Recognition of Articulated and Occluded Objects

Grinnell Jones, III, and Bir Bhanu, *Fellow*, *IEEE*

Abstract—A model-based automatic target recognition (ATR) system is developed to recognize articulated and occluded objects in Synthetic Aperture Radar (SAR) images, based on invariant features of the objects. Characteristics of SAR target image scattering centers, azimuth variation, and articulation invariants are presented. The basic elements of the new recognition system are described and performance results are given for articulated, occluded and occluded articulated objects and they are related to the target articulation invariance and percent unoccluded.

Index Terms—Articulation invariance, automatic target recognition, azimuth variance, empirical performance modeling, geometric hashing, synthetic aperture radar images.

1 Introduction

Automated object recognition in SAR imagery is a significant problem because new image collection platforms can produce far more imagery (terabytes per day per aircraft) than the declining ranks of image analysts are capable of handling. In this paper, the problem scope is the recognition subsystem itself, starting with SAR chips of target vehicles and ending with the vehicle identification. The specific challenges we address are automated recognition of vehicles, e.g., tanks, that can be in articulated configurations and can be partially hidden (occluded) [3]. Recognition methods involving template matching are not useful in these cases because articulation or occlusion changes global features like the object outline and major axis [11], [13]. Constrained models of parts and joint articulation, used in optical images [2], [7], are not appropriate for the relatively low resolution, nonliteral nature, and complex part interactions of SAR images. Prior detection theory, pattern recognition and neural network approaches to SAR recognition all tend to use global features that are optimized for standard, nonarticulated, nonoccluded configurations [10]. Some of these SAR ATR techniques, MINACE filters [4] and invariant histograms [8], have reported limited test results for small amounts of occlusion, typically 25 percent or less. In addition, there is no published paper by others on the recognition of articulated or occluded articulated objects in SAR images.

Our approach to object recognition is specifically designed for SAR images. In contrast to the passive vision systems, the scale of the SAR image is fixed by characteristics of the radar. However, where optical images are mainly formed as a result of diffuse reflections from a noncoherent source of light, SAR images are formed primarily from specular reflections from a coherent source; they are nonliteral and vary quickly and abruptly with small pose angle variations [12]. The peaks (local maxima) in radar return are related to the physical geometry of the object. The relative locations of these scattering centers are independent of translation and serve as distinguishing features. Six-inch resolution X-band images (10.0 GHz center frequency, 1.0 GHz bandwidth, 5.6° angular span) are used to provide a rich feature set to facilitate the task of recognizing articulated and occluded objects. Because six-inch resolution SAR data is not available to the public, all the experimental data is based on images generated by the XPATCH radar signature prediction code [1].

We approach the problem of recognizing articulated and occluded objects from the fundamentals of SAR images. Based on locations of the scattering centers, we characterize SAR image variance with object pose or azimuth to establish that 360 azimuth models at 1° intervals are appropriate for recognizing articulated and occluded SAR objects. This is in contrast to others who have used 5° to 10° intervals [8], [11] for unoccluded objects. Since the radar depression angle to the target is controllable, or known, it is fixed at 15° for this study.

We identify (and measure) the invariance of SAR scatterer locations with object articulation. Based on these invariants, we develop a SAR specific recognition system that uses standard nonarticulated recognition models to successfully recognize articulated and occluded versions of these objects. This avoids the combinatorial explosion of modeling 360 turret angles times 360 aspect views for tank targets. The SAR recognition system has an off-line model construction phase and an on-line recognition process. The recognition model is a look-up table that relates relative distances among the scattering centers (in the radar range and cross range directions) to object type and azimuth. In contrast to geometric hashing [9], we consider only relative locations without preserving the basis location. This automatically handles translation without having to explicitly determine the appropriate translation, which can be expensive for highly occluded vehicle sized objects at high

• *The authors are with the Center for Research in Intelligent Systems, University of California, Riverside, CA 92521. E-mail: {grinnell,bhanu}@cris.ucr.edu*

Manuscript received 23 Dec. 1997; revised 30 Mar. 1999. Recommended for acceptance by K. Ikeuchi. For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 107611.

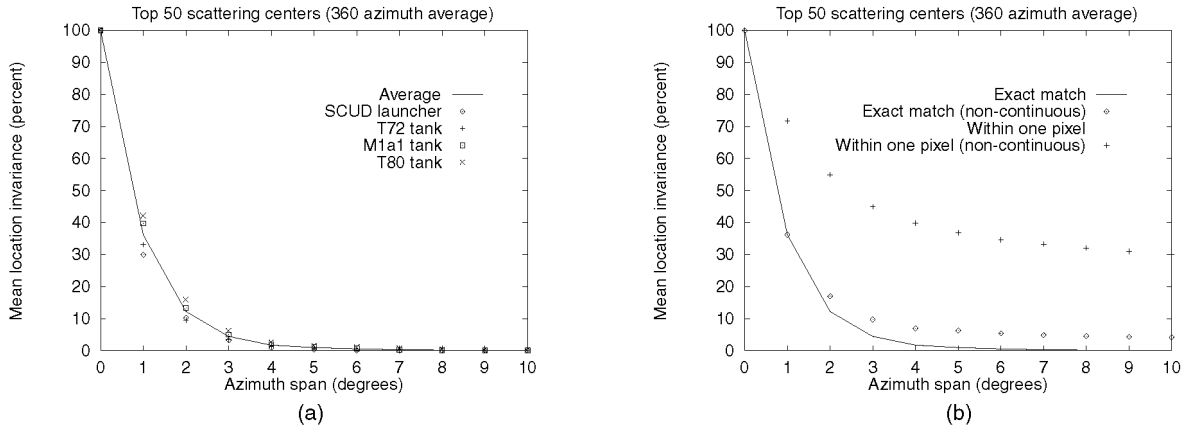

Fig. 1. Scatterer location persistence. (a) Exact match invariance. (b) Average persistence.

resolution. The recognition process is an efficient search for *positive evidence*, using relative locations of scattering centers in the test image to access the look-up table and generate votes for the appropriate object (and azimuth).

2 SAR Target Image Characteristics

The relative locations of peaks in the radar return 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 subcomponents of vehicle sized objects at six-inch 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. These peaks tend to be "spiky" (approximating an impulse function) so their locations should be relatively insensitive to noise. Furthermore, we want to explore what can be accomplished with peaks alone, before more complex features are used. At scattering centers, the amplitude of the SAR return is greater than the surrounding (eight) neighbors; if equal, we choose the last in raster scan order. The 4,320 SAR images, or target chips, of the T72, T80, M1a1, FRED tanks, and the SCUD missile launcher used in this research have a range of 52 to 284 local peaks (above a -100 dBsm signal strength threshold, expressed as a radar cross section in square meters, for negligibly small returns). Hence, an upper limit of the 50 strongest scatterers is used to avoid introducing an uncontrolled variable: the number of scattering centers actually available for some instance of an object at an azimuth angle orientation. (In practice, if some target aspect did not have 50 peaks, the performance would degrade as if the "missing" peaks were occluded.) The locations of the 50 strongest scatterers are related to the aspect and detailed geometry of the object and many of the locations may change with each aspect view.

