PROCEEDINGS OF SPIE

Probabilistic approach to model extraction from training data

Michael DeVore, Joseph O'Sullivan, Sushil Anand, Natalia Schmid

SPIE.

Probabilistic Approach to Model Extraction from Training Data

Michael D. DeVore[a], Joseph A. O'Sullivan[b], Sushil Anand[c], Natalia A. Schmid[d]

Electronic Systems and Signals Research Laboratory
Department of Electrical Engineering, Washington University, St. Louis, MO 63130

ABSTRACT

Many of the approaches to automatic target recognition (ATR) for synthetic aperture radar (SAR) images that have been proposed in the literature fall into one of two broad classes, those based on prediction of images from models (CAD or otherwise) of the targets and those based on templates describing typical received images which are often estimated from sample data. Systems utilizing model-based prediction typically synthesize an expected SAR image given some target class and pose and then search for the combination of class and pose which maximizes some match metric between the synthesized and observed images. This approach has the advantage of being robust with respect to target pose and articulation not previously encountered but does require detailed models of the targets of interest. On the other hand, template-based systems typically do not require detailed target models but instead store expected images for a range of targets and poses based on previous observations (training data) and then search for the template which most closely represents the observed image. We consider the design and use of probabilistic models for targets developed from training data which do not require CAD models of the targets but which can be used in a hypothesize-and-predict manner similar to other model-based approaches. The construction of such models requires the extraction from training data of functions which characterize the target radar cross section in terms of target class, pose, articulation, and other sources of variability. We demonstrate this approach using a conditionally Gaussian model for SAR image data and under that model develop the tools required to determine target models and to use those models to solve inference problems from an image of an unknown target. The conditionally Gaussian model is applied in a target-centered reference frame resulting in a probabilistic model on the surface of the target. The model is segmented based on the information content in regions of the target space. Modeling radar power variability and target positional uncertainty results in improved accuracy. Performance results are presented for both target classification and orientation estimation using the publicly available MSTAR dataset.

Keywords: model-based recognition, conditionally Gaussian model, model extraction

1. INTRODUCTION

A large number of approaches to automatic target recognition (ATR) from synthetic aperture radar (SAR) imagery have been reported in the literature and many of these can be divided into one of two broad classes. Approaches generally referred to as model-based often rely on physical models of the shape and composition of targets. These models are used to predict observable features in images of the target for arbitrary viewing conditions such as target pose relative to the radar platform. Recognition of an unknown target in an image consists of searching over the set of plausible poses and reporting the combination of target class and pose that most closely match the feature set actually observable in the image. The predict-extract-match-search methodology[10] of the Moving and Stationary Target Acquisition and Recognition (MSTAR) program is an example of this kind of approach. Such methods require detailed models of targets and a highly accurate characterization of electromagnetic phenomenology to support feature prediction. They have the advantage of being relatively robust with respect to viewing conditions, such as target pose or partial obscuration, not originally envisioned.

Approaches generally referred to as template-based depend on the collection of statistics characterizing targets from sample images. The statistics are collected from sample observations (training data) for a wide variety of viewing conditions and are stored as part of the ATR implementation. An image of an unknown target is classified by finding the set of statistics that most accurately characterize the image and the corresponding target class and pose are reported. Examples of this kind of approach are found in Ross, *et al.*[7] and Owirka, *et al.*[6] Such methods do not require detailed models of targets, relying instead on training observations of actual data which is seen as an

E-mail: [a]mdd2@cis.wustl.edu, [b]jao@ee.wustl.edu, [c]sushil@essrl.wustl.edu, [d]nar@essrl.wustl.edu

advantage by some concerned about the complexity of actual SAR imagery. These methods are generally not adept at classifying images obtained under conditions significantly different than those represented by the training data.

We describe a method that combines concepts from these two types of approaches, employing models of targets estimated from training images. These target models are based on stochastic models for SAR imagery that attempt to capture the complexity of SAR returns through conditionally Gaussian probability densities.[5] Target models are used to predict statistical properties of radar images of targets at arbitrary poses relative to the radar platform. Classification of SAR images is performed by maximizing the likelihood of the observation given these predictions over all possible target classes and poses. Further, the target models admit a hierarchical representation that allows inference in terms of a successively-refinable statistical characterization. Successive refinability allows ATR algorithms to proceed until some maximum quantity of resources, such as elapsed time, processor cycles, or network bandwidth, has been consumed. Further, this maximum quantity can be dynamically adjusted according to operating requirements without altering the implementation. The approach is demonstrated on a four-class problem using actual SAR data that is publicly available from the MSTAR program.

