Machine Vision and Applications (2000) 11: 289–299

Machine Vision and Applications

Adaptive target recognition

Bir Bhanu, Yingqiang Lin, Grinnell Jones, Jing Peng

Center for Research in Intelligent Systems, University of California, Riverside, CA 92521, USA; e-mail: {bhanu,yqlin,grinnell,jp}@cris.ucr.edu

Abstract. Target recognition is a multilevel process requiring a sequence of algorithms at low, intermediate and high levels. Generally, such systems are open loop with no feedback between levels and assuring their performance at the given probability of correct identification (PCI) and probability of false alarm (Pf) is a key challenge in computer vision and pattern recognition research. In this paper, a robust closed-loop system for recognition of SAR images based on reinforcement learning is presented. The parameters in model-based SAR target recognition are learned. The method meets performance specifications by using PCI and Pf as feedback for the learning system. It has been experimentally validated by learning the parameters of the recognition system for SAR imagery, successfully recognizing articulated targets, targets of different configuration and targets at different depression angles.

Key words: Target recognition – Reinforcement learning – Parameter learning

1 Introduction

Typical systems for model-based object recognition use a sequence of algorithms that operate at various stages of the process to perform recognition. These systems are usually unidirectional and without feedback. Often there are procedural (tuning) parameters associated with these algorithms that must be selected to achieve the desired performance, e.g., a given point on a ROC (receiver operating characteristic) curve demonstrating the performance of the entire system. However, automatic development of control strategies in an object recognition system has been a challenging problem in the field of computer vision and pattern recognition. In addition, the simultaneous adjustment of even a few system parameters is time-consuming and difficult and has yet to be solved satisfactorily for multistage systems. This is due to the lack of a feedback theory of multistage object recognition.

In this paper, we propose a learning-based target recognition system for SAR imagery that is capable of automatically adjusting its procedural parameters (input to target detection, discrimination and recognition algorithms), thereby achieving a desired performance specification, a point on the 'limit' ROC curve. A typical ROC curve (plot of probability of detection Pd, or in our case, probability of correct identification PCI, versus probability of false alarm Pf) is generated by varying one parameter of the system (in our case, typically, the decision rule parameter). Varying other parameters (such as the number of features used) generates families of different ROC curves. Usually, no one set of parameters provides a single ROC curve that is optimum, with the greatest PCI, for all values of Pf. Thus, we introduce the concept of a 'limit' ROC curve (an example is shown in Fig. 1), which is the highest ROC curve that can be produced (for a given set of target and 'confuser' data) by optimally varying all of the system parameters.

Our approach to the control strategy is to use an incremental method based on reinforcement learning for computing parameter settings in a multistage model-based target recognition system. This paper begins the development and evaluation of the proposed computational model by building and controlling a system that is capable of being operated on any given point on the limit ROC curve. The application of the system is the recognition of articulated targets, non-standard targets and targets at slightly different depression angles in SAR imagery. Although the system is now a single-stage system, it serves as the basis for an eventual multistage system that would also control parameters for target detection and target discrimination.

We use reinforcement learning because it is difficult, if not impossible, for a conventional search method to accomplish the same task. Simply, there are no well-defined evaluation functions at each of the stages of object recognition. Furthermore, if a method uses a point on the limit ROC curve for evaluation, then it is not clear how the process should proceed in a systematic way. Finally, at each stage, any such method will have to delay its decision as to where to search next until the recognition results become available. However, this need not be the case for the approach presented in this paper.

Correspondence to: B. Bhanu

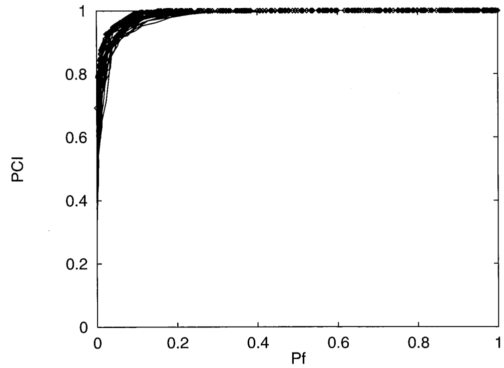


Fig. 1. Limit curve (*top thick curve*) of five ROC curves of articulated MSTAR data. Each ROC curve is for a different number of scattering centers (26, 32, 37, 43, 47). The decision threshold is varied along each ROC curve

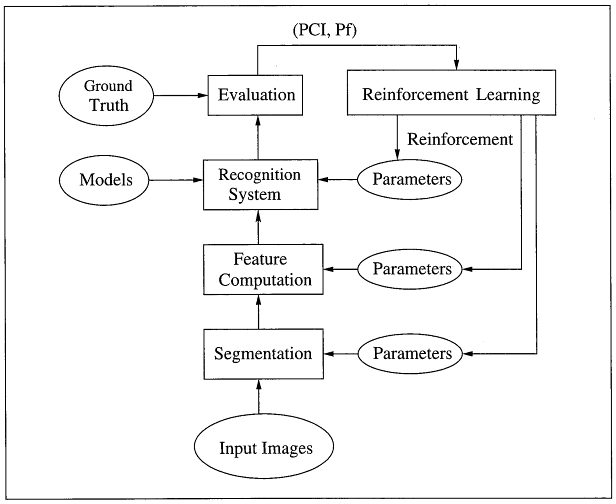


Fig. 2. Learning integrated SAR recognition system

2 Related work and our contributions

There is no published work in image processing, computer vision and pattern recognition on learning to perform recognition at a given point on the ROC curve. Procedural parameter learning for the recognition algorithm and the adaptive selection and combination of different algorithms in a learning integrated system are *unsolved* problems in the field of image processing, computer vision and pattern recognition [1]. Most techniques in computer vision and pattern recognition do not involve any learning to improve future performance with experience.

Burges et al. [4] describe a method for coupling recognition and segmentation by the principle of heuristic over-segmentation. The basic idea is that a segmentation algorithm generates a graph that summarizes a large number of segmentation hypotheses that are scored by a recognition algorithm. A globally optimal decision is then made that combines uncertainties in segmentation and recognition. Each time a new input comes in, an over-segmented hypothesis graph must be generated and traversed in order to classify the input. In another related work by Le Cun et al. [9], graph generation is actually learned by minimizing global errors that take into account both segmentation hypotheses and recognition scores. Bengio and Le Cun [2] describe a method for fitting segmentation parameters to maximize the likelihood of a model of an object. Peng and Bhanu [10, 11] develop approaches in which reinforcement learning techniques are used to close the loop between segmentation and recognition. Draper et al. [5] use reinforcement learning to select a control policy for recognizing a rectangular rooftop in aerial images.

The original contributions of the adaptive target recognition approach presented in this paper are that performance specification (given as a point on the ROC curve) can be used as feedback to influence the target recognition system, thereby providing a target recognition system with adaptability in real-world applications. Experimental results are presented using XPATCH and MSTAR data for adapting system parameters for different articulations, different configurations and slightly different depression angles.