The typical rigid body rotational transformations for viewing objects in the visual world do not apply much for the specular radar reflections of SAR images. This is because a *significant* number of features *do not* typically persist over a few degrees of rotation. Persistence can be expressed in terms of the percentage of the scattering center locations that are unchanged over a certain span of azimuth angles. It can be measured (for some base azimuth θ_0) by rotating the pixel locations of the scattering centers from an image of an object at azimuth θ_0 by an angle $\Delta\theta$ and comparing the resulting range and cross-range locations (r_1, c_1) with the locations (r_2, c_2) from an image of the same object at azimuth $\theta_0 + \Delta\theta$. The images were generated with a consistent origin, so no translation was required for registration. For the locations to match, the centroid of the rotated scatterer pixel from the image at θ_0 azimuth is within the pixel boundaries of a corresponding scatterer in the image at $\theta_0 + \Delta\theta$ (i.e., $\max\{|\mathrm{r}_2 - \mathrm{r}_1|, |\mathrm{c}_2 - \mathrm{c}_1|\} \leq 0.5$). To determine scattering center locations that persist over a span of angles, there is an additional constraint that, for a matching scattering center to "persist" at the kth span $\Delta\theta_k$, it must have been a persistent scattering center at all smaller spans $\Delta\theta_j$, where $0 \leq j < k$. Averaging the results of these unchanged scattering center locations over 360 base azimuths gives a "mean location invariance" for an object.

Fig. 1a gives the mean location invariance as a function of azimuth angle span for the strongest 50 scattering centers of the SCUD launcher, T72, M1a1, and T80 tanks, and the average of the four objects. Note that, while an average of 36.3 percent of the scattering center locations remained unchanged for a 1° azimuth change, only 1.0 percent were unchanged over a 5° azimuth span. Fig. 1b shows the average results for the four objects with various definitions of persistence. The "average" results of Fig. 1a are the "exact match" case in Fig. 1b. In the "exact match (noncontinuous)" case, the constraint that the scatterer exist for the entire span of angles is relaxed (i.e., scintillation is allowed) and small numbers of scatterers (6.3 percent at 5°) "persist" at larger spans. In the "within one pixel" case, where the scatterer location is allowed to move into one of the adjacent pixel locations ($\max\{|\mathrm{r}_2 - \mathrm{r}_1|, |\mathrm{c}_2 - \mathrm{c}_1|\} \leq 1.5$), the "location invariance" is initially much higher (71.7 percent vs. 36.3 percent at 1° in Fig. 1a), but declines rapidly to 15.1 percent at 5°. In addition to movement "within one pixel," if the scattering center is allowed to scintillate, then significant "persistence" can be obtained (34.6 percent at 5° and 30.0 percent at 10°). The "within one pixel (noncontinuous)" results in Fig. 1 are consistent with the one foot resolution ISAR results of Dudgeon et al. [5], whose definition of persistence allowed scintillation (i.e., in that

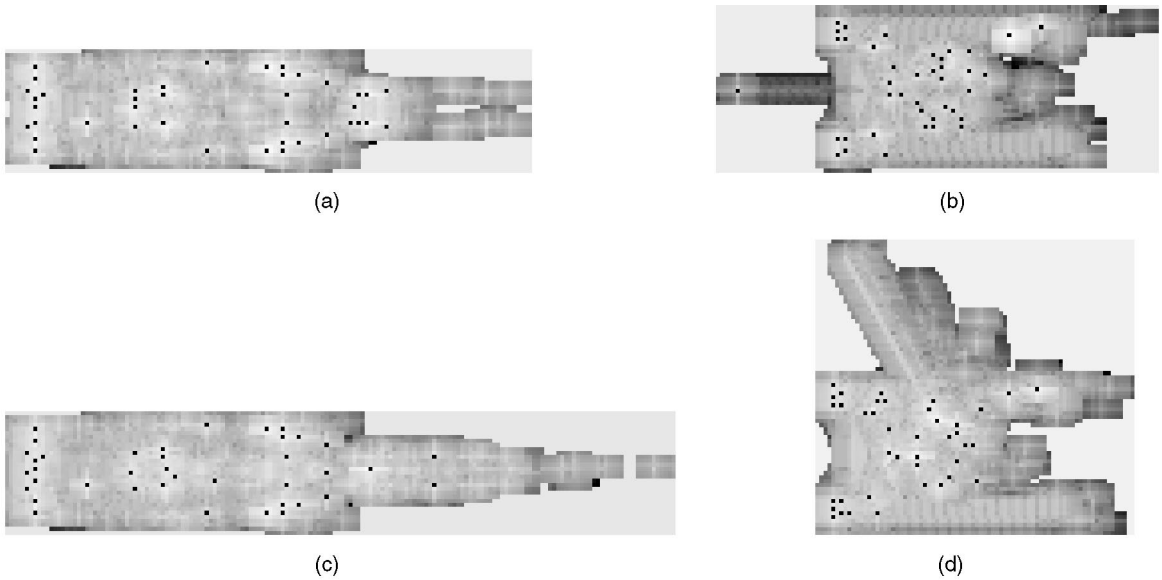

Fig. 2. SCUD launcher and T72 tank articulation examples. (a) SCUD launcher, missile down, (b) T72 straight turret, (c) SCUD launcher, missile up, (d) T72 -60° turret.

work, a point was required to be present/absent for two consecutive angles, 1° apart, to appear/disappear, thus a feature point would be "persistent" if it appears and then disappears in images separated by 1°). However, for recognition purposes, we do not want the scatterer locations to wander and scintillate, so the lower "exact match" results in Fig. 1 are appropriate. Because of azimuthal variation and the goal of recognizing articulated and occluded objects, in this research, we use 360 azimuth models (1° intervals) for each target.

The existence of articulation invariants in six-inch resolution SAR imagery can be seen in Fig. 2, where the locations of scattering centers are indicated by the black squares. In the example of the SCUD launcher, with the radar directed (from the left in Fig. 2) at the front (cab end) of the launcher, many of the details from the launcher itself are independent of whether the missile is erect or down. In the similar view of the T72 tank, many of the details from the hull are independent of the turret angle. An example of articulation invariance is shown in Fig. 3, which plots the percentage of the strongest 50 scattering centers for the T72 tank that are in exactly the same location (the "exact match" case) with the turret rotated 60° as they are with the turret straight forward, for each of 360 azimuths. (The remaining peaks in the top 50 are in different locations due to the articulation of the turret.) Later results (Figs. 7 and 8) show that recognition failures tend to occur when the articulation invariance is low. The mean values of articulation invariance for the SCUD launcher, T72, M1a1, and T80 tanks are 52.76 percent, 46.57 percent, 36.57 percent, and 58.40 percent, respectively. The smaller average articulation invariance for the M1a1 tank is expected because the M1a1 tank has a comparatively much larger turret than the other tanks.

