Recognizing target variants and articulations in synthetic aperture radar images

Bir Bhanu, MEMBER SPIE
Grinnell Jones III
Center for Research in Intelligent Systems
University of California
Riverside, California 92521

Abstract. The focus of this paper is recognizing articulated vehicles and actual vehicle configuration variants in real synthetic aperture radar (SAR) images. Using SAR scattering-center locations and magnitudes as features, the invariance of these features is shown with articulation (e.g., rotation of a tank turret), with configuration variants, and with a small change in depression angle. This scatterer-location and magnitude quasiinvariance is used as a basis for development of a SAR recognition system that successfully identifies real articulated and nonstandard-configuration vehicles based on nonarticulated, standard recognition models. Identification performance results are presented as vote-space scatterplots and receiver operating characteristic curves for configuration variants, for articulated objects, and for a small change in depression angle with the MSTAR public data. [S0091-3286(00)03403-6]

Subject terms: automatic target recognition; model-based recognition; quasiinvariance of features; geometric and magnitude constraints; multiattribute hashing.

Paper 990245 received June 17, 1999; revised manuscript received Aug. 20, 1999; accepted for publication Aug. 20, 1999.

1 Introduction

In this paper we address the target recognition problem, starting with chips of military target vehicles from real synthetic aperture radar (SAR) images at 1-ft resolution and ending with vehicle identification. The specific challenges for the recognition system are the need for automated recognition of vehicles with articulated parts (like the turret of a tank) or that have significant external configuration variants (fuel barrels, searchlights, etc.) or that are at a small change in depression angle.

Table 1 compares the recognition results based on different approachs, all using real SAR images from the MSTAR public data.[1] This is an active area of current research, new approaches are evolving, and this comparison should be viewed as a snapshot in time. The results are presented in terms of probability of correct identification (PCI) for cases with target articulation, depression-angle change, and target configuration variants. Many of the results are for forced recognition, where the automatic target recognition (ATR) system is forced to make a target decision (i.e., there is no "unknown" class). Other results with an "unknown" or "reject" class are presented at a given probability of false alarm (Pfa) or probability of miss (Pmiss). The MSTAR predict approach[2] uses CAD object models as the basis to generate synthetic SAR signature model predictions that are matched to the real SAR data; all the other approaches build models from real SAR images. The template-matching approaches[1,3–5] use (scaled) intensity values of the target chip as features, but require either an accurate estimate of the target pose or an exhaustive search of a large database of model templates. Others[6,7] use the Radon transform as a way to reduce the number of models required to handle the SAR signature variations with target azimuth rotation. None of these approaches is specifically designed to accommodate articulated objects, and the one attempt to recognize articulated objects[5] confirms that the template-matching approach is not well suited for such objects. Our previous work in this area[8–10] was specifically designed for articulated objects; it primarily used the location invariance of SAR scattering centers and was largely based on simulated SAR data.

The approach in this paper is based on using SAR scattering-center locations and magnitudes as features that are quasiinvariant with articulation, a small change in depression angle, or target configuration variants. Our recognition system uses standard nonarticulated models of the objects (at 1-deg azimuth increments) to recognize the same objects in nonstandard and articulated configurations. The recognition process is an efficient search for positive evidence, via table lookups based on information in the test image, that generates votes for the appropriate object (and azimuth). The key contributions of this paper are that it:

1. quantifies the azimuthal variance of scattering-center locations in real SAR data
2. demonstrates that quasiinvariant scattering-center *locations* exist and that their *magnitudes* and *shape factor* are also quasiinvariant for (a) articulation, (b) configuration variants, and (c) a small depression-angle change for actual vehicles in real SAR data
3. develops a new recognition system based on scattering center location, magnitude, and shape factor that achieves significant vehicle recognition performance for articulation, configuration variants, and small depression-angle changes with real SAR data.

 0091-3286/2000/$15.00

Table 1 Comparison with related work on target recognition using real MSTAR SAR data.

