-72	Percent
BMP-2	580	0	0	7	98.81%
BRDM-2	4	252	0	7	95.82%
BTR-70	12	0	183	1	93.37%
T-72	0	0	0	582	100%

Note: The row headings indicate the true target class for all images represented in each row. Each row contains the number of images from the test set that were chosen as belonging to the target class indicated in the corresponding column heading. The last column lists the correct recognition rate for each row. The overall percentage of correctly classified test images was 98.10%.

TABLE VII
Recognition Test as Four Class Problem Under Configuration Variants EOC, When Training Done Using Both 15° and 17° Depression Images, for 80×80 Images

	BMP-2	BRDM-2	BTR-70	T-72	Percent
BMP-2 #1	385	0	3	40	89.95%
BMP-2 #2	367	0	15	45	85.95%
T-72 #2	70	2	18	336	78.87%
T-72 #4	106	25	40	402	70.16%
T-72 #5	48	20	3	502	87.61%
T-72 #6	118	10	5	440	76.79%
T-72 #7	83	17	2	465	82.01%

Note: The row headings indicate the true target class for all images represented in each row. Each row contains the number of images from the test set that were chosen as belonging to the target class indicated in the corresponding column heading. The last column lists the correct recognition rate for each row. The overall percentage of correctly classified test images was 81.22%.

variation EOC (both large and small training sets), and version variation EOC (both large and small training sets), respectively. For the results represented in these confusion matrices, all thresholds τ_a from (26) were set to zero. Table V shows that a very high recognition rate was achieved by both the BMP-2 and T-72 whose models incorporated images from vehicles of differing configurations and, in the case of the T-72, different versions. All images which were misclassified were incorrectly classified as either

TABLE VIII
Recognition Test as Four Class Problem Under Configuration Variants EOC, When Training Done Using Only 17° Depression Images, for 80×80 Images

	BMP-2	BRDM-2	BTR-70	T-72	Percent
BMP-2 #1	375	0	5	48	87.62%
BMP-2 #2	355	0	17	55	83.14%
T-72 #2	79	0	13	334	78.40%
T-72 #4	106	29	42	396	69.11%
T-72 #5	58	7	3	505	88.13%
T-72 #6	113	12	8	440	76.79%
T-72 #7	87	11	5	464	81.83%

Note: The row headings indicate the true target class for all images represented in each row. Each row contains the number of images from the test set that were chosen as belonging to the target class indicated in the corresponding column heading. The last column lists the correct recognition rate for each row. The overall percentage of correctly classified test images was 80.43%.

TABLE IX
Recognition Test as Four Class Problem Under Version Variants EOC, When Training Done Using Both 15° and 17° Depression Images, for 80×80 Images

	BMP-2	BRDM-2	BTR-70	T-72	Percent
T-72 #3	37	0	16	366	87.35%
T-72 #8	69	11	10	482	84.27%
T-72 #9	85	8	16	464	80.98%
T-72 #10	104	22	29	418	72.95%
T-72 #11	135	16	3	419	73.12%

Note: The row headings indicate the true target class for all images represented in each row. Each row contains the number of images from the test set that were chosen as belonging to the target class indicated in the corresponding column heading. The last column lists the correct recognition rate for each row. The overall percentage of correctly classified test images was 79.30%.

BMP-2 or T-72. Table VI shows that when registration uncertainty was introduced and the recognition algorithms accounted for the uncertainty, the overall performance was not decreased from the case of SOC.

Table VII, which represents performance under the configuration variation EOC with a large set of training data, shows that the variants of the BMP-2 fared better in the performance analysis than did most of the T-72 variants. Comparison with Table VIII, which represents the same EOC but with a smaller set of training data, shows that less training data resulted in worse performance for all but two test vehicles. Tables IX and X contain confusion matrices for the version variation EOC with large and small training sets, respectively. All test targets yielded better recognition performance when models were constructed from the larger training sets. Comparison of Table IX with Table VII shows that when large training sets were used, the T-72 version variants performed similarly to the T-72 configuration variants with correct classification rates between 70% and 88%. Comparison of Table X with Table VIII,