3 SAR Recognition System

An appropriate local coordinate reference frame is critical to reliably establishing locations of features in SAR images of articulated and occluded objects because global coordinates and global constraints do not work with articulated and occluded objects, as illustrated in Fig. 2, where the center of mass and the principal axes change with articulation. (In addition, these conditions and effects are a difficult problem for any approach that tries to estimate the object pose to reduce the search space.) In the geometry of a SAR sensor, the image range (radar beam direction) and cross-range directions are known and any local reference point chosen, such as a scattering center location, establishes a reference coordinate system. (The scattering centers are local maxima in the radar return signal.) The relative distance and direction of the other scattering centers can be expressed in radar range and cross-range coordinates and naturally tessellated into integer buckets that correspond to the radar

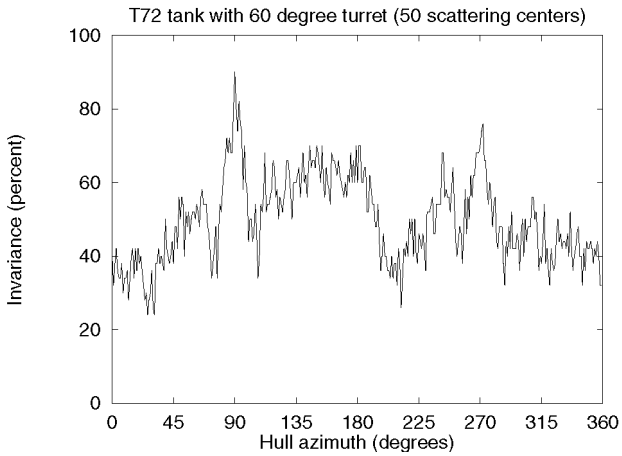

Fig. 3. T72 articulation invariants.

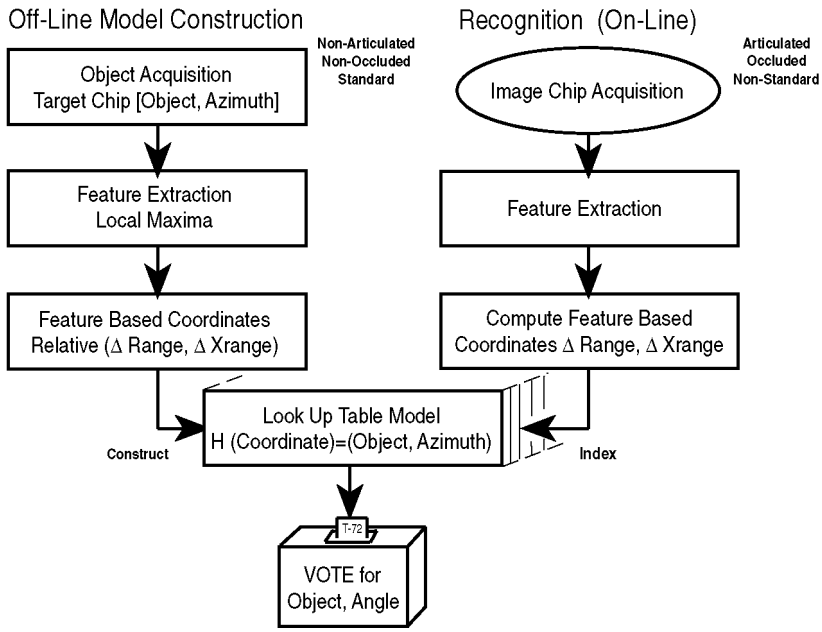

Fig. 4. SAR recognition system.

range/cross-range bins. The recognition system takes advantage of this natural system for SAR images, where a single basis point performs the translational transformation and fixes the coordinate system to a "local" origin. For ideal data, picking the location of the strongest scattering center as the reference point is sufficient. However, for potentially corrupted data where any feature point could be spurious or missing (due to the effects of noise, articulation, occlusion, nonstandard configurations, etc.), the process needs to continue, using other scattering centers as the reference point to ensure a valid origin is obtained. We used all the scattering center locations in turn as reference points; thus, to handle occlusion and articulation, the size of the look-up table models and the number of relative distances that are considered in a test image is increased from M to $\mathrm{M}(\mathrm{M}-1)/2$, where M is the number of scattering centers used. Heuristics could be applied to use fewer reference points for increased efficiency. The complete 360 aspect recognition table for a single object using 50 scatterers takes about 10 Mbytes of computer memory.

The basic recognition system architecture, Fig. 4, is an off-line model construction process and a similar on-line recognition process. The approach is based on local features and local reference coordinate systems. A systematic method is employed for constructing recognition models of objects that are not articulated. These models are stored in a look-up table and, then, during the on-line phase, local image features are used to index into the look-up table of models and recognize the same objects in articulated configurations. The image features used are the positions of the scattering centers (local maxima in the signal strength). The number of scattering centers used is a design parameter that is optimized based on experimental results. The positions of the scattering centers are expressed in relative distances in the known SAR range and cross-range coordinates. The model construction technique extracts these relative distances of the scattering centers from the nonarticulated training data for all 360 azimuths for each target type. The model database is basically a table that relates these distances to object labels (target type and azimuth). In the recognition phase, the relative positions of scattering centers are obtained from the articulated (or occluded) test data. These relative distances are indices into a look-up table which provides the associated object label(s) that are used to accumulate evidence for target identification. The process is repeated with different scattering centers as reference points, providing multiple "looks" at the database. The recognition algorithm makes $\mathrm{M}(\mathrm{M}-1)/2$ queries of the lookup table to accumulate evidence for the appropriate target type and azimuth angle. The only models (object, azimuth labels) associated with a specific lookup table entry are the "real" model and any other models that happen, by coincidence, to have a feature pair with the same relative distance. While these collisions may occur at one relative location, the same random object-azimuth pair does not keep showing up often at other relative locations, while the "correct" object-azimuth pair does. In contrast to many model-based approaches [6], we are not "searching" all the models. Instead, we are doing table look-ups based on the relative distances between the strongest scatterers in the test image. In practice, this look-up process is quite fast, e.g., for 50 peaks it takes a SPARC 20 workstation less than one-half second to completely process a test target chip.