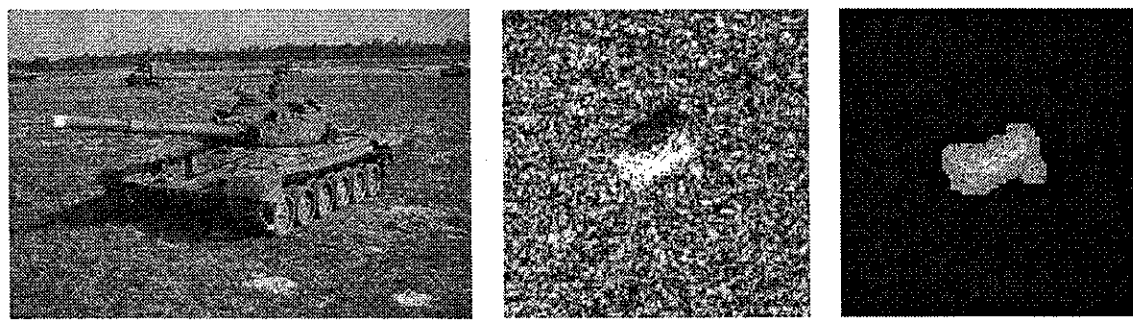
Approach	Ref.	Prob. of correct id. Art.	Depr.	Config.	Remarks
MSTAR predict models, mean-squared-error matching	2	—	—	0.74–0.78	Forced, avg. of 11 T72s; estimates
Template matching,	4	—	0.99	0.93*	Forced; *(config. and depr.)
correlation and	3	—	—	0.40–0.75	11 T72s at 0.10 Pfa
mean squared error	5	0.35	0.93	0.73	Forced recognition
	1	—	—	0.79	At 0.10 Pmiss
Radon transform, neural net	7	—	0.93	—	Forced recognition
Radon transform, hidden Markov models	6	—	0.94	—	Forced recognition
This paper		1.00	0.99	0.95	Forced recognition
		0.97	0.82	0.67–0.85	At 0.10 Pfa

2 Invariances of SAR Scattering Centers

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 1-ft resolution; however, there is a wealth of peaks corresponding to scattering centers. The relative locations of SAR scattering centers, determined from local peaks in the radar return, are related to the aspect and physical geometry of the object, are independent of translation, and serve as distinguishing features. Target regions of interest (ROIs) are found in the MSTAR SAR chips by reducing speckle noise using the Crimmins algorithm in Khoros,[11] thresholding at the mean plus two standard deviations, dilating to fill small gaps among regions, eroding to have one large ROI and small 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. An example photograph, SAR target chip image, and ROI (with scattering-center locations shown as black dots) are given in Fig. 1 for T72 tank (serial number) #a64. The parameters used in extracting ROIs are held constant for all the results reported. In addition to the scatterer locations, the magnitudes and shapes of the peaks are also features that we use in this paper.

2.1 *Azimuthal Variance of Scatterer Locations*

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. Since the radar depression angle is generally known, the significant unknown target rotation is (360 deg) in azimuth. Azimuth persistence or invariance can be expressed in terms of the percentage of 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 at azimuth θ_0 by an angle $\Delta\theta$ and comparing the resulting range and cross-range locations with the scatterer locations from an image of the same object at azimuth $\theta_0+\Delta\theta$. More precisely, because the images are in the radar slant plane, we actually project from the slant plane to the ground plane, rotate in the ground plane, and



(a) (b) (c)

Fig. 1 Example (a) photograph, (b) SAR image, and (c) ROI (with peaks) for T72 tank #a64 (photograph not to scale).

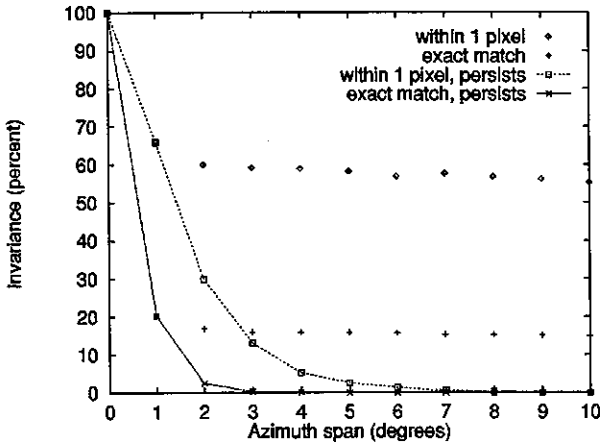


Fig. 2 Scatterer location persistence, T72 #132.

project back to the slant plane. Since the objects in the chips are not registered, we calculate the azimuth invariance as the maximum number of corresponding scattering centers (whose locations match within a given tolerance) for the optimum integer pixel translation. This method of registration by finding the translation that yields the maximum number of correspondences has the limitation that for very small or no actual invariance it may find some false correspondences and report a slightly higher invariance than in fact exists. 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 k'th 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 persistent-scattering-center locations over 360 base azimuths gives the mean azimuth invariance of the object.