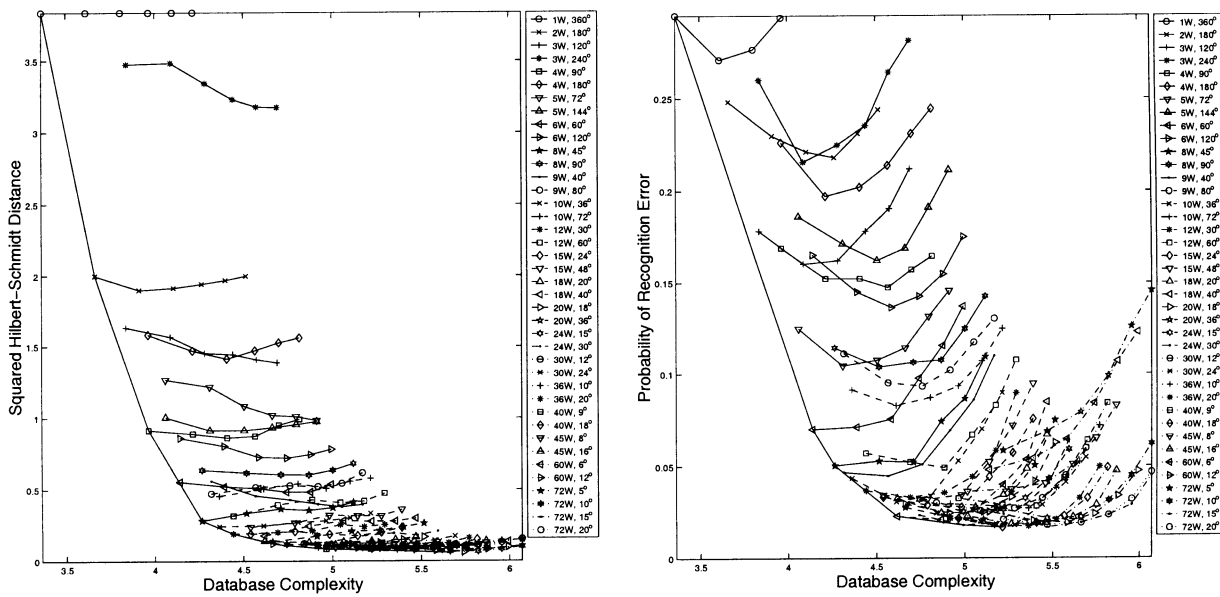


Fig. 5. Performance-complexity pairs for range of parameterizations in four class problem under SOC. Left and right panes show average orientation estimation error and average recognition error rate over all test images as function of database complexity which is defined as log of number of floating point values in database. Each point marker in scatter plot denotes combination of performance and complexity that is achievable with combination of number of training windows and training window sizes as given in legend. Each of six point markers along the curves represent successively larger image sizes from 48×48 to 128×128 . Solid convex outer boundary represents best achievable performance at any given complexity.

TABLE X
Recognition Test as Four Class Problem Under Version Variants EOC, When Training Done Using Only 17° Depression Images, for 80×80 Images

	BMP-2	BRDM-2	BTR-70	T-72	Percent
T-72 #3	42	0	17	360	85.92%
T-72 #8	80	8	7	477	83.39%
T-72 #9	100	18	12	443	77.31%
T-72 #10	120	29	16	408	71.20%
T-72 #11	145	30	3	395	68.94%

Note: The row headings indicate the true target class for all images represented in each row. Each row contains the number of images from the test set that were chosen as belonging to the target class indicated in the corresponding column heading. The last column lists the correct recognition rate for each row. The overall percentage of correctly classified test images was 76.86%.

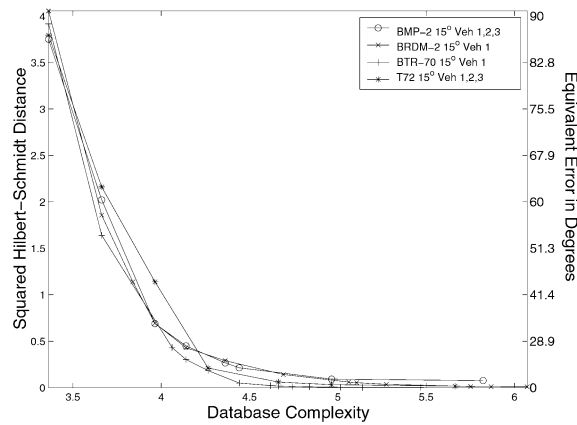
however, shows that when small training sets were used, the T-72 version variants, with classification rates ranging from 69% to 86%, performed slightly worse overall than did the BMP-2 configuration variants which exhibited rates ranging from 69% to over 88%. The tendency for images to be misclassified as either BMP-2 or T-72 which was mentioned in the case of SOC is apparent in these sets of tests as well. Unlike the SOC tests, in these EOC tests the BMP-2 and T-72 models were trained with images of only one vehicle.