3 Overview of the approach

In the general multilevel recognition process, the parameters used in the image segmentation, feature extraction and recognition modules can be adaptively controlled with feedback, based on reinforcement learning, as shown in Fig. 2. The results of recognition (PCI, Pf) are evaluated and a feedback signal based on the discrepancy between the desired and actual performance is generated. This signal is used as the input to reinforcement learning which provides appropriate feedback in selecting parameters that are used in the segmentation, feature extraction and recognition modules. Segmentation, feature extraction and recognition are carried out with new parameters and the process continues until the desired performance is obtained or the number of allowed loops has been exceeded.

3.1 Image segmentation

The image segmentation process extracts small areas of the image that are likely to contain targets. The goal of this stage is to eliminate most of the sensor data from further consideration and find small regions potentially containing the targets of interest. Detecting targets in SAR images typically involves a pre-screen stage (e.g., a constant false-alarm-rate thresholding technique [13]) and a discriminator stage to separate the targets from the background clutter.

The experimental results in this paper are based on the XPATCH and MSTAR (Public) data set, so we do not currently implement and control the initial image segmentation to provide the SAR target chips. However, the discrimination of target features from the background in the target chips is required and implemented as described in the next section.

3.2 Feature computation

The locations and magnitudes of the return from SAR scattering centers are characteristic features that are related to the geometry of the object. The typical detailed-edge and straight-line features of man-made objects in the visual world, do not have good counterparts in SAR images for sub-components of vehicle-sized objects at 1-foot resolution. A typical SAR target shows a wealth of peaks corresponding to scattering centers and has no obvious lines or edges within the boundary of the vehicle.

The MSTAR (Public) data set, used to provide SAR target chips for these experiments, has targets and background clutter from a grassland environment. Target regions of interest (ROI) are found in the MSTAR SAR target chips by reducing speckle noise using the Crimmins algorithm in Khoros [7], thresholding at the mean plus two standard deviations, dilating to fill small gaps among regions, eroding to have one large ROI and little regions, discarding the small regions with a size filter and dilating to expand the extracted ROI. The scattering centers are extracted from the SAR magnitude data (within the boundary contour of the ROI) by finding local eight-neighbor maxima.

In the current implementation, the parameters used to extract the ROIs (e.g., signal and region size thresholds) are fixed, while the number of scattering centers used, N, is a variable parameter of the recognition algorithm described in the next section.

3.3 SAR recognition algorithm

The basic SAR recognition algorithm is an off-line model construction process and a similar on-line recognition process (the detailed algorithms are given in [6]). The approach is designed for SAR and is specifically intended to accommodate recognition of occluded and articulated objects. Standard non-articulated models of the objects are used to recognize these same objects in non-standard, articulated and occluded configurations. The models are a look-up table and the recognition process is an efficient search for *positive evidence*, using relative locations of the scattering centers in the test image to access the look-up table and generate votes for the appropriate object (and azimuth pose).

The relative locations and magnitudes of the N strongest SAR scattering centers (local maxima in the radar return signal) are used as characteristic features (where N, the number of scattering centers used, is a design parameter). Any local reference point, such as a scattering center location, can be chosen as a 'basis point' to establish a reference coordinate system. The relative distance and direction of other scattering centers can be expressed in radar range and cross-range coordinates and naturally tessellated into integer buckets that correspond to the radar range/cross-range bins. For ideal data, picking the location of the strongest scattering center as the basis point is sufficient. However, for potentially corrupted data where any scattering center could be spurious or missing (due to the effects of noise, target articulation, occlusion, non-standard target configurations, etc.), we use all N strongest scattering centers in turn as basis points to ensure that a valid basis point is obtained. Using a technique like geometric hashing [8], the models are constructed using the relative positions of the scattering centers in the range and cross-range directions as the *initial* indices to a look-up table of labels that give the associated target type, target pose, basis point range and cross-range positions and the magnitudes of the two scatterers. Because of the specular radar reflections of SAR images, a significant number of features do not typically persist over a few degrees of rotation. Consequently, we use 360 models (at 1° azimuth increments) for each object.

The recognition process uses the relative locations of the N strongest scattering centers in the test image to access the look-up table and generate votes for the appropriate object, azimuth, range and cross-range translation. Constraints are applied to limit the allowable percent difference in the magnitudes of the data and model scattering centers. (Limits on allowable translations are also imposed for computational efficiency.) To accommodate some uncertainty in the scattering center locations, the eight-neighbors of the nominal range and cross-range relative location are also probed and the translation results are accumulated for a 3×3 neighborhood in the translation sub-space. The recognition process is repeated with different scattering centers as basis points, providing multiple 'looks' at the model database to handle spurious scatterers that arise due to articulation, occlusion or configuration differences. The basic decision rule used in the recognition is to select the object-azimuth pair (and associated "best" translation) with the highest accumulated vote total. To handle identification with 'unknown' objects, we introduce a criterion for the quality of the recognition result that the votes for the potential winning object exceed some minimum threshold T. By varying the decision rule threshold, we obtain a form of ROC curve with PCI vs. Pf. We call the algorithm a 6D recognition algorithm since, in effect, we use the range and cross-range positions and the magnitudes of pairs of scattering centers.

There are several parameters in the recognition procedure that could be optimized: the number of scatterers used, N (or alternatively use all scatterers above a certain signal strength); the constraint on the allowable difference in magnitude between the data and the model (we use a fixed value of $\pm$ 10%); the limits on allowable translations (fixed at $\pm$ 5 pixels in both range and cross-range directions) and the decision rule. Several possibilities exist for the decision rule: a vote threshold, T; a specific vote threshold for each object, T_i; a vote ratio threshold, VR (i.e., the ratio of votes for the potential winning object to the votes for the next best different object exceeds some value VR). Based on experiments, we found that the vote ratio approach did not work for distinguishing the BRDM 'confuser' vehicle from the T72 tank when the alternative was the ZSU 23/4, so we used the vote threshold, T, decision rule for the MSTAR data (for XPATCH data, we used vote ratio).