Figure 2 shows an example of the mean scatterer-location invariance (for the 40 strongest scatterers) as a function of azimuth angle span using T72 tank #132, with various definitions of persistence. In the "exact match" cases the center 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 $\theta_0+\Delta\theta$. In the "within 1 pixel" cases the scatterer location is allowed to move into one of the eight adjacent pixel locations. Note that for a 1-deg azimuth span, while only 20% of the scatterer locations are invariant for an "exact match," 65% of the scatterer locations are invariant "within 1 pixel." The cases labeled "persists" in Fig. 2 enforce the constraint that the scatterer exist for the entire span of angles; in practice, very few scatterers continuously persist for even 5 deg. In the upper two cases (not labeled "persists") scintillation is allowed and the location invariance declines slowly with azimuth span. The "within 1 pixel" results (that allow scintillation) are consistent with the 1-ft ISAR results of Dudgeon et al.,[12] whose definition of persistence allowed scintillation. Because of the higher scatterer-location invariance with 1-deg azimuth span, in this research we use azimuth models at 1-deg increments for each target, in contrast to others who have used 5-deg,[13] 10-deg,[14] and 12-deg models.[15]

2.2 *Scatterer-Location Invariance*

Many of the scatterer locations are invariant to target conditions such as articulation or configuration variants or to a small change in depression angle. Because the object and ROI are not registered, we express the scattering-center location invariance with respect to articulation, configuration differences, or depression-angle changes as the maximum number of corresponding scattering centers (whose locations match within a stated tolerance) for the optimum integer pixel translation.

Given an original version of a SAR target image with n scattering centers, represented by points at pixel locations $P_i=(x_i,y_i)$ for $1 \leq i \leq n$ and a translated, distorted version $P'_j=(x'_j,y'_j)$ $(1 \leq j \leq n)$ at a translation $t=(t_x,t_y)$, we define a *match* between points P'_j and P_i as

$$M_{ij}(t)=\begin{cases}1 & \text{if } |x'_j-t_x-x_i| \leq l \text{ and } |y'_j-t_y-y_i| \leq l,\\ 0 & \text{otherwise,}\end{cases}$$

where $l=0$ for an "exact" match and $l=1$ for a match "within one pixel." The scatterer location invariance, L_n, of n scatterers, expressed as a percentage of matching points, is given by

$$L_n=\max_t \frac{100}{n}\sum_{j=1}^{n} \min\left(\sum_{i=1}^{n} M_{ij}(t),1\right),$$

where each point P'_j is restricted to at most one match.

Figure 3 shows the location invariance, L_{40}, of the strongest 40 scattering centers with articulation for MSTAR T72 tank #a64 and ZSU 23/4 antiaircraft gun #d08 (at a 30-deg depression angle) as a function of the hull azimuth. The combined average invariance for both vehicles is 16.5% for an exact match of scattering centers and 56.5% for a location match within one pixel (3×3 neighborhood) tolerance. Similarly, Fig. 4 shows the percentage of the strongest 40 scattering-center locations that are invariant for configuration variants, T72 #812 versus #132 and BMP2 vehicle #C21 versus #9563, at a 15-deg depression angle. Figure 5 shows the percentage of scattering-center location invariance for T72 #132 and BMP2 #C21 at 17- versus 15-deg depression angles. Note that the mean scatterer-location invariance for T72 tank #132 with a 2-deg change in depression angle shown in Fig. 5(a), 17.8% for exact match and 61.6% within 1 pixel, is similar to the results with the identical tank shown in Fig. 2 for a 2-deg azimuth span (for the appropriate cases where scintillation is allowed). The mean and standard deviation for percentage of location invariance (for 40 scatterers and depression angle ϕ) are shown in Table 2 for articulated versions of the T72 and ZSU23/4, for configuration variants of the T72 and BMP2, and for depression-angle changes with the T72 and BMP2.

2.3 *Scatterer Magnitude Invariance*

Using a scaled scatterer amplitude (S), expressed as a radar cross section in square meters, given by $S=100+10\log_{10}(i^2+q^2)$, where i and q are the components of the complex radar return, we define a percentage amplitude