C. Performance and Complexity

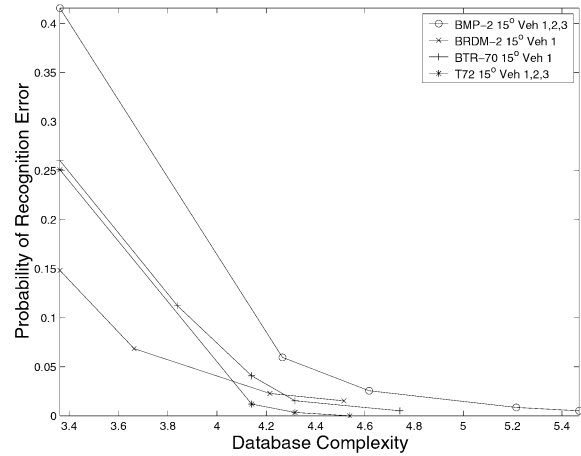
We conducted an analysis of performance versus variation in model complexity. The model complexity

is a function of the target chip size and model resolution (number of images in the trained database set). Here, complexity is defined as the logarithm of the total number of pixels stored in the trained database per target type. The logarithm of the number of computations required to evaluate the likelihood function is roughly proportional to this complexity measure. Orientation estimation performance, in terms of the average squared Hilbert-Schmidt distance between the actual and estimated orientations, and target recognition performance, in terms of probability of error in recognition, were obtained for forty combinations of model resolution and training window width and for six different target chip sizes for a total of 240 combinations.

The left pane of Fig. 5 details the orientation estimation performance averaged over all four targets under SOC as a function of database complexity. Each curve shown in the legend represents the average orientation estimation error for a different combination of number of azimuth windows and window widths for six different image sizes. For example, the solid curve with x-shaped point markers shows the average orientation estimation error when models containing only two windows, each trained over 180° of azimuth are used. The left most point marker on this curve shows that for 48×48 images, an average orientation estimation error of 2 can be achieved and the database complexity for this combination is about 3.7. Each point marker to the right represents the combination of performance and complexity attainable when images of size 64×64



Estimation Error vs. Complexity



Recognition Error vs. Complexity

Fig. 6. Performance versus complexity curves. Each curve indicates best achievable performance as function of database complexity for given target. Orientation estimation error is average of squared Hilbert-Schmidt distance over test set, recognition performance is in terms of probability of recognition error, complexity is defined as logarithm of total number of pixels used in trained database per target type.

through 128×128 are used, respectively. These curves are relatively flat indicating that, for the most part, changes in image size do not result in large changes in orientation estimation performance. The solid convex outer boundary represents the best achievable performance as a function of database complexity over the model parameters tested. The combinations of performance and database complexity indicated by points on a segment of this convex boundary between any two model parameters are achievable by selecting one of the models with probability P and the other with probability $1 - P$. This convex boundary drops sharply at the left side of the graph but becomes nearly horizontal at database complexities between 4.5 and 5 indicating that increasing complexity beyond this point does not significantly improve orientation estimation performance.

The right pane of Fig. 5 details the target recognition performance averaged over all four targets under SOC as a function of database complexity. Here, the curves indicating performance as a function of image size for a given combination of number of windows and width of training intervals are not predominately horizontal but in fact show significant curvature. This indicates that, unlike the case of orientation estimation, the target recognition performance is quite dependent on image size. The convex boundary showing best achievable performance as a function of database complexity in this graph is very similar to that of the left pane. It drops sharply on the left edge of the graph but levels off between database complexities 4.5 and 5 indicating that increases in database complexity beyond this range do not yield significant increases in performance.

Fig. 6 summarizes the performance as a function of complexity for all four targets in terms of

orientation estimation error and probability of recognition error. These graphs show the best achievable performance (lowest achievable error) line as a function of database complexity for each target. This figure shows that for all four targets, the average error, in terms of both orientation estimation and target recognition, at first decreases rapidly with database complexity, but then tapers off, changing only slightly as complexity is increased. Orientation estimation performance versus database complexity is very similar for all targets while target recognition performance varies considerably among the four.

VI. PERFORMANCE IN THE TEN CLASS PROBLEM

Orientation estimation and target recognition performance of the conditionally Gaussian model were assessed under SOC for an expanded set of ten target classes. These ten targets include the four from the previous section with six additions. A summary of the datasets used is provided in Table XI. For each of the ten classes, training was performed using images formed with a 17° depression angle and testing was performed using 15° depression angle images. This database was trained with 72 overlapping windows each of width 10° .