The conditionally Gaussian model for SAR imagery and algorithms for recognition and rejection of confuser targets is described in Section 2. Estimation of target models from training observations accounting for variability of target pose is discussed in Section 3. Section 4 details the approach to successively-refinable models and describes how these can facilitate the search for maxima of the likelihood function. Results from a four-class experiment are demonstrated in Section 5 and conclusions follow in Section 6.

2. CONDITIONALLY GAUSSIAN MODEL

The conditionally Gaussian model for SAR image formation holds that the radar returns from region l in the scene corresponding to pixel i in the image are instances of a complex Gaussian random variable conditioned on the contents of the scene, its position and orientation relative to the radar platform, and the power in the radar illumination. The observations have mean zero and the returns from nonoverlapping regions are independent. If the pixels in a SAR image $\mathbf{R}$ correspond to nonoverlapping regions in the scene, the probability density function for an observed SAR image can be written as the product over all pixels i of conditional Gaussian density functions. That is,

$$p(\mathbf{r}|a, \Theta, c^2) = \prod_{i=1}^{N} \frac{1}{\pi c^2 \sigma_i^2(a, \Theta)} e^{-|r_i|^2/c^2 \sigma_i^2(a, \Theta)}, \tag{1}$$

where r_i is the complex value observed in pixel i, $\sigma_i^2(a, \Theta)$ is the variance of the radar reflection from the region of the scene corresponding to pixel i, a and Θ are the class and pose of the target in the scene, respectively, and c^2 is a constant over all pixels in the image accounting for radar power fluctuation from some nominal power.

In general, the pose parameter Θ is a vector with components representing the location of target a in $\Re^3$ and orientation in $\mathbf{SO(3)}$ relative to the radar platform. The pose parameter also specifies the states of all articulated components of the target, such as turrets, barrels, and hatches. For the experiments we present, we will restrict attention to recognition of ground-based vehicles on flat terrain, with unarticulated components, and imaged from a known, fixed depression angle. In this restricted case, Θ can be simplified to represent the location of the target in the radar slant plane, $\mathbf{s} \in \Re^2$, and the azimuth angle θ relative to the radar platform.

Given an image $\mathbf{r}$ of an unknown target with unknown pose, joint recognition and pose estimation can be performed by maximizing the probability density in (1) over all combinations of a, Θ, and c^2. That is, the generalized likelihood ratio test (GLRT) is employed as

$$\begin{bmatrix} \hat{a} \\ \hat{\Theta} \\ \hat{c}^2 \end{bmatrix} = \underset{[a, \Theta, c^2]^T}{\operatorname{argmax}} p(\mathbf{r}|a, \Theta, c^2). \tag{2}$$

It is generally desirable to restrict the product in (1) to only those pixels i in the image that correspond to regions in the scene occupied by the target of interest. This can be accommodated by introducing a binary mask $\mathbf{I}(a, \Theta)$, an indicator function that takes the value one for every pixel representing a region on the target and zero elsewhere. Because the number and location of pixels contributing to the product now vary with the hypothesized target and pose, the simple GLRT of (2) is no longer appropriate. In a related dimensionality reduction problem, Schmid and O'Sullivan[9] introduce a density function representing what they term a null-hypothesis and show approximate

asymptotic results indicating improved recognition performance over unmasked data in simulated high resolution radar experiments. This density has the form of (1) but does not represent the presence of any particular target, rather it is a mathematical device that introduces a free parameter to facilitate the decision rule. Alternatively, we can think of the null-hypothesis as representing a null-target a_0 for which all radar return variances are some constant value, $\sigma_i^2(a_0, \Theta) = \xi^2$ for all Θ. The GLRT is modified to compute likelihood ratios relative to this null-hypothesis as

$$\begin{bmatrix} \hat{a} \\ \hat{\Theta} \end{bmatrix} = \operatorname*{argmax}_{[a,\Theta]^T} \prod_{i=1}^{N} I_i(a, \Theta) \left[\frac{1}{\pi c^2 \sigma_i^2(a, \Theta)} e^{-|r_i|^2 / c^2 \sigma_i^2(a, \Theta)} \right] / \left[\frac{1}{\pi \xi^2} e^{-|r_i|^2 / \xi^2} \right], \tag{3}$$

where

$$\hat{c}^2(a, \Theta) = \frac{1}{\sum_i I_i(a, \Theta)} \sum_{i=1}^{N} \frac{|r_i|^2}{\sigma_i^2(a, \Theta)} I_i(a, \Theta). \tag{4}$$

is the maximum-likelihood estimate of c^2. DeVore, $et\ al.$[3] demonstrate this approach on MSTAR data for $\xi^2 = 0.0028$ and show that the results are relatively insensitive to choices of ξ^2 over the interval $[10^{-16}, 0.0084]$.