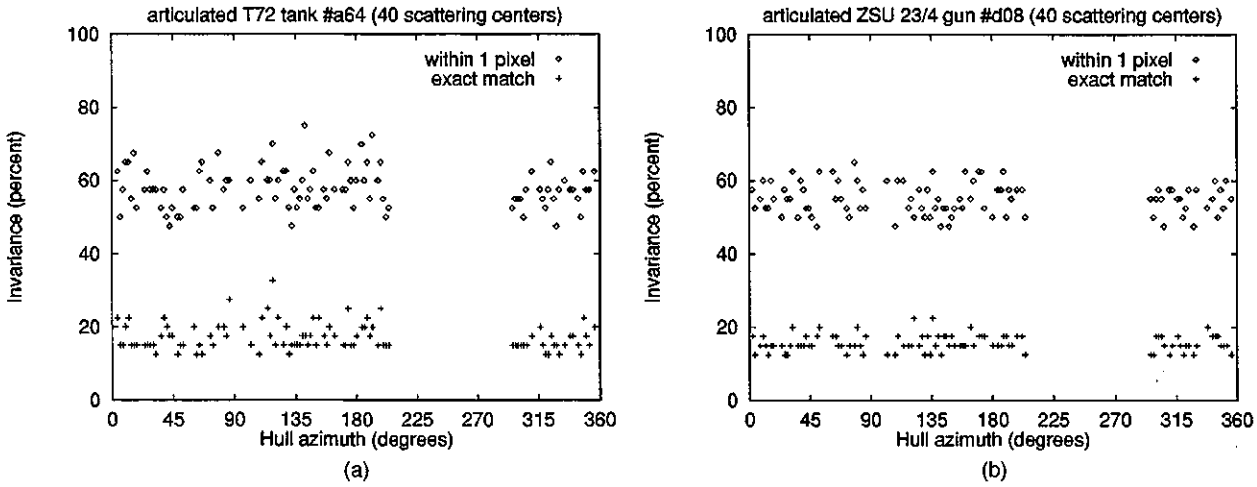


Fig. 3 Scatterer location invariance with articulation: (a) T72 tank, (b) ZSU 23/4.

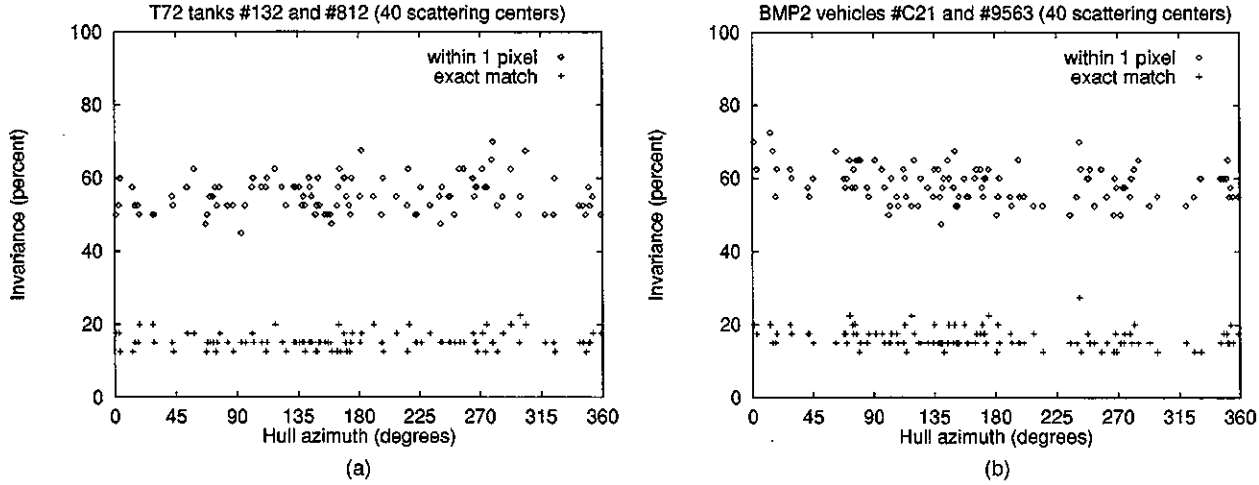


Fig. 4 Scatterer location invariance with configuration: (a) T72 tank, (b) BMP2.