A. Orientation Estimation Performance

Orientation estimation performance was computed for each target class and for a range of image sizes. The left pane of Fig. 7 shows how the overall estimation error varies with image size. It can be seen that overall estimation performance drops as the image size gets too small or too large with a minimum error of around 0.05, corresponding to slightly more than 9° , at image sizes of 80×80 and

TABLE XI
MSTAR Dataset Used in Ten Class Performance Analysis

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

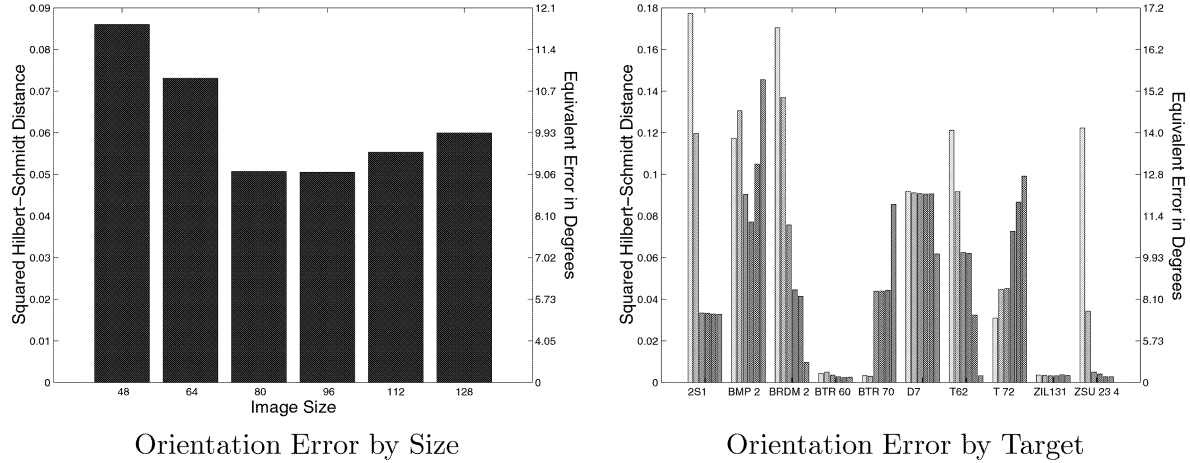


Fig. 7. Left Pane: Average squared Hilbert-Schmidt distance over all targets as function of image size. Bar labels indicate number of pixels along one edge of square SAR image. Right Pane: Average squared Hilbert-Schmidt distance for each target and image size. For each target, left-most bar represents estimation error for 48×48 pixel images and the right-most bar for 128×128 pixel images.

96×96 pixels. The right pane of Fig. 7 shows how this estimation error varies with both image size and target class. For the target classes 2S1, BRDM2, T62, and ZSU 23 4, the estimation error grows sharply as the image size shrinks and consequently begins to encroach on the edges of the target and its shadow. While for the BMP2, BTR 70, and T72, the estimation error grows as the image size gets large and includes more background noise which effectively decreases the SNR.

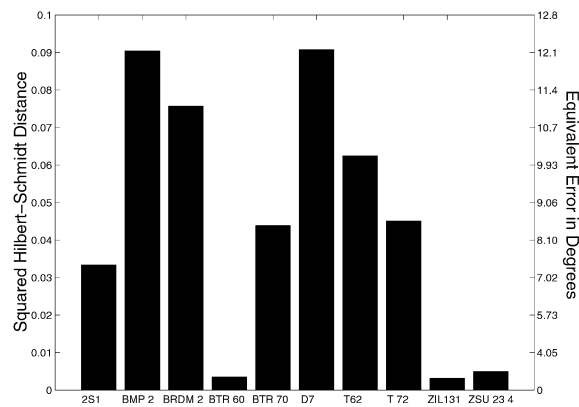
The orientation estimation error for all targets is shown in the left pane of Fig. 8 for the 80×80 pixel case. The right pane of Fig. 8 contains a graph showing the percentage of images, also of size 80×80 , from each target class which have an orientation estimation error in excess of 175° . It can be seen that, with the exception of the target class BTR 70, for those targets with a large overall estimation error, a significant portion of that error is due to images which yield an estimate that is incorrect by 180° . Such targets thus have a 180° ambiguity at certain orientations. For example, the target type BMP-2, when at an orientation of approximately 90° (broadside to the sensor, facing left), is several times

estimated to have an orientation of approximately 270° (broadside to the sensor, facing right), and vice-versa.