3.4 Multistage learning algorithm

Reinforcement learning provides a framework for constructing a general mapping from images to parameter settings in a multistage model-based object recognition system [10]. One set of effective methods for reinforcement learning is given by the theory of dynamic programming. Given a Markov decision problem, these methods involve first determining the "optimal action-value function," the Q function, that assigns to each state-action pair a value measuring the average total (discounted) reward obtained when a particular action is taken in the given state and the optimal policy is followed

1. Initialization: $\hat{Q}(x,\bar{p}) \leftarrow 0$ for all $x,\bar{p}$, where x is an image, a segmented image or features extracted and $\bar{p}$ is an instance of segmentation parameters $\bar{a}$, feature extraction parameters $\bar{b}$ or recognition algorithm parameters $\bar{c}$.
2. LOOP:
   For each image i in the training set do
   a) Segment image i with segmentation parameters $\bar{a} = (a_1, a_2, \cdots, a_n)$ recommended by ϵ-greedy policy; i_s is the resulting segmented image.
   b) Update: $\hat{Q}(i,\bar{a}) \leftarrow \hat{Q}(i,\bar{a}) + \alpha\{\gamma\hat{V}(i_s) - \hat{Q}(i,\bar{a})\}$ where $\hat{V}(i_s) = \max_{\bar{b}} \hat{Q}(i_s,\bar{b})$, α is the learning rate.
   c) Perform feature extraction with feature extraction parameters $\bar{b} = (b_1, b_2, \cdots, b_n)$ recommended by ϵ-greedy policy from the segmented image i_s.
   d) Update: $\hat{Q}(i_s,\bar{b}) \leftarrow \hat{Q}(i_s,\bar{b}) + \alpha\{\gamma\hat{V}(i_f) - \hat{Q}(i,\bar{a})\}$, $\hat{Q}(i,\bar{a}) \leftarrow \hat{Q}(i,\bar{a}) + \alpha(\lambda\gamma)\{\hat{V}(i_f) - \hat{V}(i_s)\}$, where $\hat{V}(i_f) = \max_{\bar{c}} \hat{Q}(i_f,\bar{c})$.
   e) Perform recognition with parameters $\bar{c} = (c_1, c_2, \cdots, c_n)$ recommended by ϵ-greedy policy from the features (i_f) extracted.
   f) Evaluate the recognition results and get the reinforcement r.
   g) Update: $\hat{Q}(i_f,\bar{c}) \leftarrow \hat{Q}(i_f,\bar{c}) + \alpha\{r - \hat{Q}(i_f,\bar{c})\}$, $\hat{Q}(i_s,\bar{b}) \leftarrow \hat{Q}(i_s,\bar{b}) + \alpha(\lambda\gamma)\{r - \hat{V}(i_f)\}$ and $\hat{Q}(i,\bar{a}) \leftarrow \hat{Q}(i,\bar{a}) + \alpha(\lambda\gamma)^2\{r - \hat{V}(i_s)\}$
3. UNTIL terminating condition

Fig. 3. Main steps of the multistage reinforcement learning algorithm for parameter adjustment

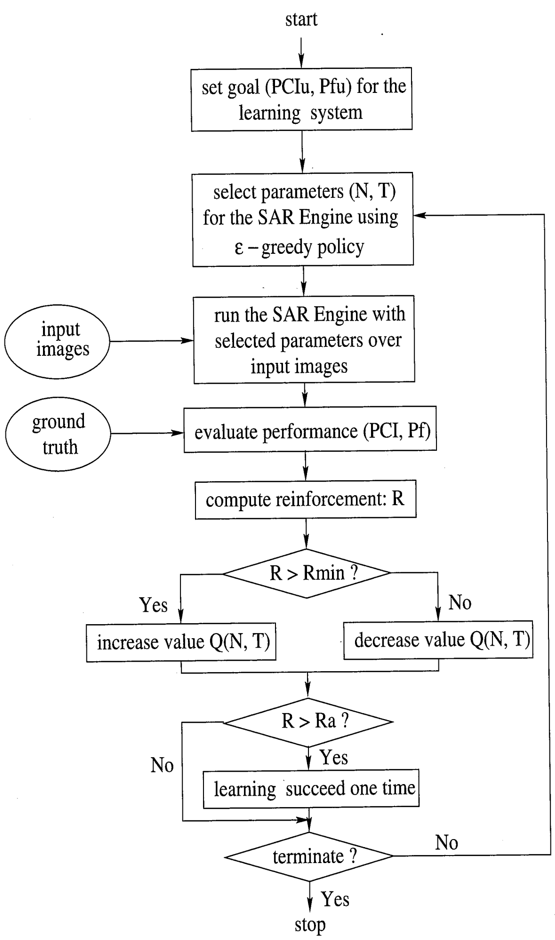


Fig. 4. Learning integrated SAR recognition system

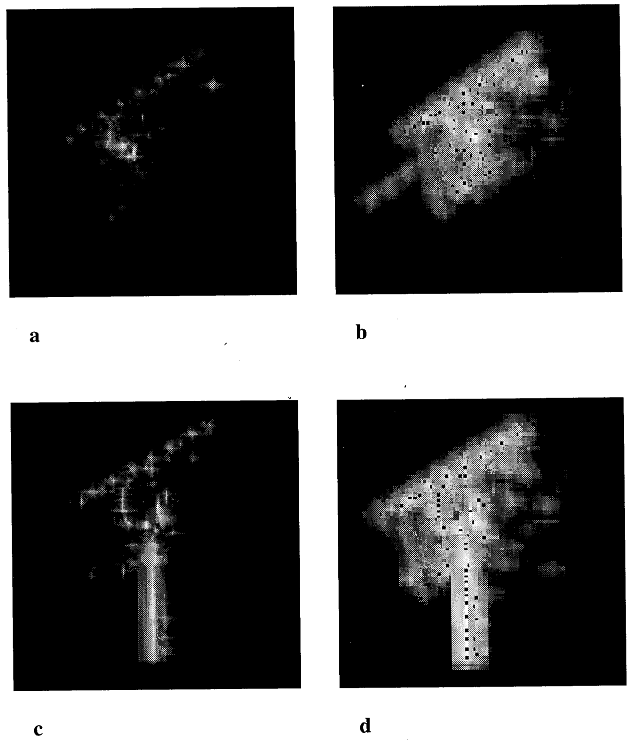


Fig. 5a–d. XPATCH T72 (30 deg hull azimuth) chips and peaks superim posed on an 8-bit image. **a** and **b**: Turret straight. **c** and **d**: Turret at 6 degrees

thereafter. That is, using the notation that x denotes the cur rent state, a the current action, r the resulting immediat reward, and y the resulting next state from taking a in x then

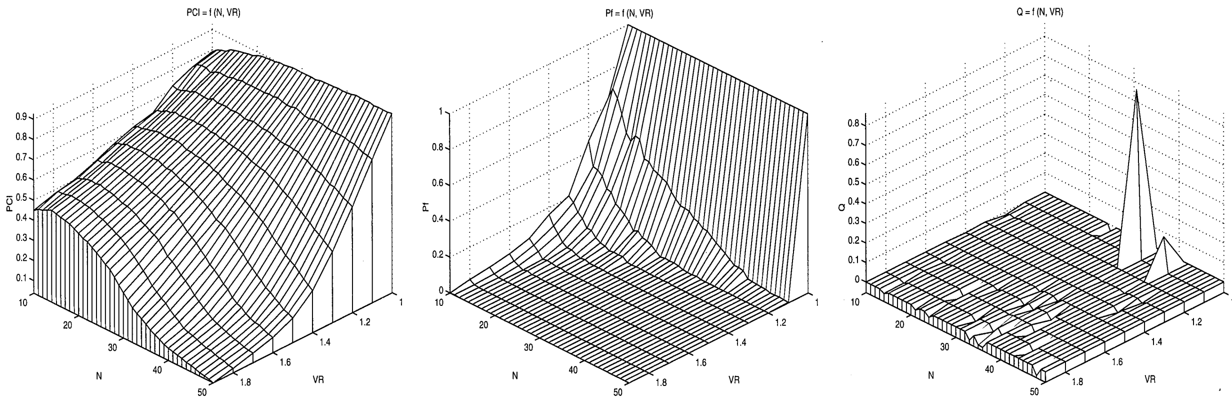
$$Q(x,a) = R(x,a) + \gamma \sum_y P_{xy}(a)V(y)\,, \qquad (1$$

where $R(x,a) = E\{r|x,a\}$ with E denoting the expectatio operator, $V(x) = \max_a Q(x,a)$, $P_{xy}(a)$ is the probability c making a state transition from x to y as a result of apply ing action a, and $\gamma \in [0,1)$ is a discount factor. Once th Q function is known, it is straightforward to determine th optimal policy. For any state x, the optimal action is simpl $\arg\max_a Q(x,a)$. Note that both x and a can be vectors.