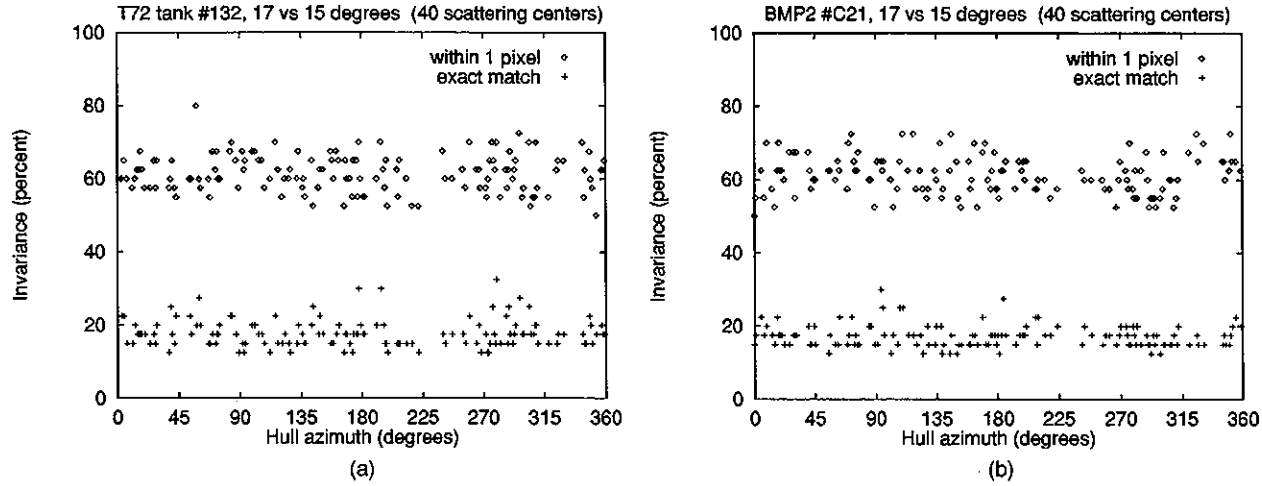


Fig. 5 Scatterer location invariance with depression angle: (a) T72 tank, (b) BMP2.

Table 2 Scatterer percentage location invariance for MSTAR targets with articulation, configuration variants, and depression-angle changes.

Change in	Depression angle (deg)	"Exact match" invariance		"Within 1 pixel" invariance	
		Mean	S.d.	Mean	S.d.
Articulation:					
T72 #a64	30	17.17	1.47	57.83	2.23
ZSU #d08	30	15.69	0.91	55.05	1.72
Average		16.45		56.47	
Configuration variants:					
T72: #812 vs #132	15	15.34	0.89	55.34	1.91
#s7 vs #132	15	15.40	0.83	56.68	1.95
BMP2: #9563 vs #c21	15	16.34	0.84	58.52	1.97
#9566 vs #c21	15	16.17	0.99	57.93	1.97
Average		15.83		57.15	
Depression angle:					
T72 #132	17–15	17.76	1.52	61.55	2.05
BMP2 #c21	17–15	17.19	1.23	61.31	2.11
Average		17.47		61.43	

change A_{jk} as $A_{jk}=100(S_j-S_k)/S_j$. (This form allows a larger variation for the stronger signal returns.) A location-and-magnitude match $Q_{jk}(t)$ is given by

$$Q_{jk}(t)=\begin{cases}1 & \text{if } M_{jk}(t)=1 \text{ and } |A_{jk}|\leq l_A,\\ 0 & \text{otherwise,}\end{cases}$$

where l_A is the percentage amplitude-change tolerance. The scatterer magnitude-and-location invariance I_n, expressed as a percentage of n scatterers, is given by

$$I_n=\max_t \frac{100}{n}\sum_{k=1}^{n}\min\left(\sum_{j=1}^{n} Q_{jk}(t),1\right).$$

Figure 6 shows the probability mass functions (PMFs) for percentage amplitude change for the strongest 40 articulated versus nonarticulated scattering centers of T72 tank #a64 and ZSU 23/4 gun #d08. Curves are shown both for the cases where the scattering-center locations correspond within one pixel tolerance and for all the combinations of scatterers whose locations do not match. Similarly, Fig. 7 shows the PMFs for percentage amplitude change for the strongest 40 scattering centers with configuration variants, T72 #812 versus #132 and BMP2 #c21 versus #9563, at a 15-deg depression angle. In addition, Fig. 8 shows this for 17- versus 15-deg depression angles for T72 #132 and BMP2 #c21. The mean and standard deviation for these matching and nonmatching scatterers and the crossover points for the PMFs are given in Table 3. Table 4 shows the mean and standard deviation for the percentage location and magnitude invariance (within a 1-pixel location toler-