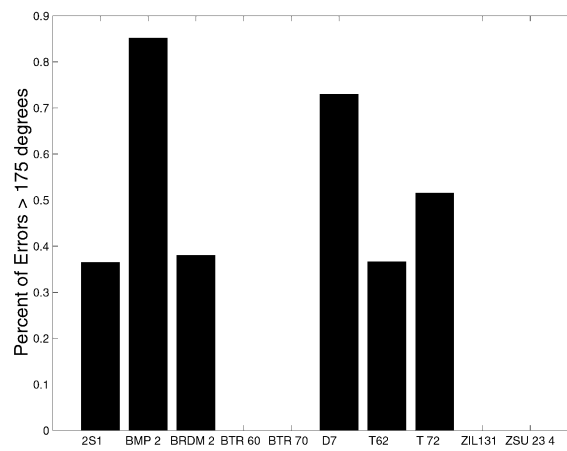
B. Ten Class Target Recognition Performance

Fig. 9 shows the experimentally obtained ROC for both one of the best and one of the worst performing target classes. These plots, derived from testing 80×80 pixel images, show the probability of correct recognition versus the probability of false alarm in a binary hypothesis test between one target and each of the remaining nine targets. In the left pane of Fig. 9 it can be seen that the performance of the ZIL131 is nearly ideal against all targets with a probability of correct recognition close to one while simultaneously having a probability of false alarm which is close to zero. In the right pane of Fig. 9 it can be seen that for the 2S1, a probability of false alarm of nearly 2% must be tolerated in order to achieve a probability of correct recognition as high as 98% in tests against the BRDM 2.

A confusion matrix for the ten class recognition problem is shown in Table XII for the case where all



Orientation Error by Target



Percent of Images Yielding 180° Errors

Fig. 8. Left Pane: Average Hilbert-Schmidt distance squared for 80×80 pixel images by target. Right Pane: Percentage of 80×80 pixel images which result in estimation errors near 180° .

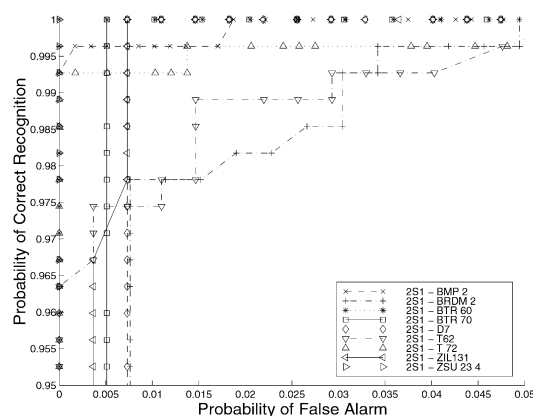
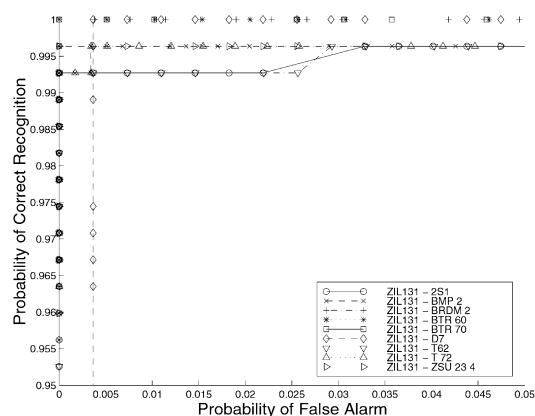


Fig. 9. Left Pane: Experimental ROC for the binary hypothesis test between ZIL131 and each of the other targets. Right Pane: Experimental ROC for binary hypothesis test between 2S1 and each of the other targets.

thresholds τ_a from (26) are zero. This table shows the total number of images from each target class indicated in the row headings which are determined to be from the target class indicated in the column headings. These results are for images of size 80×80 pixels. Note that, as indicated by the ROC curves shown in Fig. 9, correct recognition of the 2S1 was somewhat below the performance of the ZIL131. The overall percent of correct classification for this set of data is 97.18%.

Table XIII details the overall percent of correct classification as a function of image size, again with all zero threshold values. As previously noted for the case of orientation estimation performance, overall performance of target recognition drops off as image sizes become too small or too large with the best results for images which are 80×80 pixels.

VII. CONCLUSIONS AND FUTURE WORK

We have extended to SAR data a conditionally Gaussian model and have applied it to target

recognition and orientation estimation. We have detailed all algorithms involved and, using actual SAR data which is publicly available, demonstrated reproducible results. These results are quite promising regarding the usefulness of the conditionally Gaussian model with recognition rates over 97% for a ten class problem under standard operating conditions, and over 81% and 79% for a four class problem under configuration and version variation extended operating conditions, respectively. Further, we have shown that both configuration and version variants of the same target class can be incorporated into a single model which yields good ATR performance across the range of variation for which it was trained.