After a decision has been reached concerning the most likely target and pose, it is often desirable to determine whether the image has come from the image declared or whether it represents a target not considered. This problem of confuser rejection can be approached by noting that $|r_i|^2$ is a one-sample estimate of the variance of the value in pixel i. If the estimated variance differs significantly from the expected variance given $\hat{a}$, $\hat{\Theta}$, and $\hat{c}^2$ over the entire image, we may conclude that the image is of a different target. This difference can be quantified in terms of the average relative entropy between a conditionally Gaussian distribution with variance $|r_i|^2$ and one with variance $c^2 \sigma_i^2(\hat{a}, \hat{\Theta})$. The image can be rejected as a confuser if the relative entropy exceeds some threshold γ as

$$\frac{1}{\sum_i I_i(\hat{a}, \hat{\Theta})} \sum_{i=1}^{N} D(p(\cdot|\sigma^2 = |r_i|^2) \| p(\cdot|\sigma^2 = c^2 \sigma_i^2(\hat{a}, \hat{\Theta}))) > \gamma, \tag{5}$$

where $D(p\|q) = \int p \log(p/q)$ is the relative entropy between density functions p and q.

3. TARGET MODEL ESTIMATION

The variance functions parameterizing the conditionally Gaussian model (1) are unknown in general and must be estimated from training data. The variance functions are themselves parameterized by the continuous-valued pose Θ, however a practical limitation of data collection is that we can collect sample images at only a finite number of poses. The problem of estimating an arbitrary function of continuous parameters from a finite set of data is ill-posed and so we consider restricting the class of functions under consideration. In the examples considered, we impose a piecewise constant approximation in which the variance in the radar return from a region l on the target surface when imaged from azimuth angle θ, $\hat{\zeta}_l(a, \theta)$ is piecewise constant in θ over N_w intervals. Since translations of the scene in the range direction affects the reflected power uniformly over the image and this quantity is accounted for in the variable c^2 from the previous section, the range to the target need not be explicitly accounted for in the representation of $\hat{\zeta}_l$. Also, since a translation of the scene in the cross-range direction results in a simple shift of the resulting image, cross-range position need not be accounted for in $\hat{\zeta}_l$.

Given a set of M training images $\{\mathbf{T}_1', \mathbf{T}_2', \ldots, \mathbf{T}_m'\}$ of a target a taken from azimuth angles $\{\phi_1, \phi_2, \ldots, \phi_M\}$ at positions $\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M\}$, registration in azimuth and position will result in images such that pixel location i in each image corresponds to the same region l on the target. Letting $U(-\phi, -\mathbf{s})$ denote a translation and rotation transformation of a SAR image, the set of images $\mathcal{T} = \{\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_M\}$, where $\mathbf{T}_j = U(-\phi_j, -\mathbf{s}_j)\mathbf{T}_j'$, will be registered with respect to an image at $0°$ azimuth. An estimate of $\hat{\zeta}_l(a, \theta)$ over the kth piecewise constant azimuth interval can then be determined from the set of registered training images representing azimuth angles close to the interval center, $\phi_k = 2\pi k/N_w$. That is,

$$\hat{\zeta}_l(a, \theta_k) = \frac{1}{|\mathcal{I}_k|} \sum_{\mathbf{r} \in \mathcal{I}_k} |r_i|^2, \tag{6}$$

where $\mathcal{I}_k = \{\mathbf{T}_j \in \mathcal{T} : \phi_j \in [\phi_k - d/2, \phi_k + d/2]\}$ for some training interval width d. If $d > 2\pi/N_w$ then some training images may contribute to the estimate of $\hat{\zeta}_l^2(a, \phi_k)$ for more than one interval k.