The basic decision rule used in the recognition system is to select the object-azimuth pair with the highest accumulated vote total. In cases of a tie, the choice is arbitrarily awarded to the larger object in a consistant manner. In this forced recognition case, the system is forced to make a choice among the models in its decision table and cannot produce an answer of "unknown." (Unless otherwise stated, all the results are based on

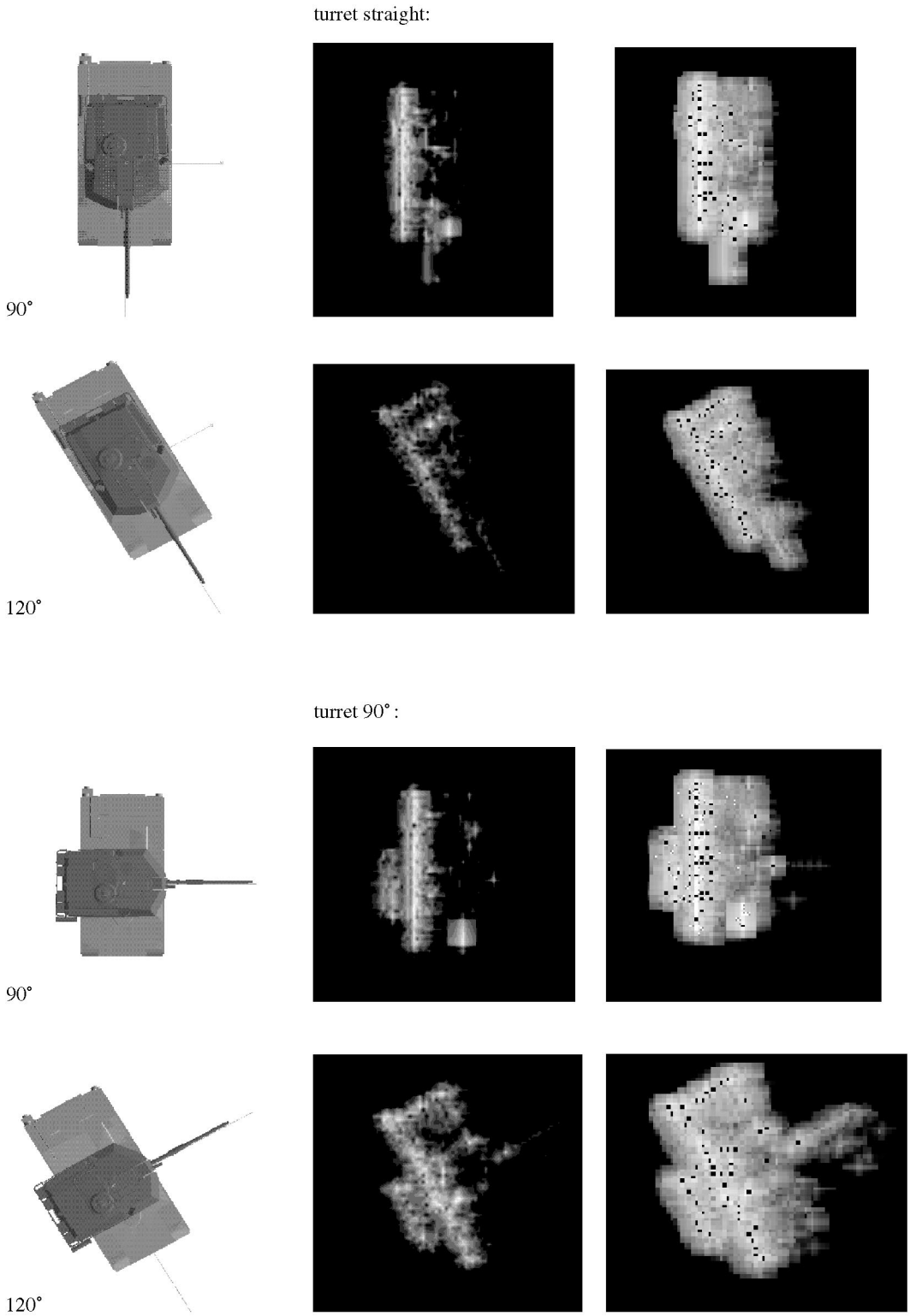

Fig. 5. Examples of M1a1 tank geometry, SAR image, and scattering centers for 90° and 120° hull azimuths.

forced recognition.) One method of introducing a criteria for the quality of the recognition result is to define a decision rule that the ratio of votes for the potential winning object, v_1, to the votes for the second place different object, v_2, is greater than some minimum ratio r. This can be expressed as rule 1: $v_1/v_2 > r$ (for $r > 1.0$). Now, if vote ratio is less than r, the test result is given as an "unknown" object. This not only introduces a quality constraint, it also allows the recognition engine to handle test cases with objects that are not contained in the table

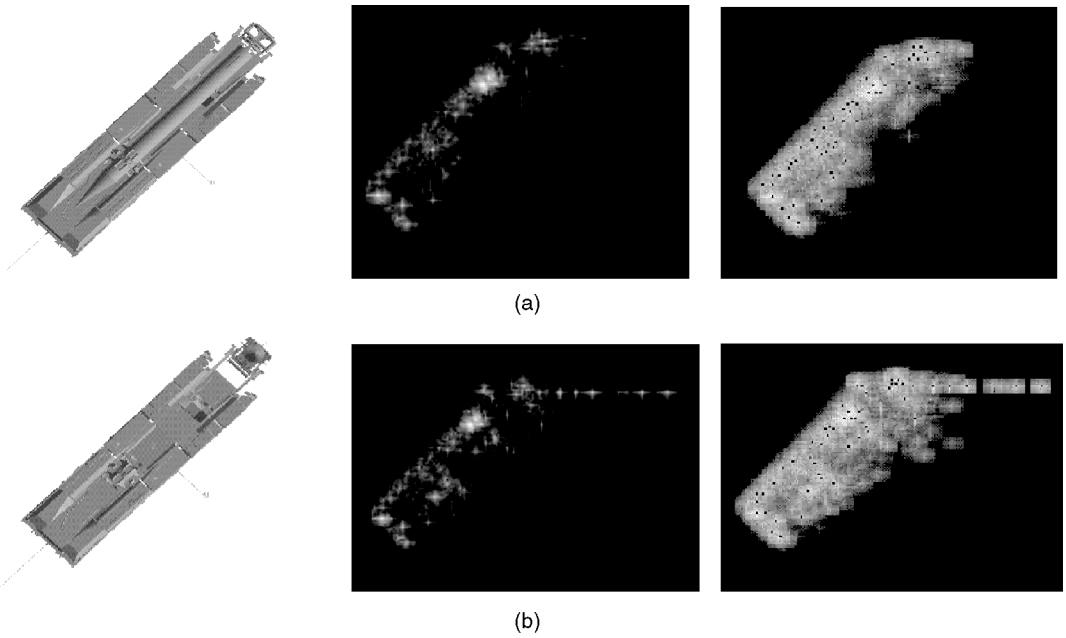

Fig. 6. Examples of SCUD Launcher geometry, SAR image, and scattering centers. (a) 45° azimuth, missile down. (b) 45° azimuth, missile up.

of models (i.e., true unknown objects). If we formulate the recognition problem as a hypothesis about the object identification of a test case, we can define the quantities:

$$P_f = P\{\text{decide any object } | \text{unknown is true}\};$$
$$P_d = P\{\text{decide any object}|\text{object is true}\};$$
$$P_m = P\{\text{decide unknown}|\text{object is true}\};$$
$$\text{PCI} = P\{\text{decide correct object}|\text{object is true}\}.$$

For historical reasons, the first three above are referred to as P_f = "Probability of false alarm," P_d = "Probability of detection," and P_m = "Probability of a miss." PCI is the "Probability of correct identification". The usual ROC (receiver operating characteristic) curve for object detection is a graph of P_d versus P_f as a function of varying the design parameters of the system. In our case, for object recognition, by using the more stringent definition, PCI, we obtain analogous ROC curves by varying r for decision rule 1.