The particular method employed in this work for learnin the Q function is the $Q(\lambda)$ algorithm, where $\lambda \in [0,1]$. Thi algorithm is an example of the "temporal difference" (TD method. Like Q learning, $Q(\lambda)$ learning works by maintai ing an estimate $\hat{Q}$ of the Q function and updating it so th Eq. 1 comes to be more nearly satisfied for each state-actio pair encountered [10]. The key steps of the multistage rei forcement learning algorithm, suited for the system show in Fig. 2, are given in Fig. 3.

3.5 Single-stage learning algorithm

Figure 4 shows the flowchart of the single-stage learnir algorithm. While our learning framework is intended for s lecting parameters in a multistage recognition system, tl



a PCI vs N, VR **b** Pf vs N, VR **c** Q vs N, VR

Fig. 6a–c. Adapting parameters for articulation with XPATCH data

Table 1. Optimum and learned results for various MSTAR Data

MSTAR articulated data								
Desired	Optimal results				Learned results			
$Pf \leq$	PCI	Pf	N	T	PCI	Pf	N	T
0.01	0.844	0.007	35	2400	0.844	0.007	35	2400
0.05	0.948	0.040	31	1700	0.948	0.040	31	1700
0.1	0.990	0.092	31	1600	0.948	0.040	31	1700
0.2	1	0.115	39	2500	0.948	0.040	31	1700
0.3	1	0.115	39	2500	0.948	0.040	31	1700
MSTAR depression angle change								
Desired	Optimal results				Learned results			
$Pf \leq$	PCI	Pf	N	T	PCI	Pf	N	T
0.01	0.621	0.009	41	5000	0.621	0.009	41	5000
0.05	0.791	0.050	47	6600	0.791	0.050	47	6600
0.1	0.876	0.096	49	7000	0.874	0.085	50	7400
0.2	0.949	0.196	41	4100	0.926	0.126	41	4300
0.3	0.978	0.284	50	6600	0.926	0.126	41	4300
MSTAR configuration change								
Desired	Optimal results				Learned results			
$Pf \leq$	PCI	Pf	N	T	PCI	Pf	N	T
0.01	0.313	0.005	34	3200	0.313	0.005	34	3200
0.05	0.542	0.031	33	2700	0.542	0.031	33	2700
0.1	0.617	0.093	36	3100	0.614	0.062	35	2900
0.2	0.718	0.196	31	2100	0.684	0.134	36	3000
0.3	0.778	0.273	31	2000	0.742	0.201	35	2700

algorithm described here is focused on learning parameter values for the SAR recognition algorithm of a single stage to achieve given PCI and Pf. It is a variant of the algorithm mentioned in the previous section. In the notation of $Q(x, a)$ used before, x is the set of all input images, which is omitted in the following equations, and $a = (N, T)$. There are two parameters to be learned: the number of features (N) to be used and the voting decision threshold (T).

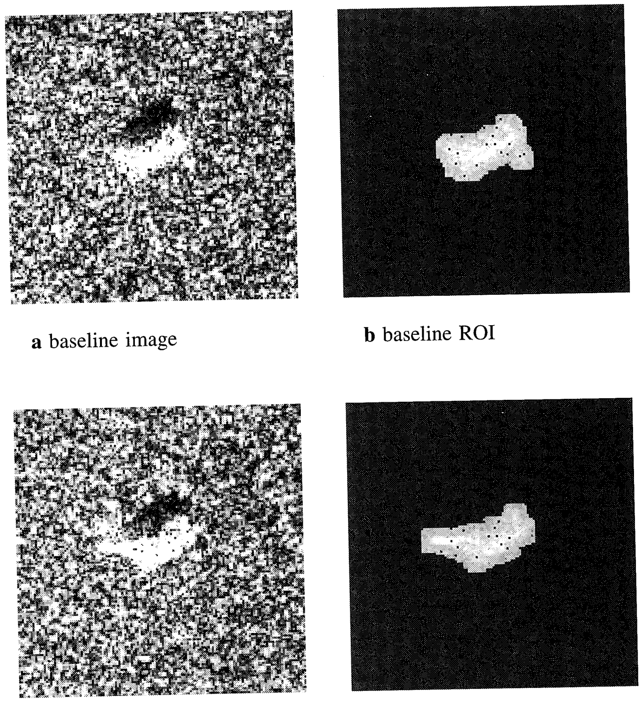
Initial conditions. When the algorithm starts to run, it asks the user to specify the performance point (PCI_u, Pf_u) on the ROC curve. Each pair of parameters (N, T) of the SAR algorithm is associated with a value $Q(N, T)$, which measures the goodness of this pair of parameters. The larger the Q value, the better the pair of parameters. All the initial Q values are 0. All pairs of parameters (N, T) are marked SELECTABLE.

Select parameters for the SAR algorithm. Only those pairs of parameters marked SELECTABLE and with non-negative Q value can be selected by the learning algorithm. If no pair of parameters meets this requirement, the learning algorithm stops automatically. The parameters are selected using an ϵ-greedy policy.

1. Select the pair of parameters (N, T) whose corresponding Q value is the maximum among all the Q values with probability $1 - \epsilon$. If several pairs of parameters have the maximum Q value, choose one of them randomly.
2. Select parameters (N, T) randomly with probability ϵ.

The value of ϵ determines the probability that the learning algorithm searches the parameter space of the SAR recognition algorithm. Parameter ϵ has an initial value ϵ_0 (we used 0.9), which means the algorithm searches the space with higher probability. Each time when the learning succeeds, the ϵ decreases by $\Delta\epsilon$ (we used 0.1). So, as the learning algorithm runs and the learning succeeds again and again, the probability of search decreases and the learning algorithm gradually focuses on good parameters. The value of ϵ should be greater than 0, since we want the learning algorithm to continue exploring the parameter space of SAR recognition algorithm, even though it explores less actively as the learning proceeds. In our experiments, the value of ϵ_{min} is 0.1.