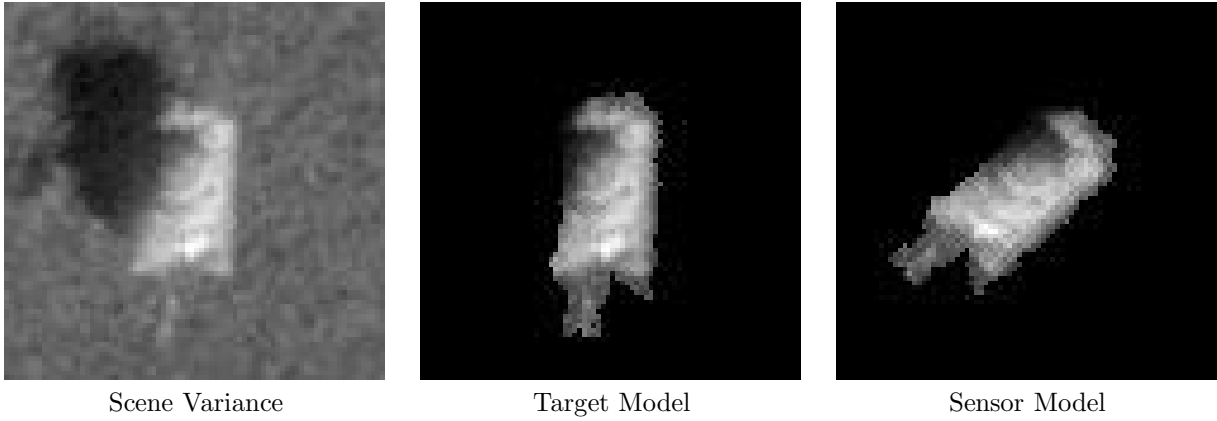| Scene Variance | Target Model | Sensor Model |

Figure 1. Estimated scene variance, segmented target model, and resulting sensor output model for target T-72 with a 45° azimuth angle.

Because the variances $\hat{\zeta}_l(a, \theta)$ are defined over the target surface, they are useful for determining the segmentation $\mathbf{S}(a, \theta)$ of Section 2. That is, we can consider the segmentation of a target model rather than segmentations of individual SAR images. An information-based segmentation can be realized by determining the average relative entropy between distributions parameterized by the estimated variances and the null-hypothesis distribution of the previous section. That is, $S_l(a, \theta) = 1$ and the region l in the scene is determined to be on the target surface if

$$\frac{1}{N_w} \sum_k D(p(\cdot|\sigma^2 = \hat{\zeta}_l^2(a, \theta_k)\|p(\cdot|\sigma^2 = \xi^2)) > \eta. \tag{7}$$

Otherwise, $S_l(a, \theta) = 0$.

The variance in pixel location i in a SAR image can be determined from the target model of (6) for any pose as the variance of the corresponding scene region l and the collection of these pixel variances constitutes a sensor model of the target. This is found through the inverse of the transformation U and the variance image is

$$\hat{\boldsymbol{\sigma}}^2(a, \theta, \mathbf{s}) = U(\theta, \mathbf{s})\hat{\boldsymbol{\zeta}}(a, \theta_k), \tag{8}$$

where θ_k represents the azimuth interval containing azimuth θ. Note that while the functions $\hat{\zeta}_l$ are approximated as piecewise constant in θ, the sensor model in (8) can vary over all θ. Similarly, the indicator function $\mathbf{I}$ for pixels in an image is related to the segmentation $\mathbf{S}$ through the same transformation. That is,

$$\mathbf{I}(a, \theta, \mathbf{s}) = U(\theta, \mathbf{s})\mathbf{S}(a, \theta_k). \tag{9}$$

Figure 1 shows estimated scene variances $\hat{\boldsymbol{\zeta}}$, the corresponding segmented target model with $\eta = 1$, and transformed sensor model $\hat{\boldsymbol{\sigma}}$ for the target T-72 from the MSTAR dataset for $\theta = 45°$.

4. SUCCESSIVELY REFINABLE MODELS

The masked sensor models from (8) and (9) are estimates of the only unknowns in the algorithm of (3) and (4). However, the maximization is complicated because the pose parameters $\Theta = [\theta, \mathbf{s}]$ are continuous-valued and a brute force search of the parameter space is difficult. We therefore discretize the parameter space resulting in a search over a finite set of parameters. A natural discretization of target location is to restrict consideration of $\mathbf{s}$ to shifts corresponding to integral numbers of pixels in the SAR image. Assuming the target region of interest has been well localized, we can establish some maximum distance in range and cross-range directions for consideration. Discretization of azimuth can be accomplished by approximating $\hat{\boldsymbol{\sigma}}^2$ as piecewise constant in θ over some number of