4 Performance Results and Analyses

The XPATCH code is used to generate six-inch resolution (hh polarization) SAR target chips at 360 azimuth angles (at a 15° depression angle, 90° squint angle) from CAD models of the various objects with numbers of surface facets ranging from 5,345 to 32,954. XPATCH is considered the best radar signature prediction code available in the world; it has been extensively validated (by others) against primitive objects and full scale vehicles [1]. For four objects, the nonarticulated set used for model construction is 1,440 (4×360) target chips. There are seven sets of articulated test data (SCUD Launcher with the missile up, and the T72, M1a1, and T80 tanks with 60° and 90° turret angles) with an additional 2,520 target chips. The scattering center locations are obtained at local maxima in the signal amplitude. Examples, at various azimuths, of the object geometry, SAR image, and (strongest 50) scattering center locations are shown for both nonarticulated and articulated cases of the M1a1 tank (Fig. 5), and the SCUD launcher (Fig. 6). (Figs. 5 and 6 are not to scale; the XPATCH image is displayed at eight intensity levels and the image with the scattering centers at 256 levels). Various controlled experiments were performed to thoroughly investigate our approach.

4.1 Articulated Objects

The experimental results of 2,520 trials with articulated test objects for the recognition system using 50 scattering centers and a four object table (SCUD missile launcher, T72, M1a1, and T80 tanks) are shown as a confusion matrix in Table 1. The overall performance is a 93.14 percent probability of correct identification (PCI). The azimuth accuracy is also shown in Table 1, where "e" is an exact pose match and "c" indicates a correct pose within $\pm 5^\circ$. Note that these azimuth results include cases with the correct pose, but the wrong object. The azimuth results are reported for the hull angle. In the case of the M1a1 tank, decreased azimuth accuracy results when identifications are based on the (relatively large) turret rather than the hull.

The number of scattering center locations used is a design parameter that can be tuned to optimize performance of the recognition system. For the objects and articulations used in Table 1, the maximum performance is achieved at 50 scattering centers (93.14 percent), but virtually the same performance could be found at 42 scattering centers (93.10 percent). A more optimal system with 35 scattering centers achieves similar performance, 92 percent PCI, with slightly less than half the storage and twice the speed of 50 scattering centers. Below 10 scatterers (82 percent PCI) the performance drops off rapidly.

The detailed recognition results can be related to the articulation invariance of articulated objects. The recogni-

TABLE 1
Confusion Matrix for Articulated Objects (50 Scatterers, e = Exact Pose, c = Pose within $\pm 5^\circ$)

Test	Identification results				Pose accuracy			
articulated	Non-articulated models				Non-articulated models			
test targets	SCUD	T72	M1a1	T80	SCUD	T72	M1a1	T80
	down	0^o turret			down	0^o turret		
SCUD missile up	360				360e			
T72 60^o turret		335	7	18		333e	1c	2c
90^o turret		327	8	25		323e		
M1a1 60^o turret		1	300	59			261c,254e	4c,1e
90^o turret		2	305	53		2c	272c,261e	4c,1e
T80 60^o turret				360				356c,355e
90^o turret				360				360e

tion failures for the T72 tank with the turret at 90° are plotted (as "diamonds") on the curve of percent invariance vs. azimuth in Fig. 7. These results show that recognition failures generally occur for azimuths where the percent invariance is low. Fig. 8 shows how the PCI varies with the percent invariance. The points at low invariance values are misleading because they are due to a few correct identifications for the M1a1 tank, where the invariance (measured with respect to the hull) is low, yet a correct identification is made from a number of features which are on the large turret. The overall recognition system performance is almost perfect for invariance values greater than 40 percent (i.e., down to 60 percent spurious data).

Test data with positional noise is simulated by adding Gaussian noise with zero mean and standard deviation sigma (in units of six-inch resolution pixels) to the range and cross-range locations of the scattering centers. For the articulated object cases, the SCUD launcher had the missile erect and the T72, M1a1, and T80 tanks had their turrets rotated 90°. The average and individual test object recognition performance results with varying amounts of positional noise are shown in Fig. 9a for a four object recognition table (with 20 scattering centers). For larger values of positional noise, the two physically smaller objects (T80 and M1a1) are correctly identified more often than average because they tend to be characterized by relatively common

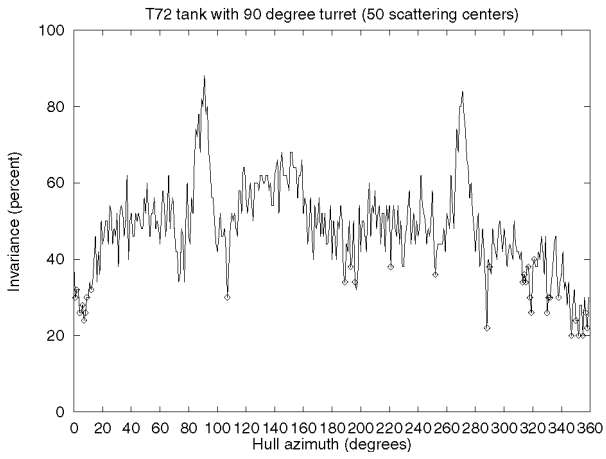

Fig. 7. T72 recognition failure plot (◇) on articulation invariance curve.

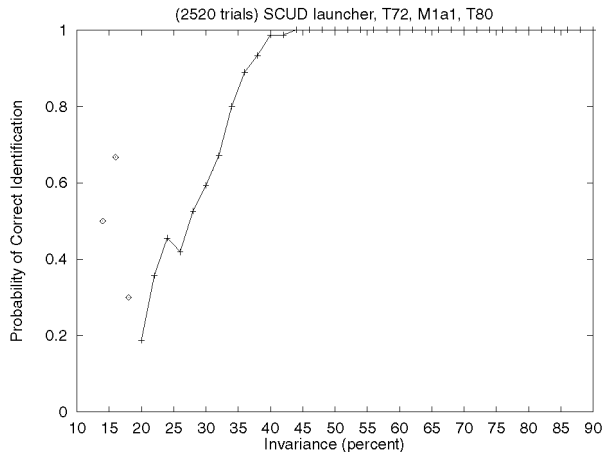

Fig. 8. Recognition rate and articulation invariance.