A direct comparison with the results reported by others in the field is not possible because of wide differences in data sets used and approach parameterizations, however numbers presented in several recent papers indicate that this approach yields at least comparable results. In [22], the authors report a three class experiment in which the probability of detection was fixed at 90%. Under SOC (BMP2-3,

TABLE XII
Recognition as Ten Class Problem for 80×80 Pixel Images

	2S1	BMP 2	BRDM2	BTR 60	BTR 70	D7	T62	T72	ZIL131	ZSU 23 4	
2S1	262	0	0	0	0	0	4	8	0	0	95.62%
BMP 2	0	581	0	0	0	0	0	6	0	0	98.98%
BRDM2	5	3	227	1	0	14	3	5	4	1	85.31%
BTR 60	1	0	0	193	0	0	0	0	0	1	98.97%
BTR 70	4	5	0	0	184	0	0	3	0	0	93.88%
D7	2	0	0	0	0	271	1	0	0	0	98.91%
T62	1	0	0	0	0	0	259	11	2	0	94.87%
T72	0	0	0	0	0	0	0	582	0	0	100%
ZIL131	0	0	0	0	0	0	2	0	272	0	99.27%
ZSU 23 4	0	0	0	0	0	2	0	1	0	271	98.91%

Note: The row headings indicate the true target class for all images represented in each row. For each row, the numbers listed indicate how many images from the test set were chosen as belonging to the target class indicated in the corresponding column heading. The rightmost column shows the overall percentage of correct classification for each target.

TABLE XIII
Overall Percentage of Images Correctly Recognized as Function of Image Size

Image Size	48×48	64×64	80×80	96×96	112×112	128×128
Percentage	96.21	96.90	97.18	96.87	96.18	95.05

BTR70-2, and T72-1) their experiments yielded a 95.19% correct recognition rate. When restricted to configuration variations (BMP2-1, BMP2-2, BTR70-1, BTR70-4, T72-2) a rate of 79.55% was reported. When restricted to version variations (BTR70-3 and T72-3) a recognition rate of 80.43% was reported. In [12], the authors report a ten class experiment that, when restricted to configuration variations (which does not include the version variant T72 #3), yielded a 95.69% correct recognition rate with conventionally processed SAR images.

The scientific framework of examining algorithm performance across a wide range of parameterizations as described in this paper provides a rigorous setting in which performance can be analyzed and different ATR methods can be compared. The recent work of O'Sullivan and Schmid [19] describes analytical performance results for binary detection problems versus the size of training sets. The performance complexity tradeoff described here can be used as a starting point for future research on this important problem.

There are several approaches that we and others have used [9, 12] to incorporate the shapes of objects by separating target and clutter pixels. The application of a theory of target-clutter segmentation to our algorithm is a subject of current and future research as is the incorporation of theoretical approaches to the rejection of confusor targets. Other ongoing research [3, 16] involves a direct comparison of our model with models of others [12, 13, 25]. Finally, comparison with a Rician statistical model is being pursued [2].

ACKNOWLEDGMENTS

These MSTAR data were obtained courtesy of Thomas J. Burns, the MSTAR program manager for DARPA.

REFERENCES

- [1] Chiang, H. C., Moses, R. L., and Irving, W. W. (1999) Performance estimation of model-based automatic target recognition using attributed scattering center features. In *Proceedings of the 10th International Conference on Image Analysis and Processing*, Sept. 1999.
- [2] DeVore, M. D., Lanterman, A. D., and O'Sullivan, J. A. (2000) ATR performance of a Rician model for SAR images. In F. A. Sadjadi (Ed.), *Automatic Target Recognition X* (Proceedings of SPIE), **4050** (2000).
- [3] DeVore, M. D., and O'Sullivan, J. A. 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.
- [4] Grenander, U., Miller, M. I., and Srivastava, A. (1999) Hilbert-Schmidt lower bounds for estimators on matrix lie groups for ATR. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**, 8 (Aug. 1999), 790–802.
- [5] Jacobs, S. P. (1997) Automatic target recognition using high resolution radar range-profiles. Ph.D. dissertation, Washington University, St. Louis, MO, May 1997.
- [6] Jacobs, S. P., and O'Sullivan, J. A. (1997) High-resolution radar models for joint tracking and recognition. In F. A. Sadjadi (Ed.), *Automatic Target Recognition VII* (Proceedings of SPIE), **3069** (1997), 94–105.