$$d_1 = 360° \qquad d_2 = 180° \qquad d_3 = 90° \qquad d_4 = 45°$$

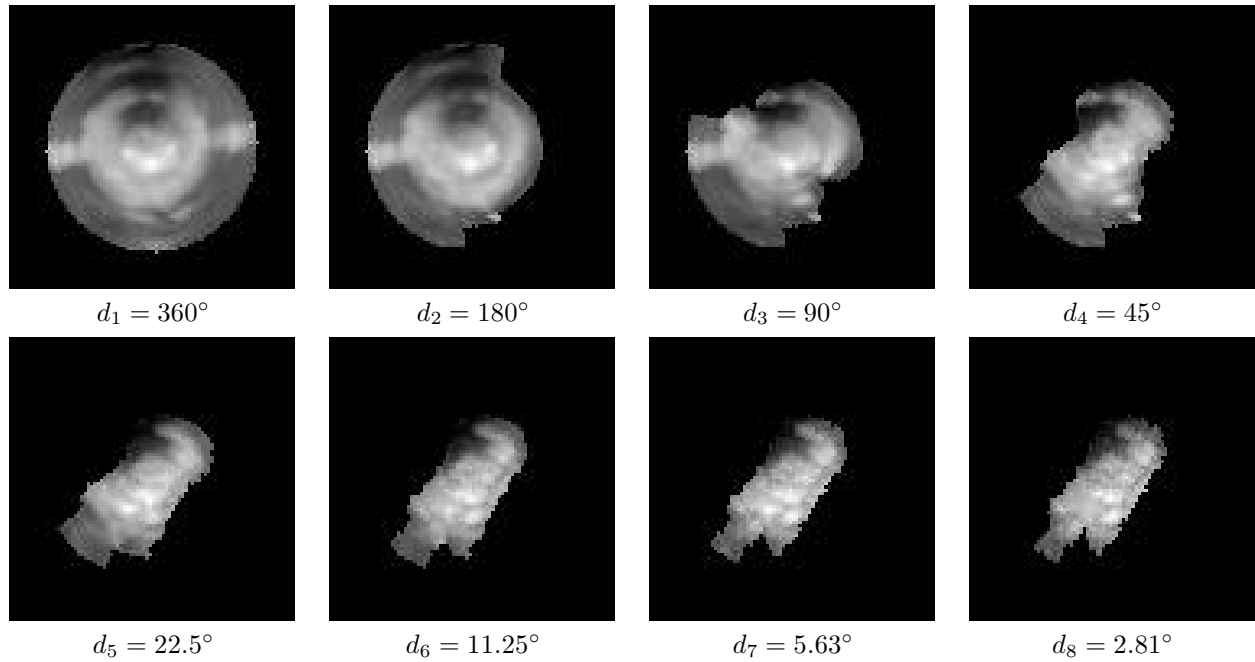$$d_5 = 22.5° \qquad d_6 = 11.25° \qquad d_7 = 5.63° \qquad d_8 = 2.81°$$

Figure 2. Sensor image models of the target T-72 at 30° azimuth under approximations d_1 through d_8.

intervals, N_d, of width d. The kth interval consists of the azimuth angles in $W_k = [2\pi k/N_d - d/2, 2\pi k/N_d + d/2]$. The approximate variance of the ith pixel over the kth interval is the average

$$\tilde{\sigma}^2_{d,i}(a, \theta_k, \mathbf{s}) = \frac{\int_{\{\theta \in W_k : I_i(a,\theta,\mathbf{s})=1\}} \hat{\sigma}^2_i(a, \theta, \mathbf{s}) d\theta}{\int_{\theta \in W_k} I_i(a, \theta, \mathbf{s}) d\theta}. \tag{10}$$

This integral can itself be approximated by a summation, choosing Δ_θ to be a small angle that evenly divides $d/2$

$$\tilde{\sigma}^2_{d,i}(a, \theta_k, \mathbf{s}) \approx \frac{1}{n} \sum_{j=-d/2\Delta_\theta}^{d/2\Delta_\theta} \sigma^2_i(a, 2\pi k/N_d + j\Delta_\theta, \mathbf{s}), \tag{11}$$

where $n = \sum_j I_i(a, 2\pi k/N_d + j\Delta_\theta, \mathbf{s})$ is the number of pixels on target contributing to the average.