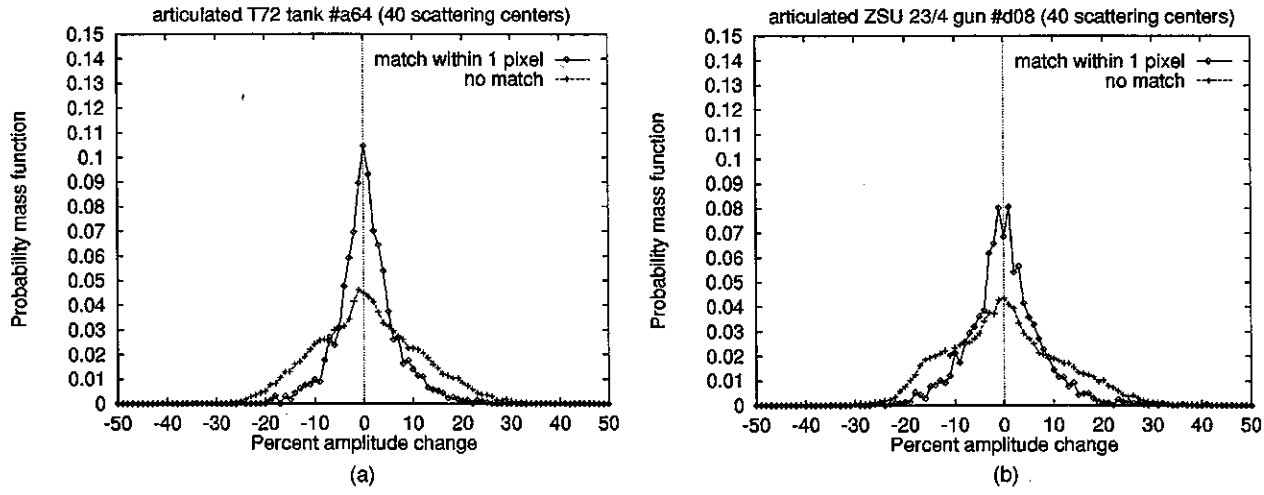


Fig. 6 Scatterer magnitude invariance with articulation: (a) T72 tank, (b) ZSU 23/4.

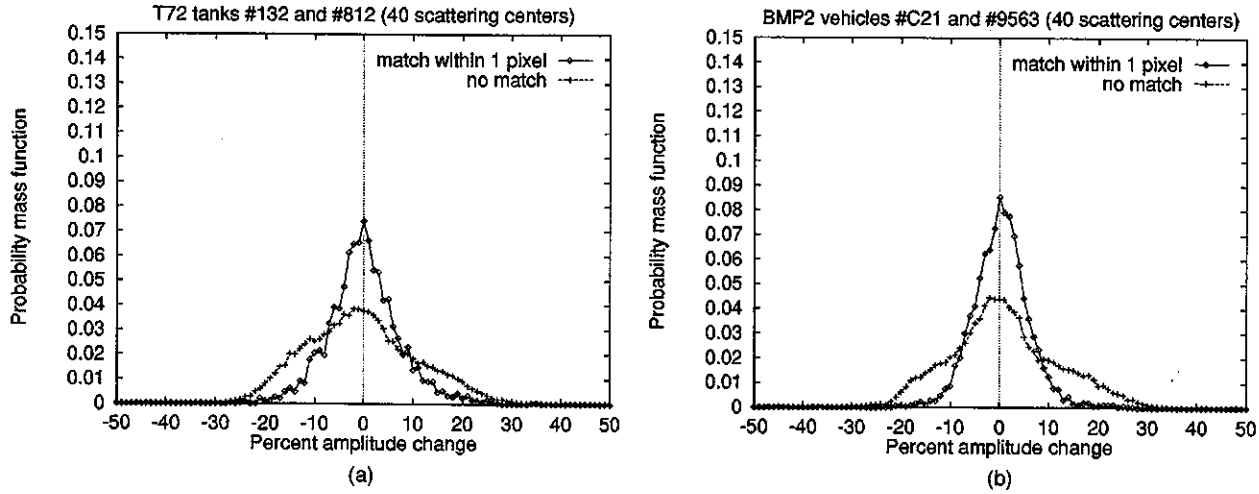


Fig. 7 Scatterer magnitude invariance with configuration: (a) T72 tank, (b) ZSU 23/4.

ance and an amplitude-change tolerance of l_A) of the strongest 40 scatterers for these same articulation, configuration-difference, and depression-angle change cases.

2.4 *Peak Shape Factor*

A *shape factor* is used as a measure of the sharpness of the local peak in the radar return associated with a scattering center. We define the shape factor $F = S_k / \Sigma_{i=1}^{8} S_i$, where S_k is the amplitude of the peak and the S_i's are the amplitudes of the eight neighbors. Figure 9 shows the PMFs for percentage shape-factor change for the strongest 40 scattering centers of T72 #812 versus #132 (at 15-deg depression angle). Curves are shown both for cases where the scattering-center locations correspond within one pixel tolerance and for all the combinations of scatterers whose locations do not match. For the cases with locations that match within one pixel, the mean and standard deviation of the percentage shape-factor change are 1.3 and 15.7, while for the nonmatching cases they are 5.3 and 31.3, respectively.