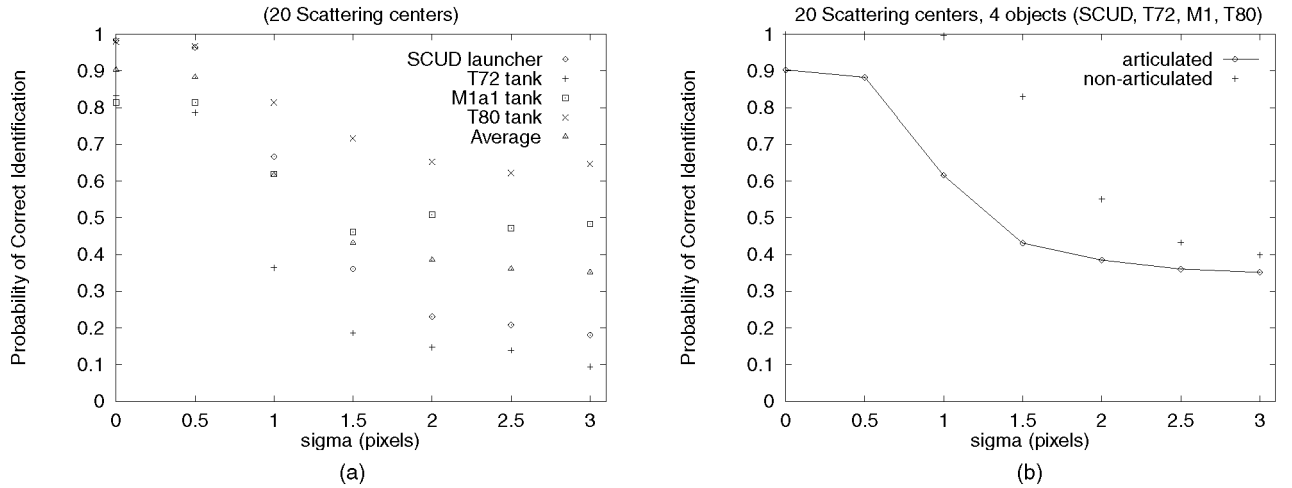

Fig. 9. Effect of positional noise on articulated and nonarticulated object recognition. (a) Articulated objects. (b) Articulated vs. nonarticulated.

short distances and positional noise often results in collisions with other T80 and M1a1 relative distances. Conversely, perturbations in the relatively rare long distances are more likely to find vacancies in the look-up table, so the larger objects are correctly identified less often. Fig. 9b compares the effect of positional noise on the recognition of articulated objects and the nonarticulated versions of these same objects (with 20 scattering centers). Since the recognition models are based on the nonarticulated objects, this illustrates the effects of positional noise alone, compared with articulation and positional noise.

4.2 Occluded (Nonarticulated) Objects

The occluded test data is simulated by starting with a given number of the strongest scattering centers and then removing the appropriate number of scattering centers encountered in order, starting in one of four perpendicular directions d_i (where d_1 and d_3 are the cross range directions, along and opposite the flight path, respectively, and d_2 and d_4 are the up range and down range directions). Then, the same number of scattering centers (with random magnitudes) are added back at *random locations* within the original bounding box of the chip. This keeps the number of scatterers constant and acts as a surrogate for some potential occluding objects. Each data set is 5,760 test cases (4 objects $\times$ 4 directions $\times$ 360 azimuths). For the nonarticulated occluded tests, (the objects are the T72, M1a1, and T80 tanks with turret at 0° and the SCUD launcher with the missile down), there are 51 data sets (for 10, 30, and 50 scattering centers with 10 to 90 percent occlusion in 10 percent steps and the same for 20 and 40 scatterers plus 55, 65, and 75 percent occlusion) for a total of 293,760 test cases. Actually, only 50 data sets with a total of 288,000 test cases are used because the data set of 10 scattering centers with 90 percent occlusion has less than two valid scattering centers for each case.

The performance of the recognition system with nonarticulated occluded objects is shown in Fig. 10 in terms of the PCI as a function of percent occlusion with the "number of scattering centers used" as a parameter. The results of 288,000 test cases are shown, where each point for a specific percent occlusion and number of

TABLE 2
Confusion Matrix for 60 Percent Occluded Object Identification (40 Scatterers)

60% occluded test targets	Non-articulated models			
	SCUD down	T72	M1a1	T80
		0^o turret		
SCUD	1435	0	2	3
T72	0	1414	11	15
M1a1	0	0	1408	32
T80	0	0	17	1423

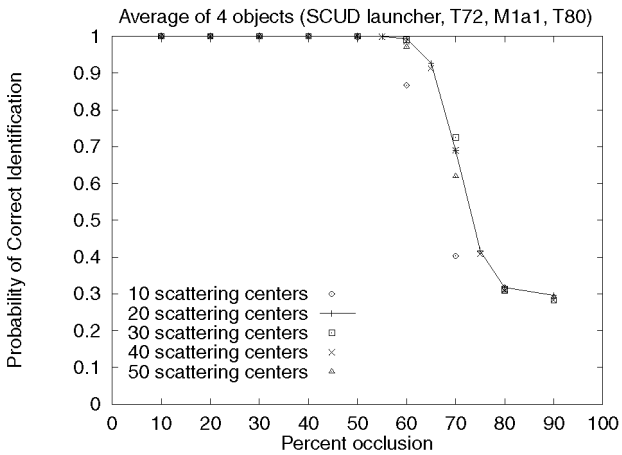

Fig. 10. Effect of occlusion on recognition rate.