Compute reinforcement. The learning algorithm compares the performance point (PCI, Pf) returned by the SAR recognition algorithm with the performance point (PCI_u,



a baseline image **b** baseline ROI

c articulated image **d** articulated ROI

Fig. 7a–d. MSTAR SAR images and ROIs (with peaks shown as *black dots* superimposed on the ROI) for T72 tank #a64 at 56° azimuth

Pf_u) pre-specified by the user and calculates the distance D between them. Then it computes the reinforcement R using

$$R = \left(1 - \frac{D}{\sqrt{2}}\right) \times S\,, \qquad (2)$$

where S is a distance scale factor (we used S = 1.0).

Update Q value. If the reinforcement R $\geq R_{min}$ (we used R_{min} = 0.9), the algorithm increases $Q(N,T)$ using the following formula:

$$Q(N,T) = Q(N,T) + \alpha[R - Q(N,T)]\,, \qquad (3)$$

where α was set to 0.1. Then the algorithm compares the new $Q(N,T)$ with the old $Q(N,T)$. If the difference between them is smaller than Q_Δ (we used Q_Δ = 0.001), the parameter (N,T) is marked UNSELECTABLE, and the ϵ is set to its initial value ϵ_0, which means that the learning algorithm should explore the parameter space more actively.

By mathematical induction, it is easy to prove that if the same parameter (N,T) is selected infinite times, its Q value $Q(N,T)$ converges to R in Eq. 3.

If reinforcement R $\leq R_{min}$, the algorithm decreases the $Q(N,T)$ using

$$Q(N,T) = Q(N,T) + \alpha[R - R_{min} - Q(N,T)]\,. \qquad (4)$$

Then, the pair of parameters (N,T) is marked UNSELECTABLE, which means that the pair of parameters (N,T) is not good and will not be selected any more.

Learning succeeds one time. If the learning succeeds, that is, the reinforcement received is greater than R_a (we used R_a = 0.95), the value of ϵ is decreased by $\Delta\epsilon$, which means the probability of searching is decreased.

Termination condition. If the learning succeeds more than a specified number of times, the learning algorithm stops. Otherwise, the learning algorithm tries another pair of parameters (N,T) and the loop runs again. It is worth noting that the algorithm will stop eventually (in the worst case after an exhaustive search). The searching strategy used in learning is critical to the efficiency of the learning algorithm. If the searching strategy is good, the learning algorithm can find the good parameters and converge quickly.

4 Experimental results

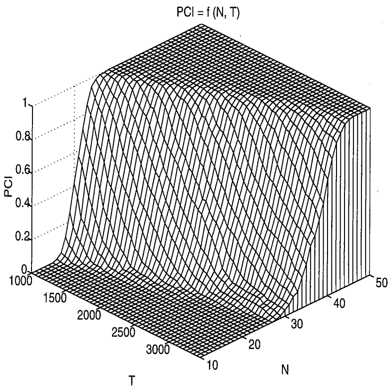
We conducted two sets of experiments: one with articulated targets using XPATCH simulated SAR data and the other with articulated targets, depression angle changes and target configuration variants using real SAR images from the MSTAR public data.

4.1 Results on XPATCH data

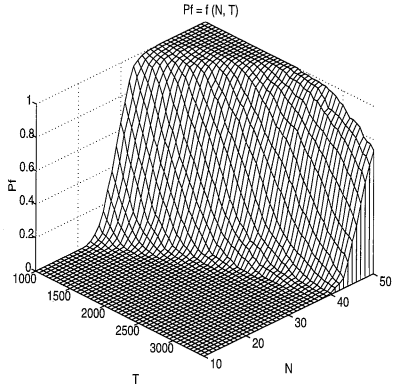
In these experiments, the input test images were articulated versions of the T72, M1A1 and T80 tanks (with the turret at 60° and 90° with respect to the hull), and the SCUD missile launcher in the erect position, where non-articulated versions of these same objects (for tanks turret at 0° and missile lancher in the down position) are used as models. The FRED tank was used as an unknown "confuser" vehicle. There are a total of 1440 model (non-articulated) images and 2880 test images. Examples of SAR target chips and scattering center locations are shown in Fig. 5 for articulated and non-articulated versions of the T72 tank.

Results with XPATCH data are based on using a 2D recognition algorithm [14] that is an earlier, simpler version of the 6D recognition algorithm described previously. The 2D algorithm uses only the relative range and cross-range distances; it does not compute the appropriate translation; it only considers the 'exact' scatterer location; and it does not use the magnitude information. For these experiments we used a vote ratio (VR) decision rule as defined in Sect. 3.3.

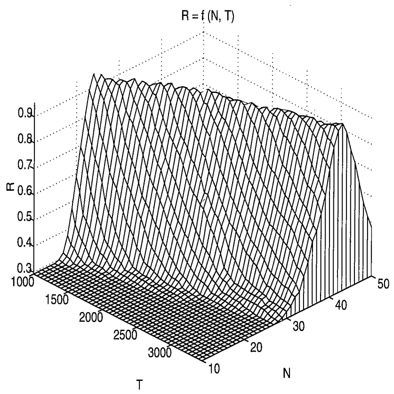
The parameters of the recognition algorithm used by the learning algorithm are number of scattering centers (N) and vote ratio (VR). Once the input images and parameters (N, VR) are given, after the recognition algorithm runs over input images with the given parameters, we get the ROC curve (PCI vs. Pf) of the recognition algorithm. Figure 6a shows the effect of N and VR on PCI, and Fig. 6b shows the effect of N and VR on Pf. The desired solutions are high PCI and low Pf value. In order to get high PCI, we should use low VR, and in order to get low Pf, we should use high VR. So, these two parameters are in conflict. The goal of the learning algorithm is to find the good parameters as fast as possible. In this experiment, the user specified 0.80 PCI and 0.10 Pf. After the learning algorithm succeeds ten times, the Q value of each pair of parameters is shown in Fig. 6c, where the best pair of parameters found by the learning algorithm is N of 35 and VR of 1.1. This pair of parameters is very close to the optimum pair of parameters (36, 1.1). The PCI and Pf resulting from (35, 1.1) are 0.80 and 0.12, respectively, very close to the PCI and Pf specified by the user.