The approximation interval width d is a parameter that will directly influence the performance of the resulting ATR system. It might reasonably be expected that smaller values of d will result in higher recognition accuracy. Smaller values of d will also result in more computational effort due to the larger search space for the maximization in (3). Motivated by multi-resolution representations of images, such as those described by Said and Pearlman[8] and others, we consider a sequence of approximations in θ characterized by decreasing interval width as $d_1 = 2\pi, d_2 = \pi, \ldots, d_m = 2\pi/2^{m-1}$. From (11) the variances over neighboring intervals are related through successive approximations by

$$\tilde{\sigma}^2_{d_m,i}(a, \theta_k, \mathbf{s}) = \frac{1}{n_1 + n_2} \left[n_1 \tilde{\sigma}^2_{d_{m+1},i}(a, \theta_{2k}, \mathbf{s}) + n_2 \tilde{\sigma}^2_{d_{m+1},i}(a, \theta_{2k+1}, \mathbf{s}) \right], \tag{12}$$

where n_1 and n_2 are the number of pixels on target contributing to the averages in (11) for each neighboring interval. The number of terms contributing to the approximation $\tilde{\sigma}^2_{d_m,(i)}(a, \theta_k, \mathbf{s})$ is $n_1 + n_2$. Figure 2 shows variance images of the target T-72 at 30° azimuth for approximations d_1 through d_8.

This set of approximations induces a resolution tree representation of the sensor model with a single variance value for all 360° of azimuth at the root of the tree, representing approximation d_1. The next level of the tree,

representing approximation d_2, contains a pair of values equal to the average variance over two nonoverlapping $180°$ intervals and satisfying the relationship in (12). Level l in the tree represents approximation d_l and consists of 2^{l-1} variances such that the $2k$ and $2k+1$ values also satisfy (12). There is one such resolution tree for each target represented by the ATR system. The maximization in (3) can be viewed as a search through all such trees for the combination of azimuth interval, target class, and target position that maximizes the likelihood at some level m in the trees.

This search can be organized in several ways. A straightforward approach is to conduct a brute force search for all targets through the leaves of the tree at level m, considering all positions $\mathbf{s}$ for each leaf. If the evaluations of likelihoods in (3) are performed target-by-target for each leaf in the tree, the ATR system will be able to report a target classification that evolves over time as the most likely target class found up to any given point. Other possible approaches can exploit the tree structure, starting with approximation d_1 and considering successively more refined approximations. One such strategy is to explore subtrees, evaluating the likelihood for each target class and each target position for the single set of variances at d_1 and for the two sets at d_2. From there, the search is conducted over only the two intervals at d_l that are related through (12) to the most likely azimuth interval considered at d_{l-1}. When the final level m is reached, likelihoods are computed for each interval according to their distance from the most likely interval at level d_{m-1}. A third approach is to conduct a breadth-first search over the resolution tree, computing likelihoods for the intervals at level d_l in an order determined by the most likely intervals from level d_{l-1}.

By interleaving the computations for each target with computations for the nodes of the resolution tree, the most likely target found up to any point in time can be reported. In this way, the accuracy of an ATR system becomes a function of the quantity of system resources consumed. In this context, the term system resources may include processor cycles, memory storage, elapsed time, etc. For the examples of the next section, we will consider accuracy as a function of the average number of bits communicated between the database which stores the model parameters $\tilde{\boldsymbol{\sigma}}^2_{d_m}$ and the processing units. In related work, DeVore, *et al.*[2] explore the relationship between accuracy and the throughput of ATR systems. The tradeoff between accuracy and database complexity has been used[1,4] to characterize and compare alternate approaches to ATR from SAR imagery.

The number of bits communicated between processor and database for any given target class is the total number of bits required to describe the variance approximations $\tilde{\boldsymbol{\sigma}}^2_{d_l}(a, \theta_k)$ used in computations plus the number of bits used to specify the order in which the variance approximations are processed, if any. Since changes in target location $\mathbf{s}$ can be accommodated by simple shifts of $\tilde{\boldsymbol{\sigma}}^2_{d_l}$, the number of bits is independent of the number of target locations considered. For the direct leaf-search through level m, no processing order must be specified and the number of bits to communicate a single approximation is lower bounded by the sum of the number of bits to specify a bounding box enclosing N pixels, N bits to represent the indicator function $\mathbf{I}(a, \theta)$, and b_f bits for every nonzero element of $\mathbf{I}(a, \theta)$ to describe the corresponding variance at that pixel location, where b_f bits are used in the representation of floating-point values. An alternative encoding method is to specify, for each nonzero element of $\mathbf{I}(a, \theta)$, an index into an array of size N and the corresponding value of $\tilde{\boldsymbol{\sigma}}^2_{d_l}(a, \theta_k)$. For the experiments in the next section, we chose the smaller of these representations for each mask function.