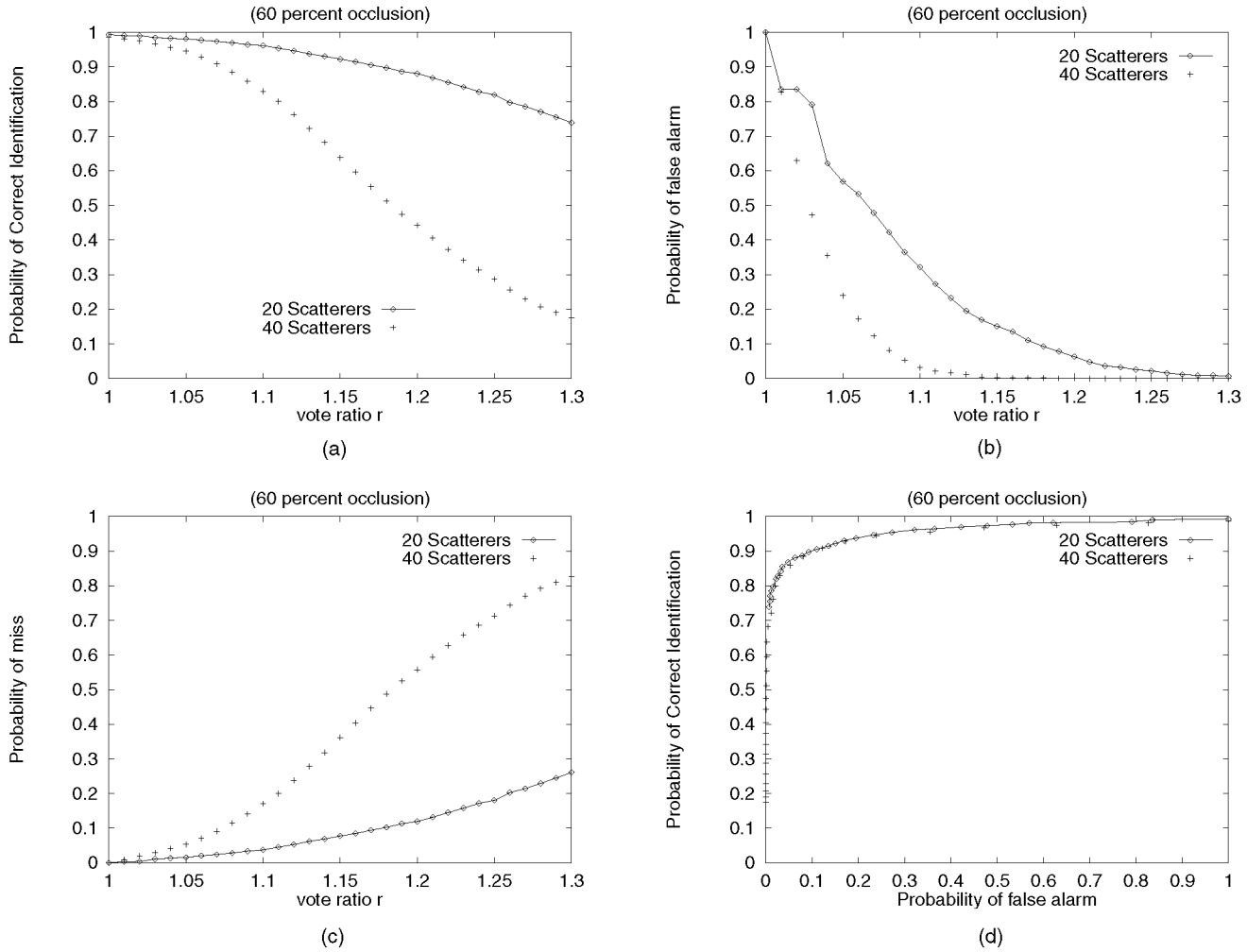

Fig. 11. Receiver operating characteristic (60 percent occlusion). (a) Probability of correct identification. (b) Probability of false alarm. (c) Probability of miss. (d) Receiver operating characteristic.

scattering centers is the average PCI for all four occlusion directions, the four objects and the 360 azimuths (5,760 test cases). The overall recognition system performance is almost perfect for up to 60 percent occlusion. (This corresponds to results shown in Fig. 8 for articulation invariance of 40 percent and above.) By 80 to 90 percent occlusion, the results are not much better than the 0.25 PCI one would expect by chance from the four possible objects. Typical results for 40 scattering centers with 60 percent occlusion are shown as a confusion matrix in Table 2 (with four occlusion directions and 360 azimuths, there are 1,440 cases for each object).

The decision rule used to determine the recognition system result is a design parameter that can be varied to optimize the system performance. Using rule 1 (which states that the ratio of votes for the potential winning object, v_1, to the votes for the second place different object, v_2, must be greater than some minimum ratio r), results are obtained for a four object table (SCUD launcher, T72, M1a1, and T80) with test data that are occluded versions of these same four objects and FRED tank as an unknown object. Figs. 11a, 11b, and 11c give the probability of correct identification, probability of false alarm, and probability of miss, respectively, as a function of r for 20 and 40 scattering centers and 60 percent occlusion. The resulting ROC curves (PCI vs. P_f) for 20 and 40 scatterers and 60 percent occlusion are shown in Fig. 11d. Fig. 12 gives the ROC curves of the 40 scattering center recognition system for 50 to 70 percent occlusion using decision rule 1. At the upper-right point on each ROC curve (maximum PCI and maximum P_f), r is equal to 1.0, which corresponds to the forced recognition case.

An illustrative confusion matrix for 60 percent occlusion with $r = 1.1$ (40 scatterers) is shown in Table 3 (with four occlusion directions and 360 azimuths, there are 1,440 cases for each object). Table 3 has only one misidentification (one T80 test case misclassified as an M1a1) with a PCI of 0.829, compared with the forced recognition ($r = 1.0$) confusion matrix in Table 2 for the same (60 percent occlusion, 40 scatterer) case, which has 80 misidentifications with a PCI of 0.986. This is typical because, as vote ratio increases, not only do some of the "weaker" identifications move into the unknown column, but also many of the misidentifications become unknowns.

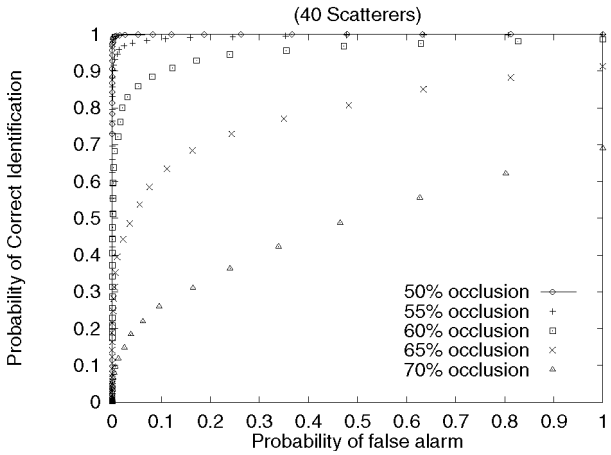

Fig. 12. Receiver operating characteristics (40 scatterers).

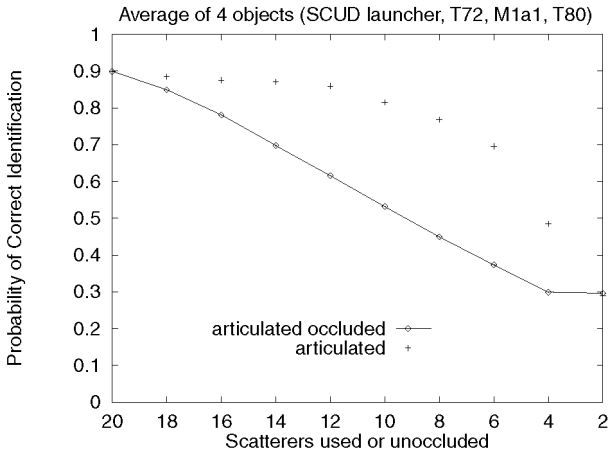

Fig. 13. Articulated object and occluded articulated object performance results.

4.3 Occluded Articulated Objects

The occluded articulated data is produced in the same manner as the nonarticulated occluded data. The same tanks are used, but with a 90° turret rotation and the missile is erect on the SCUD launcher. There are nine occluded articulated data sets (for 20 scattering centers with 10 to 90 percent occlusion) for a total of 51,840 test cases. Fig. 13 compares the performance results of the articulated unoccluded and occluded objects for cases with the same number of valid scatterers (i.e., "scatterers used" in the unoccluded cases or "unoccluded scatterers" in the occluded cases). In the occluded data, the valid points are "clustered" in a neighborhood which gets smaller as the occlusion increases (and the number of valid scatterers decreases). These relatively worse results for the naturally "clustered" occluded articulated data, compared with the more widely distributed unoccluded articulated data for the same number of valid scattering centers, illustrate the importance of the relatively rare long distances.