3 Invariant-Based SAR Recognition System

Our invariant-based recognition system uses standard non-articulated models of the objects (at 1-deg azimuth increments) to recognize the same objects in nonstandard and articulated configurations. Using a technique like geometric hashing,[16] the relative positions of the scattering centers in the range and cross-range directions are indices to a lookup table of labels that give the associated target type and pose. This is an efficient search for *positive* evidence that generates votes for the appropriate object (and azimuth). The models and recognition system have evolved from the earlier 2-D version,[8–10] using only the relative distances and the "exact" scatterer locations, to the current 6-D and 8-D versions that use more local features and accommodate a "within 1 pixel" scatterer location uncertainty. In contrast to many model-based approaches,[17] we are not *searching* all the models; instead we are doing table lookups based on relative distances between the strongest scatterers in the test image. We use a local coordinate system where the origin is the scatterer used as the basis for computing the relative

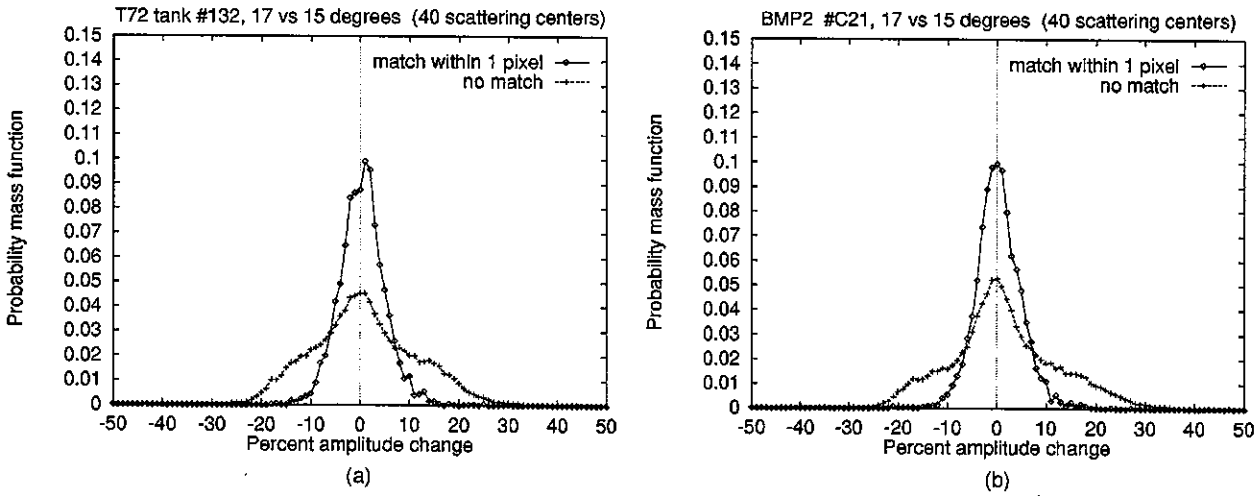


Fig. 8 Scatterer magnitude invariance with depression angle: (a) T72 tank, (b) BMP2.

Table 3 Scatterer percent amplitude change.

Change in	Within 1 pixel Mean	Within 1 pixel S.d.	No match Mean	No match S.d.	Cross-over
Articulation:					
T72 #a64	0.51	5.91	0.75	10.44	−5/+6
ZSU #d08	0.06	7.44	0.08	11.37	±9
Configuration variants:					
T72: #812 vs #132	0.15	7.29	−0.38	11.12	±8
#s7 vs #132	0.48	6.69	2.20	11.15	±9
BMP2: #9563 vs #c21	0.35	5.72	0.94	10.88	−8/+9
#9566 vs #c21	0.48	6.20	0.56	10.68	−7/+8
Depression angle:					
T72 #132	0.43	4.66	0.84	10.45	−7/+8
BMP2 #c21	0.37	4.65	1.53	10.91	−7/+8

locations of the other scatterers. For ideal data one could use the strongest scatterer as the origin; however, any given scatterer could actually be spurious or missing due to the effects of noise, articulation, occlusion, or nonstandard configurations. Thus, we model and use all the scattering-center locations in turn as the origin, so the size of the lookup-table models and the number of nominal relative distances considered in a test image are each $n(n-1)/2$, where n is the number of the strongest scattering centers used.

The off-line model construction algorithm extracts these relative distances of the scattering centers from sets of training-data target chips at 1-deg azimuth increments for each target type. The relative distances are the indices to a lookup table, and in the 2-D version each entry in the table is a list of labels that give the appropriate object type and azimuth. In the 6-D version the model lookup table labels contain four additional features: the range and cross-range

Table 4 Scatterer percentage location and magnitude invariance (for locations within one pixel and amplitude tolerance I_A).