Fewer bits may be expected with more sophisticated coding schemes, but may also require more effort from the processing units for decoding. The subtree search method requires approximately half that amount for levels 2 through m, because the relationship in (12) can be exploited to determine half the approximations $\tilde{\boldsymbol{\sigma}}^2$ but does require an additional number of bits to communicate the integer n_1 for half the intervals plus a single bit for each level 2 through $m-1$ to communicate which interval yielded a larger likelihood. The breadth-first search has similar communication requirements except that enough bits to communicate a complete ordering of the likelihoods at each level must be transmitted to the database. For the experiments of the next section, we assume all models fit within a 128×128 pixel region so 14 bits are sufficient to describe a bounding box, and we assume the use of floating-point values for pixel variances that are 64 bits long.

5. RESULTS

The algorithms for classification and confuser rejection of Section 2 were applied to publicly available SAR data from the MSTAR program. The data were segregated into nonoverlapping subsets for estimating target models and for assessing recognition performance. Table 1 characterizes the training and testing data corresponding to the four target models that were estimated. Images taken from a $17°$ depression angle were used for training and those from a $15°$ depression angle were used for testing. Two of the target classes, BMP-2 and T-72, had target models estimated

Class	Vehicle	Serial No.	Training Set		Testing Set	
			Depression	Images	Depression	Images
BMP-2	#1	9563	17°	233	15°	195
	#2	9566		231		196
	#3	c21		233		196
BRDM-2	#1	E-71	17°	298	15°	263
BTR-70	#1	c71	17°	233	15°	196
T-72	#1	132	17°	232	15°	196
	#2	812		231		195
	#3	s7		228		191

Table 1. Images from the MSTAR dataset used for estimation of target models and testing images used to assess algorithm performance.

Class	Vehicle	Serial No.	Depression	Images
2S1	#1	b01	15°	274
BTR-60	#1	k10yt7532	15°	195
D7	#1	92v13015	15°	274
T-62	#1	A51	15°	273

Table 2. Images from the MSTAR dataset used for evaluating confuser rejection capabilities.

from images of multiple vehicles. Table 2 describes confuser images, images of targets without corresponding target models. These confuser images were used in tests of the confuser rejection algorithm (5).

The left panel of Figure 3 shows the percentage of incorrectly classified test images from Table 1 as a function of the average number of bits per target model transmitted between the target model database and processing unit for each test image. These results include approximations d_1 through d_9. The three search strategies described in Section 4 were employed and configured to report the largest likelihood value computed at each step in the computation of (3). The leaf search explores the likelihoods under approximation d_9 in an essentially arbitrary order and the classification error rate drops from a value of 57% when only a single orientation is considered to just under 2% after all 256 azimuth intervals have been processed. The decline in error rate is approximately linear over this range. The error rate under the subtree search declines sharply from 40% to around 13% but then levels out until near the end of the search. This suggests that the wrong branch of the resolution tree is often followed early in the search. The breadth-first search yields an error rate that drops sharply to near its minimum value over the first one-sixth of the plot. This suggests that in most cases relatively coarse approximations are sufficient for classification. Each of the three search strategies eventually evaluate all likelihoods at approximation d_9 and so all have the same final error rate.

The right panel of Figure 3 shows the percentage of correctly classified test images from Table 1 as a function of the percentage of confuser images from Table 2 that were not rejected. The curve is based on classification results through approximation d_9 and is parameterized by the threshold γ of (5). The classification of a test image is considered to be in error if it is either classified as the wrong target or if it is incorrectly rejected as a confuser target. For this reason the curve reaches a maximum of just over 98% even when no confuser images are rejected. The graph shows that a rejection threshold γ can be chosen to yield a 91% correct recognition rate while passing 10% of confuser images for the dataset considered.