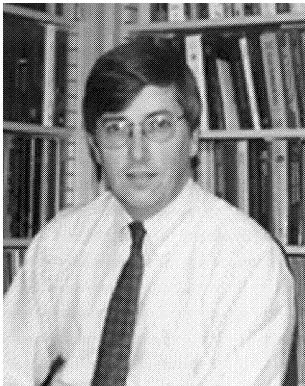
- [7] Jacobs, S. P., and O'Sullivan, J. A. (1997)
Automatic target recognition using sequences of high resolution radar range-profiles.
IEEE Transactions on Aerospace and Electronic Systems, **36**, 2 (April 2000), 364–382.
- [8] Kedia, V. (1998)
Performance analysis of automatic target recognition using high resolution radar.
Master's thesis, Washington University, St. Louis, MO, Aug. 1998.
- [9] Lanterman, A. D., O'Sullivan, J. A., and Miller, M. I. (1999)
Kullback-Liebler distances for quantifying clutter and models.
Optical Engineering, **38**, 12 (Dec. 1999), 2134–2146.
- [10] Miller, M. I., Grenander, U., O'Sullivan, J. A., and Snyder, D. L. (1997)
Automatic target recognition organized via jump-diffusion algorithms.
IEEE Transactions on Image Processing, **6**, 1 (Jan. 1997), 157–174.
- [11] Mossing, J. C., and Ross, T. D. (1998)
An evaluation of SAR ATR algorithm performance sensitivity to MSTAR extended operating conditions.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery V* (Proceedings of SPIE), **3370** (1998), 554–565.
- [12] Novak, L. M., Owirka, G. J., Brower, W. S., and Weaver, A. L. (1997)
The automatic target recognition system in SAIP.
Lincoln Laboratory Journal, **10**, 2 (1997), 187–201.
- [13] Novak, L. M., Owirka, G. J., and Weaver, A. L. (1999)
Automatic target recognition using enhanced resolution SAR data.
IEEE Transactions on Aerospace and Electronic Systems, **35**, 1 (Jan. 1999), 157–175.
- [14] O'Sullivan, J. A., Blahut, R. E., and Snyder, D. L. (1998)
Information-theoretic image formation.
IEEE Transactions on Information Theory, **44**, 6 (Oct. 1998), 2094–2123.
- [15] O'Sullivan, J. A., and DeVore, M. D. (1999)
Performance analysis of ATR from SAR imagery.
In *Proceedings of the 33rd Annual Conference on Information Sciences and Systems*, Apr. 1999.
- [16] O'Sullivan, J. A., and DeVore, M. D. (2000)
Performance-complexity tradeoffs for several approaches to ATR from SAR images.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery VII* (Proceedings of SPIE), **4053** (2000).
- [17] O'Sullivan, J. A., Jacobs, S. P., and Kedia, V. (1997)
Orientation estimation bounds for high resolution radar data.
In *Thirty Fifth Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 1997, 112–121.
- [18] O'Sullivan, J. A., Jacobs, S. P., and Kedia, V. (1998)
Stochastic models and performance bounds for pose estimation using high resolution radar data.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery V* (Proceedings of SPIE), **3370** (1998), 576–587.
- [19] O'Sullivan, J. A., and Schmid, N. A. (1999)
Dependence of ATR system performance on size of training sets.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery VI* (Proceedings of SPIE), **3721** (1999), 730–739.
- [20] Owirka, G. J., and Novak, L. M. (1994)
A new SAR ATR algorithm suite.
In *Algorithms for Synthetic Aperture Radar Imagery* (Proceedings of SPIE), **2230** (1994), 336–343.
- [21] Ross, T. D., Bradley, J. J., Hudson, L. J., and O'Connor, M. P. (1999)
SAR ATR—so what's the problem? An MSTAR perspective.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery VI* (Proceedings of SPIE), **3721** (1999), 566–573.
- [22] Ross, T. D., Worrell, S. W., Mossing, J. C., and Bryant, M. L. (1998)
Standard SAR ATR evaluation experiments using the MSTAR public release data set.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery V* (Proceedings of SPIE), **3370** (1998), 566–573.
- [23] Srivastava, A. (1996)
Inferences on transformation groups generating patterns on rigid motions.
Ph.D. dissertation, Washington University, St. Louis, MO, July 1996.
- [24] Theera-Umpon, N., Khabou, M. A., Gader, P. D., Keller, J. M., Shi, H., and Li, H. (1998)
Detection and classification of MSTAR objects via morphological shared-weight neural networks.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery V* (Proceedings of SPIE), **3370** (1998), 530–540.
- [25] Wissinger, J., Washburn, R. B., Friedland, N. S., Nowicki, A., Morgan, D. R., Chong, C., and Fung, R. (1996)
Search algorithms for model based SAR ATR.
In E. G. Zelnio and R. J. Douglass (Ed.), *Algorithms for Synthetic Aperture Radar Imagery III* (Proceedings of SPIE), **2757** (1996), 279–293.
- [26] Worrell, S. W., Parker, S., and Bryant, M. L. (1997)
Class separability assessments and MSE algorithm robustness.
In E. G. Zelnio (Ed.), *Algorithms for Synthetic Aperture Radar Imagery IV* (Proceedings of SPIE), **3070** (1997), 294–304.