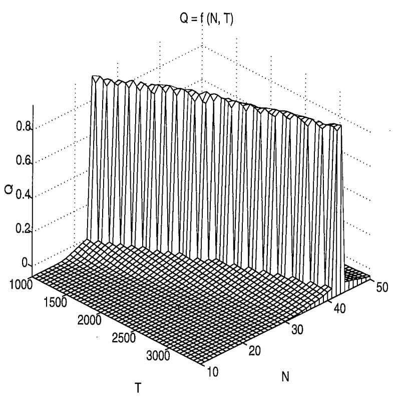
a PCI vs T, N



b Pf vs T, N

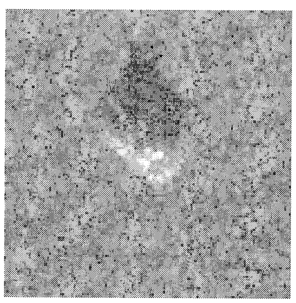


c R vs T, N

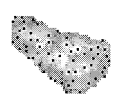


d Q vs T, N

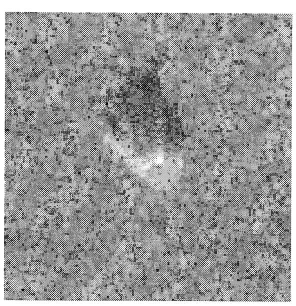
Fig. 8a–d. Adapting parameters for articulated angle changes



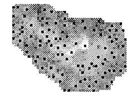
a 15° image



b 15° ROI



c 17° image



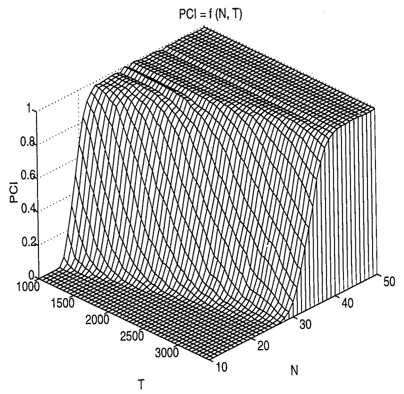
d 17° ROI

Fig. 9a–d. MSTAR SAR images and ROI (with peaks shown as *black dots* superimposed on the ROI) for T72 tank #132 at 132° azimuth and 15° and 17° depression angles

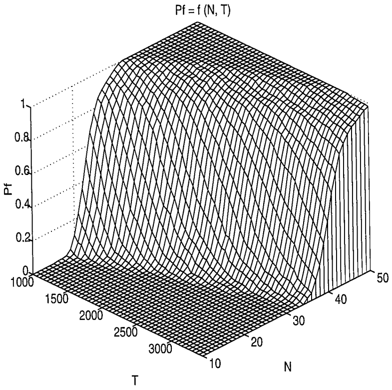
4.2 Results on MSTAR data

Articulated target results. In this experiment, the input images (test data) are articulated versions of T72(#a64) and ZSU23/4(#d08) at 30° depression angle, using the non-articulated versions of these same serial number targets as the model. BRDM2(#e71) is also included in test data, used as an "unknown" confuser vehicle. There are a total of 576 model images and 530 test images. Example SAR images and the ROI, with the locations of the scattering centers superimposed, are shown in Fig. 7 for baseline (turret straight) and articulated (turret at 315°) versions of T72 serial number (#) a64.

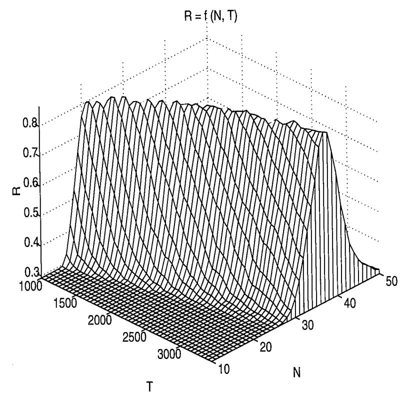
Figure 8a shows the effect of N and T on PCI, and Fig. 8b shows the effect of N and T on Pf, where N ranges from 10 to 50 and T from 1000 to 3500. We can see that as N increases, the PCI and Pf increase, and as T increases, the PCI and Pf decrease. The best results are high PCI and low Pf value. Figure 8c shows the goodness of the parameters expressed as the reinforcement R obtained from each pair of parameters when user-specified performance point is the best point (1.0, 0.0). The larger the reinforcement, the better the pair of parameters. The pair of parameters (34, 2050) are the best parameters. Figure 8d shows the Q value of each pair of parameters after the learning algorithm stops. In this experiment, the user-specified performance point is (1.0, 0.0). The Q value measures the goodness of the pair of parameters. It is consistent with the



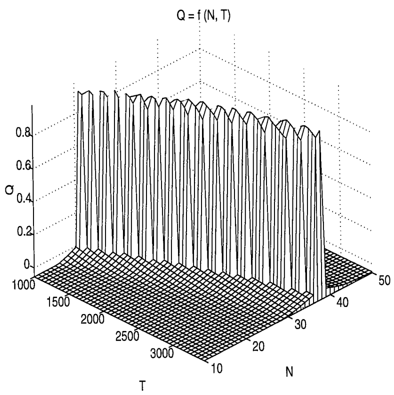
a PCI vs T, N



b Pf vs T, N

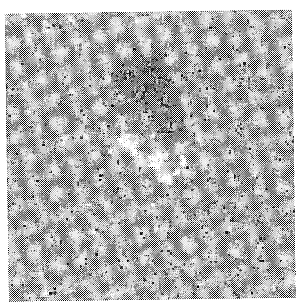


c R vs T, N

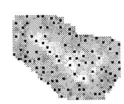


d Q vs T, N

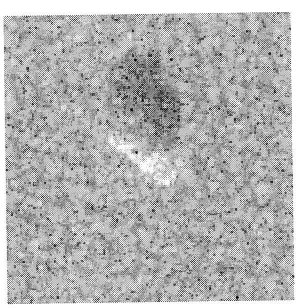
Fig. 10a–d. Adapting parameters for depression angle changes



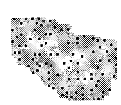
a #C21 image



b #C21 ROI



c #9563 image



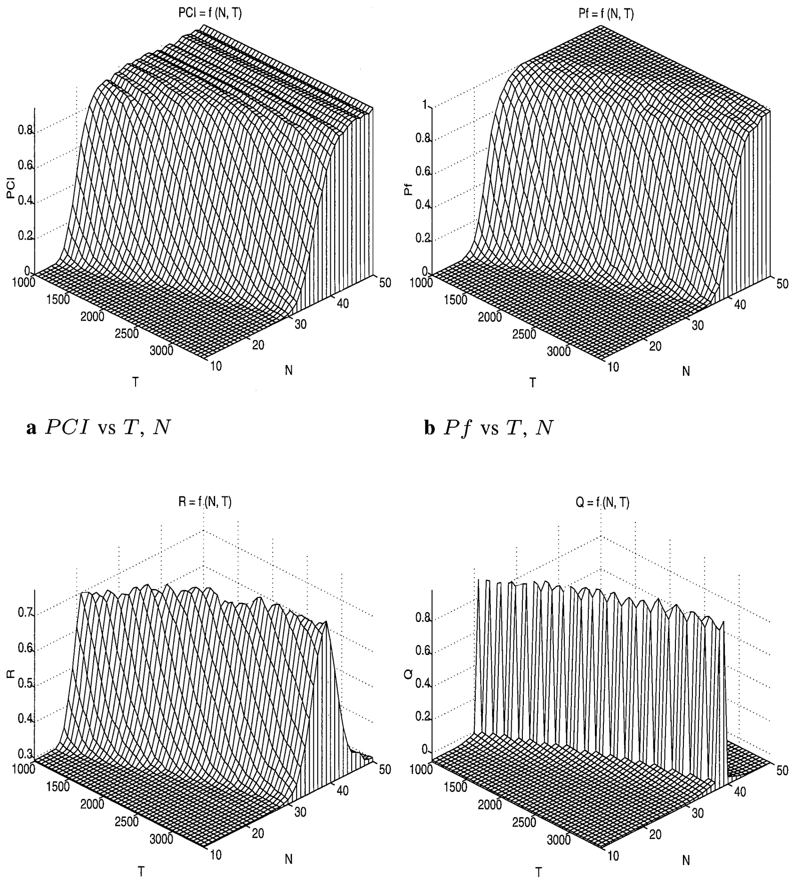
d #9563 ROI

Fig. 11a–d. MSTAR SAR images and ROIs (with peaks shown as black dots superimposed on the ROI) for BMP configuration variants #C21 and #9563 at 132° azimuth