6. CONCLUSIONS

We have demonstrated an approach to building models that account for the variation in pose of targets from sample observations. The models can be used to predict statistical characterizations of images of those objects for application to ATR. Target segmentation and the rejection of confuser vehicles is performed using empirical measures of relative information and recognition is performed via maximization of likelihood ratios. The models admit a hierarchical
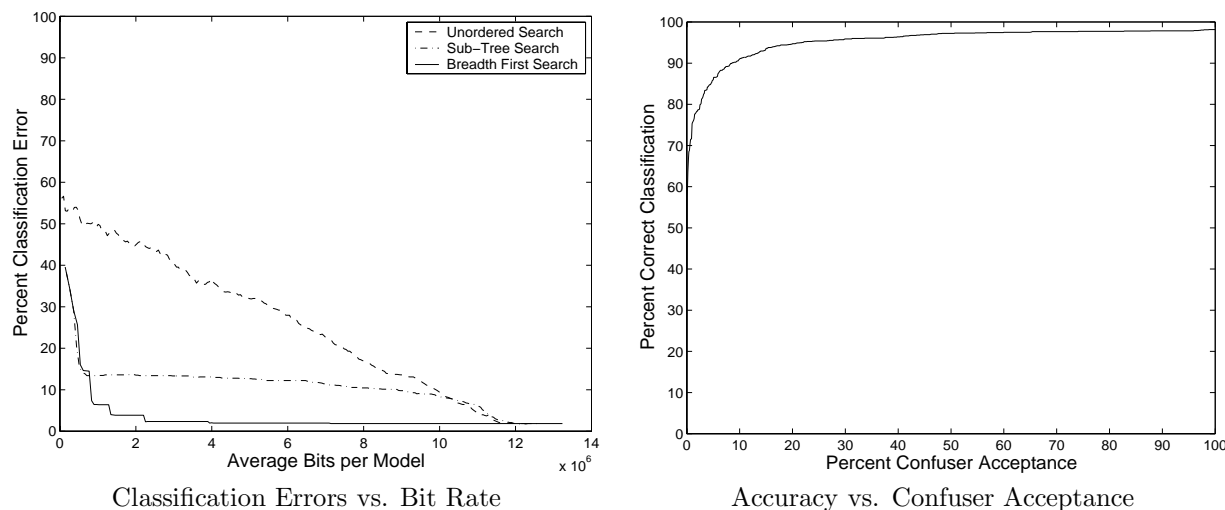
Figure 3. Left panel: Percentage of incorrectly classified images as a function of bits per target class transmitted from the model database. Right panel: Percentage of correctly classified images as a function of the percentage of confuser images not rejected.

representation that is useful in the search for combinations of target class and pose that maximize the likelihood of an observation, allowing the most likely regions of the search space to be considered first. Such a representation allows an initial target classification which can then be refined as the search space is examined through successively finer grained approximations.

These methods were demonstrated on a four-class problem with four confuser targets with images from the MSTAR dataset. The results show that a breadth-first search through the resolution tree of model approximations resulted in a classification error rate that quickly approached its minimum value of under 2% as the number of bits in the approximations was increased. Rejection of confuser targets is based on a threshold parameter which could be set to reject 90% of the images in the confuser set while correctly classifying 91% of the test images.

ACKNOWLEDGMENTS

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. D. DeVore, J. A. O'Sullivan, R. D. Chamberlain, and M. A. Franklin. Relationships between computational system performance and recognition system performance. In F. A. Sadjadi, editor, *Automatic Target Recognition XI, Proc. of SPIE*, 2001.

3. M. D. DeVore, N. A. Schmid, and J. A. O'Sullivan. Analytical and experimental performance-complexity tradeoffs in ATR. In *Proceedings of the Thirty-Fourth Asilomar Conference on Signals, Systems, and Computers*, Oct. 2000.

4. J. A. O'Sullivan and M. D. DeVore. Performance-complexity tradeoffs for several approaches to ATR from SAR images. In E. G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery VII, Proc. of SPIE*, volume 4053, 2000.

5. 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*, 37(1):91–108, Jan. 2001.

6. G. J. Owirka, S. M. Verbout, and L. M. Novak. Template-based SAR ATR performance using different image enhancement techniques. In *Algorithms for Synthetic Aperture Radar Imagery VI, Proc. of SPIE*, volume 3721, pages 302–319, 1999.

7. T. D. Ross, S. W. Worrell, J. C. Mossing, and M. L. Bryant. Standard SAR ATR evaluation experiments using the MSTAR public release data set. In E. G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery V, Proc. of SPIE*, volume 3370, pages 566–573, 1998.

8. A. Said and W. A. Pearlman. An image multiresolution representation for lossless and lossy compression. *IEEE Transactions on Image Processing*, 5(9):1303–1310, Sept. 1996.

9. N. A. Schmid and J. A. O'Sullivan. Thresholding method for reduction of dimensionality. *IEEE Transactions on Information Theory*. Accepted for Publication.

10. J. Wissinger, R. B. Washburn, N. S. Friedland, A. Nowicki, D. R. Morgan, C. Chong, and R. Fung. Search algorithms for model based SAR ATR. In E. G. Zelnio and R. J. Douglass, editors, *Algorithms for Synthetic Aperture Radar Imagery III, Proc. of SPIE*, volume 2757, pages 279–293, 1996.