TABLE 3
Typical Confusion Matrix for 60 percent Occluded Objects (40 Scatterers, $r = 1.1$).

60% occluded test targets	Identification results				
	SCUD	T72	M1a1	T80	unknown
SCUD	1372	0	0	0	68
T72	0	1238	0	0	202
M1a1	0	0	1040	0	400
T80	0	0	1	1126	313
FRED	0	0	9	36	1395

5 Conclusions

The XPATCH generated, six-inch resolution SAR imagery has great azimuthal variation that can be successfully captured by using 360 azimuth models for a given depression angle. Useful articulation invariant features are found in SAR images of vehicles. The feasibility of a new concept for a SAR recognition system to identify articulated and occluded objects based on nonarticulated recognition models is demonstrated. The performance of the recognition system can be predicted by the percent articulation invariance (or percent unoccluded) when comparing the scattering center locations of the articulated (or occluded) test images with the nonarticulated model scattering center locations. The results indicate the importance of the relatively rare long distances and suggest an explanation why this approach, which can use long distances (if available), could have an advantage over others [8] that are restricted to a "neighborhood." The power of this technique comes from the combination of a SAR specific approach for recognition, models properly accounting for azimuthal variance, articulation invariants, and the resolution of the sensor data.

Acknowledgments

This work was supported in part by grants MDA972-93-1-0010, F49620-97-1-0184, DAAG55-97-1-0036, and DAAH049510049. The contents and information do not reflect positions or policies of the U.S. Government.

References

[1] D. Andersch, S. Lee, H. Ling, and C. Yu, "XPATCH: A High Frequency Electromagnetic Scattering Prediction Code Using Shooting and Bouncing Ray," *Proc. Ground Target Modeling and Validation Conf.*, pp. 498-507, Aug. 1994.

[2] A. Beinglass and H. Wolfson, "Articulated Object Recognition, or: How to Generalize the Generalized Hough Transform," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 461-466, June 1991.

[3] B. Bhanu, D. Dudgeon, E. Zelnio, A. Rosenfeld, D. Casasent, and I. Reed, "Introduction to the Special Issue on Automatic Target Detection and Recognition," *IEEE Trans. Image Processing*, vol. 6, no. 1, pp. 1-6, Jan. 1997,

[4] D. Casasent and S. Ashizawa, "SAR Detection and Recognition Filters," *Optical Eng.,* pp. 2,729, Oct. 1997.
[5] D. Dudgeon, R. Lacoss, C. Lazott, and J. Verly, "Use of Persistent Scatterers for Model-Based Recognition," *SPIE Proc. Algorithms for Synthetic Aperture Radar Imagery,* vol. 2,230, pp. 356-368, Orlando, Fla., Apr. 1994.
[6] W.E.L. Grimson, *Object Recognition by Computer: The Role of Geometric Constraints.* MIT Press, 1990.
[7] Y. Hel-Or and M. Werman, "Recognition and Localization of Articulated Objects," *Proc. IEEE Workshop Motion of Non-Rigid and Articulated Objects,* pp. 116-123, Austin, Texas, 1994.
[8] K. Ikeuchi, T. Shakunga, M. Wheeler, and T. Yamazaki, "Invariant Histograms and Deformable Template Matching for SAR Target Recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 100-105, June 1996.
[9] Y. Lamden and H. Wolfson, "Geometric Hashing: A General and Efficient Model-Based Recognition Scheme," *Proc. Int'l Conf. Computer Vision,* pp. 238-249, Dec. 1998.
[10] L. Novak, G. Owirka, and C. Netishen, "Performance of a High-Resolution Polarimetric SAR Automatic Target Recognition System," *The Lincoln Laboratory J.,* vol. 6. no. 1, pp. 11-24, 1993.
[11] J. Verly, R. Delanoy, and C. Lazott, "Principles and Evaluation of an Automatic Target Recognition System for Synthetic Aperture Radar Imagery Based on the Use of Functional Templates," *SPIE Proc. Automatic Target Recognition III,* vol. 1,960, pp. 57-71, Orlando, Fla., Apr. 1993.
[12] D. Wehner, *High Resolution Radar,* Second ed. Boston: Artech House, 1995.
[13] J.H. Yi, B. Bhanu, and M. Li, "Target Indexing in SAR Images Using Scattering Centers and the Hausdorff Distance," *Pattern Recognition Letters,* vol. 17, pp. 1,191-1,198, 1996.

Grinnell Jones III was a national merit scholar who received his BS degree in mechanical engineering from the Massachusetts Institute of Technology in 1966, his MS in aerospace engineering (with distinction) from the Air Force Institute of Technology in 1971, and his 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. 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 databases, and systems applications.

Bir Bhanu (S'72-M'82-SM'87-F'96) received the SM and EE degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge. He received his PhD degree in electrical engineering from the Image Processing Institute, University of Southern California, Los Angeles, and the MBA degree from the University of California, Irvine. He also received his BS degree (with honors) in electronics engineering from the Institute of Technology, BHU, Varanasi, India, and his ME degree (with distinction) in electronics engineering from the Birla Institute of Technology and Science, Pilani, India.

Since 1991, he has been a professor of electrical engineering and computer science, the director of the Visualization and Intelligent Systems Laboratory, and is the founding director of the Center for Research in Intelligent Systems (CRIS) at the University of California, Riverside. Previously, he was a senior Honeywell fellow at Honeywell Systems and Research Center, Minneapolis, Minnesota. He has been on the faculty of the Department of Computer Science at the University of Utah, Salt Lake City, worked for the Ford Aerospace and Communications Corporation, INRIA-France, and the IBM San Jose Research Laboratory, California. 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 coauthor of *Computational Learning for Adaptive Computer Vision* (Plenum, to be published), *Genetic Learning for Adaptive Image Segmentation* (Kluwer, 1994) and *Qualitative Motion Understanding* (Kluwer, 1992). He received the outstanding paper award from the Pattern Recognition Society. He has also received industrial awards for technical excellence, outstanding contributions, and team efforts. He has been a guest editor of several IEEE transactions and journals and has served on the editorial board of various journals. He holds 10 U.S. and international patents and more than 150 reviewed technical publications in the areas of his interest. He was the general chair for the first IEEE Workshop on Applications of Computer Vision, chair for the DARPA Image Understanding Workshop, general chair for the IEEE Conference on Computer Vision and Pattern Recognition, and program chair for the 1999 IEEE Workshop on Computer Vision Beyond the Visible Spectrum.

Dr. Bhanu is a fellow of the IEEE and the AAAS (American Association for the Advancement of Science). He is a member of the ACM, AAAI, Sigma Xi, Pattern Recognition Society, and SPIE.