reinforcement shown in Fig. 8c. Parameters (34,2050) have the maximum Q value. The PCI and Pf results in Fig. 8a and b can be plotted as the family of ROC curves show in Fig. 13a (For a fixed value of N, varying T maps appropriate values of PCI and Pf from Fig. 8a and b to generate one of the ROC curves shown in Fig. 13a).

Depression angle change results. The input images (test data) are T72 (#132) and BMP2 (#C21) at the depression angle of 17°, using these same serial number objects at 15° as the model. BTR70(#C71) at 17° is also included in the test data as an "unknown" confuser. There are a total of 388 model images and 697 test images. Example SAR images and the ROIs, with the locations of the scattering centers superimposed, are shown in Fig. 9. The results are shown in Fig. 10. In this experiment, the user-specified performance point is (0.89, 0.16), which is the performance point when the best pair of parameters (36,3100) in Fig. 10c is selected. Parameter (36,3100) has the maximum Q value in Fig. 10d. The PCI and Pf results in Fig. 10a and b can be plotted as the family of ROC curves show in Fig. 13b.

Configuration variant results. In this experiment, a single configuration of the vehicle (BMP2#C21 and T72#132) is used as the model and the test data are two other variants of each vehicle type (BMP #9563, #9566 and T72 #812, #s7). BTR70 (#C71) is also included in the test data, as an "unknown" confuser. Both the model and data are acquired at 15° depression angle. There are a total of 388 model images



a PCI vs T, N **b** Pf vs T, N **c** R vs T, N **d** Q vs T, N

Fig. 12a–d. Adapting parameters for configuration changes

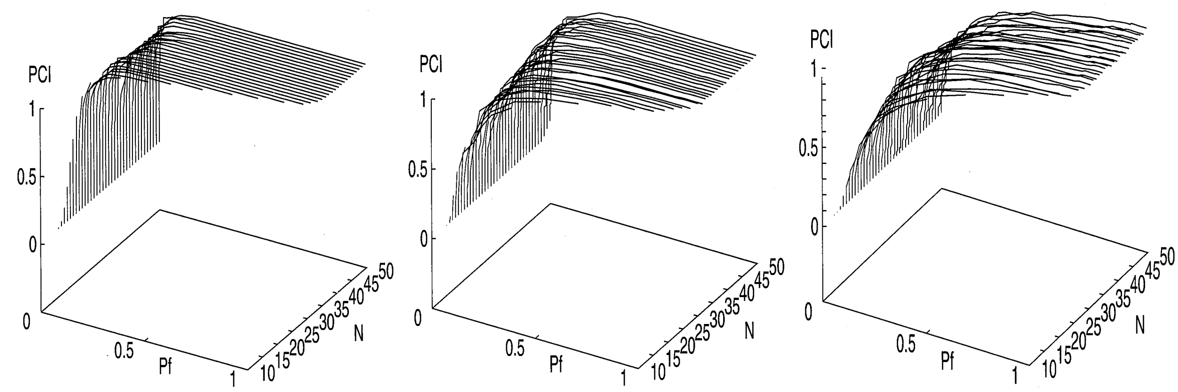
and 963 test images. Example SAR images and the ROIs, with the locations of the scattering centers superimposed, are shown in Fig. 11 for BMP #C21 and #9563 configurations at 132° azimuth. The results are shown in Fig. 12. In this experiment, the user-specified performance point is (0.72, 0.15), which is the performance point when the best pair of parameters (35,2750) in Fig. 12c is selected. Parameter (35,2750) has the maximum Q value in Fig. 12d. The PCI and Pf results in Fig. 12a and b can be plotted as the family of ROC curves show in Fig. 13c.

Analysis of results. The families of ROC curves for the MSTAR articulated data, depression angle change and configuration change data are shown in Fig. 13. These curves are superimposed to show the limit ROC curves in Fig. 14. These limit curves demonstrate the best performance the recognition system can archieve by using the optimum parameters for N and T. Table 1 shows the maximum PCI results for various specified values of Pf that can be obtained from the learning-based adaptive recognition system and the optimal results from the limit ROC curve for the MSTAR articulated data, depression angle change and configuration change data. In addition, in Table 1, the tuning parameters, N and T, that are found by the learning algorithm can be compared with the values for the optimal system.

5 Conclusion and future work

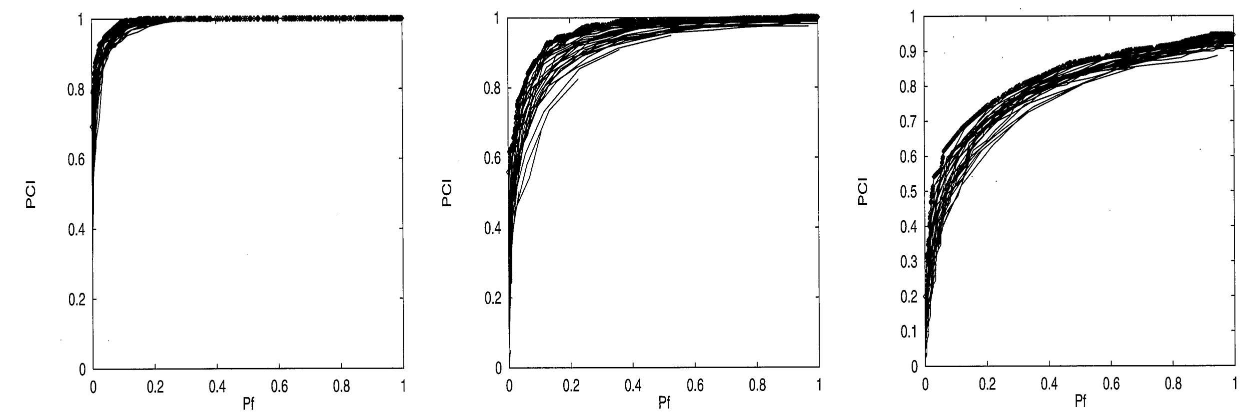
In this paper, we have demonstrated a reinforcement-learning-based control strategy to determine the optimum values for two parameters in the recognition algorithm that constitutes one stage of a multistage recognition system. This sub-set of the problem has provided a useful experimental regime where the results of the learning-based approach can be compared to the optimum solution found by exhaustive search.