Joseph A. O'Sullivan (S'83—M'86—SM'92) was born in St. Louis, MO, on January 7, 1960. He received the B.S., M.S., and Ph.D. all in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1982, 1984, and 1986, respectively.

In 1986, he joined the faculty in the Department of Electrical Engineering at Washington University, where he is now a Professor. He has joint appointments in the Department of Radiology and the Department of Biomedical Engineering. He was a founding member and is now Director of the Electronic Systems and Signals Research Laboratory. He is a member of the Magnetics and Information Science Center and the Center for Imaging Science at Washington University. His research interests include information-theoretic imaging, automatic target recognition, systems integration issues in magnetic recording, CT imaging in the presence of known high density attenuators, information hiding, and hyperspectral imaging.

Dr. O'Sullivan was Secretary of the Faculty Senate, Secretary of the Senate Council, and Faculty Representative to the Board of Trustees at Washington University from 1995 to 1998. He is a consultant for and on the Board of Directors of Abacus Controls, Inc. He was the Publications Editor for the *IEEE Transactions on Information Theory* from 1992 to 1995, is currently Associate Editor for Detection and Estimation, and was a Guest Associate Editor for the 2000 Special Issue on Information Theoretic Imaging. He has served on the organizing and program committees for several conferences and workshops. He was co-chair of the 1999 Information Theory Workshop on Detection, Estimation, Classification, and Imaging.

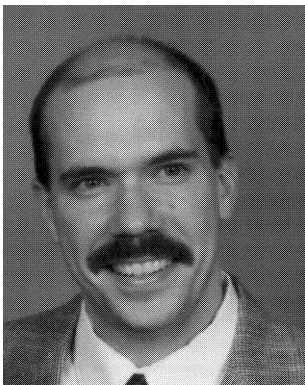
Dr. O'Sullivan is active in local IEEE activities as well, including being chair of the St. Louis section of the IEEE in 1994. He is a member of Eta Kappa Nu and SPIE. He was awarded an IEEE Third Millennium Medal.

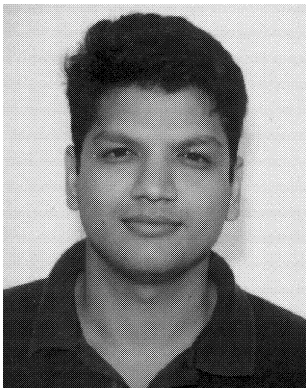


Michael D. DeVore (S'92) received B.S. degrees in electrical engineering, computer engineering, and mathematics from the University of Missouri at Columbia in 1991 and an M.S. degree in electrical engineering in 1993.

From 1993 he worked in, and from 1995 through 1998 he led, the System Integrity and Integration and Operations Support teams for Amdocs, Inc. in St. Louis. He is currently pursuing a D.Sc. degree at Washington University in St. Louis doing research in the area of automatic target recognition from synthetic aperture radar data. His research interests include target recognition, statistical model assessment, model representation and storage, and performance of inference systems subject to complexity constraints.

Mr. DeVore is a student member of SPIE and a Boeing Foundation Fellow.





Vikas Kedia received the M.S. degree in electrical engineering with a certificate in imaging science from Washington University in St. Louis, MO, in 1998. He received the Bachelor of Technology in electrical engineering from Indian Institute of Technology, Bombay, India, in 1996.

Mr. Kedia is currently employed with Citibank in New York.



Michael I. Miller received his M.S. in electrical and computer engineering in 1976 and his Ph.D. in biomedical engineering in 1983 from the Johns Hopkins University, Baltimore, MD.

He became the Director of the Whiting School of Engineering, Center for Imaging Science in July 1998. Prior to that he was the Newton Professor of Biomedical and Electrical Engineering at Washington University in St. Louis. He is co-founder of two companies working in medical image analysis and image understanding. Dr. Miller is currently the Director of the National Center for Imaging Science, a consortium of seven universities working in the areas of automated target recognition and image understanding and is a scientific partner in the National Partnership for Computing Infrastructure (NPACI).

He was the recipient of the National Science Foundation's Presidential Young Investigator Award in 1985 and the Paul Ehrlich Graduate Award in 1983.