We intend to extend this single-stage learning algorithm to other tuning parameters of the recognition algorithm and apply the multistage learning algorithm to target discrmination algorithm parameters and eventually the target detection algorithm parameters. We also plan to incorporate prediction theory [15] into adaptive target recognition research. This initial work serves as a foundation and useful benchmark to assess future progress, where the higher problem dimensionality makes it impractical to exhaustively determine the optimal solution.



a articulated data **b** depression angle change **c** configuration change

Fig. 13a–c. ROC curves for MSTAR data



a articulated data **b** depression angle change **c** configuration change

Fig. 14a–c. Limit ROC curves for MSTAR data

Acknowledgements. This research is supported in part by grant F49620-97-1-0184. The contents and information do not necessarily reflect the positions and policies of the U.S. Government.

References

1. Au W, Roberts B (1996) Overview of a self-adaptive ATR system via context-based configuration and control. In: ARPA Image Understanding Workshop, February 1996, Palm Springs, Calif., Morgan Kaufmann Publishers, Inc., San Francisco, CA, pp 585–588
2. Bengio Y, Le Cun Y (1994) Word normalization for on-line handwritten word recognition. In: Proc. Int. Conf. on Pattern Recognition, 1994, Jerusalem, Israel. IEEE Computer Society Press, Los Alamitos, CA, pp 409–413
3. Bhanu B, Jones G (1997) Performance characterization of a model-based SAR target recognition system using invariants. Proc SPIE (Algorithms for Synthetic Aperture Radar Imagery IV) 3070: 305–321. Orlando, Florida, April 1997
4. Burges C, et al. (1992) Shortest path segmentation: A method for training a neural network to recognize character strings. In: Proc. Int. Joint Conf. on Neural Networks, Vol. 3, 1992, Baltimore, Md., IEEE, New York, NY, pp 165–172
5. Draper B (1996) Learning object grouping strategies for 2D and 3D object recognition. In: Proc. ARPA Image Understanding Workshop, February 1996, Palm Springs, Calif., Morgan Kaufmann Publishers Inc., San Francisco, CA, pp 1447–1454
6. Jones G, Bhanu B, Guo J (1998) Geometrical and magnitude invariants for recognition of articulated and non-standard objects in MSTAR images. In: George E. Lukes (ed) Proc. DARPA Image Understanding Workshop, 21–23 November 1998, Monterey, Calif., Morgan Kaufmann Publishers Inc., San Francisco, CA, pp 1163–1169
7. Khoral Research Inc. (1998) Khoros Pro v2.2 User's Guide. Addison Wesley Longman, Reading, Mass.
8. Lamden Y, Wolfson H (1988) Geometric hashing: A general and efficient model-based recognition scheme. In: Proc. Int. Conference on Computer Vision, December 1988, Tampa, Florida, IEEE Computer Society, pp 238–249
9. Le Cun Y, Bottou L, Bengio Y (1997) Reading checks with graph transformer networks. In: Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Vol. 1, 1997, Munich, Germany, IEEE Computer Society Press, Los Alamitos, CA, pp 151–154
10. Peng J, Bhanu B (1998) Delayed reinforcement learning for adaptive image segmentation and feature extraction. IEEE Trans Syst Man Cybern 28(3): 482–488
11. Peng J, Bhanu B (1998) Closed-loop object recognition using reinforcement learning. IEEE Trans Pattern Anal Mach Intell 20(2): 139–154
12. Bhanu B, Peng J (1998) Adaptive integrated image segmentation and object recognition. In: Proc. Eleventh Vision Interface Conference, June 1998, Vancouver, British Columbia, Canadian Image Processing and Pattern Recognition Society, pp 471–478
13. Novak L, Owirka G, Netishen C (1993) Performance of a high-resolution polarimetric SAR automatic target recognition system. Lincoln Lab J 6(1): 11–24

14. Jones G, Bhanu B (1999) Recognition of articulated and occluded objects. IEEE Trans Pattern Anal Mach Intell 21(7): 603–613
15. Boshra M, Bhanu B (1999) Performance prediction and validation for object recognition. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 1999, Fort Collins, Colorado, pp 380–386

Bir Bhanu received the S.M. and E.E. degrees in electrical engineering and Computer Science from the Massachusetts Institute of Technology, Cambridge, the Ph.D. degree in electrical engineering from the Image Processing Institute, University of Southern California, Los Angeles. Currently he is a Professor of electrical engineering and computer science and Director of Center for Research in Intelligent Systems at the University of California, Riverside. Previously, he was a Senior Honeywell Fellow at Honeywell Systems and Research Center, Minneapolis, MN. He has been the principal investigator of various programs for DARPA, NASA, NSF, AFOSR, ARO and other agencies and industries in the areas of learning and vision, image understanding, pattern recognition, target recognition, navigation, image databases, and machine vision applications. He is the co-author of books on "Computational Learning for Adaptive Computer Vision," (Forthcoming), "Genetic Learning for Adaptive Image Segmentation" (Kluwer 1994), and "Qualitative Motion Understanding" (Kluwer 1992). He holds 10 U.S. and international patents and over 200 reviewed technical publications in the areas of his interest. Dr. Bhanu is a Fellow of the IEEE and the AAAS. He is a member of ACM, AAAI, Sigma Xi, Pattern Recognition Society, and SPIE.

Yingqiang Lin received his B.S. and M.S. degree in computer science at Fudan University, Shanghai, China in 1991 and 1994 respectively. Currently, he is a Ph.D. student in Computer Science at the Center for Research in Intelligent Systems, University of California, Riverside. His research interests include image processing, pattern recognition and machine learning.

Grinnell Jones III was a National Merit Scholar who received his BS degree in mechanical engineering from the Massachusetts Institute of Technology in 1966, MS in aerospace engineering (with Distinction) from the Air Force Institute of Technology in 1971 and MS in computer science from the University of California, Riverside in 1997. After a 25 year career in development engineering, missile operations and acquisition management with the U.S. Air Force, Lt. Col. (USAF Retired) Jones has been conducting research in automatic target recognition using Synthetic Aperture Radar imagery for the past four years with the University of California, Riverside. His research interests include object recognition, computer vision, machine learning, image and video databses, and systems applications.

Jing Peng received the B.S. degree in Computer Science from Beijing University of Aeronautics and Astronautics, China, the M.A. degree in Computer Science from Brandeis University and the Ph.D. degree in Computer Science from Northeastern University, Boston, Massachusetts . From 1994 to 1995 he was a Post Doctoral Fellow in the Visualization and Intelligent Systems Laboratory at the University of California, Riverside. During 1996 he worked as Senior Research Scientist in the Machine Vision Department at the Amherst Systems Inc. in Buffalo, NY. From January 1997 to July 1999 he was a Research Scientist in the Center for Research in Intelligent Systems at the University of California, Riverside. Since August 1999 he has been an Assistant Professor in the Computer Science Department at the Oklahoma State University. Dr. Peng's interests are in the areas of machine learning, content-based image retrieval, pattern classification, and learning and computer vision.