	I_A	Mean	S.d.
Articulation:			
T72 #a64	±9	53.47	2.63
ZSU #d08	±9	47.98	2.22
Average		50.78	
Configuration variants:			
T72: #812 vs #132	±9	48.40	2.42
#s7 vs #132	±9	50.69	2.44
BMP2: #9563 vs #c21	±9	54.38	2.34
#9566 vs #c21	±9	53.00	2.51
Average		51.68	
Depression angle:			
T72 #132	±7	56.15	2.38
BMP2 #c21	±7	55.66	2.53
Average		55.91	

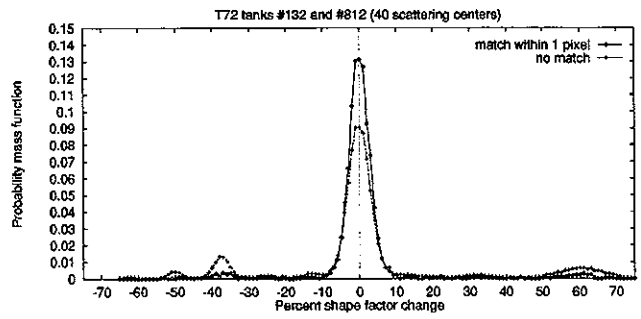


Fig. 9 Shape-factor change with configuration.

position of the origin, and the magnitudes of the two scatterers. The 8-D version adds the shape factors of the two scatterers.

Similarly, the on-line recognition algorithm extracts these relative distances of the scattering centers from the test-data target chips and uses the relative distances as indices to access the lookup table. In the 2-D version of the recognition algorithm each query of the lookup table may directly generate votes for one or more potential candidate solutions. The 2-D version accumulates votes in a 2-D object-azimuth space. The process is repeated with different scattering centers as the origin, providing multiple "looks" at the model database to handle spurious scatterers that arise from articulation, configuration differences, or noise.

The 6-D and 8-D versions of the recognition algorithm are extensions of the basic 2-D algorithm that add additional features as constraints, accommodate a "within 1 pixel" scatterer location uncertainty, and weight the votes for the more uncommon longer distances. In the 6-D version, the comparison of the test data pair of scatterers with the model lookup table result(s) also provides information on the range and cross-range translation and on the percentage magnitude changes for the two scatterers, and the 8-D system provides additional information on the percentage change in the shape factor of the two scatterers. Limits on allowable values for translations, magnitude, and shape-factor changes are used as constraints to reduce the number of false matches. (The number of scattering centers used and the various constraint limits are design parameters that are optimized, based on experiments, to produce the best recognition results.) Votes are accumulated in a 4-D space: object, azimuth, range, and cross-range translation. A (city-block) weighted voting method is used to reduce the influence of the more common small relative distances. To accommodate some uncertainty in the scattering-center locations, the eight neighbors of each nominal range and cross-range relative location are also probed, and the translation results are accumulated for a 3×3 neighborhood in the translation subspace.

To handle identification with "unknown" objects, we introduce a criterion for the quality of the recognition result (e.g., the votes for the potential winning object exceed some threshold $v_{\min}$). By varying the decision-rule parameter we obtain a form of receiver operating characteristic (ROC) curve of probability of correct identification (PCI) versus probability of false alarm (Pfa).

Table 5 Forced-recognition confusion matrix for MSTAR configuration variants (36 scatterers, ±9% amplitude tolerance).

MSTAR (public) test targets [serial number]		Identification results (configuration modeled)			
		BMP2	[#c21]	T72	[#132]
BMP2	[#9563]	106	(98.1%)	2	
	[#9566]	107	(97.2%)	3	
T72	[#812]	11		92	(89.3%)
	[#S7]	6		88	(93.6%)

4 Recognition Results

4.1 *Configuration Experiments*

In the configuration-variant experiments a single configuration of the T72 (#132) and one of the BMP2 (#C21) vehicle were used as the models, and the test data were two other variants of each vehicle type (T72 #812, #s7 and BMP2 #9563, #9566) and different "unknown" confuser test vehicles (all at 15-deg depression angle). Although more extensive T72 configuration variant data are available, only two configurations were used so that the numbers of test data for the T72 and BMP2 were comparable and the results were not artificially biased toward recognizing the T72. The forced-recognition confusion matrix for these configuration variants is shown in Table 5, with an overall recognition rate of 94.7%. (This 94.7% rate with the 6-D recognition system is directly comparable with the 68.4% rate for the prior 2-D version of the recognition system given in Ref. 9.) These results were obtained with the 6-D system using 36 scattering centers, a translation limit of ±5 pixels, and a percentage magnitude change of less than ±9%. These parameter settings were optimum for the configuration-variant experiments (the most difficult case), and the same settings were also used with the 6-D system in the results on articulation and depression-angle change given in subsequent sections. The effect on the forced-recognition PCI of the number of scattering centers used is shown in Fig. 10, and Fig. 11 shows the effect of varying the amplitude change limit.

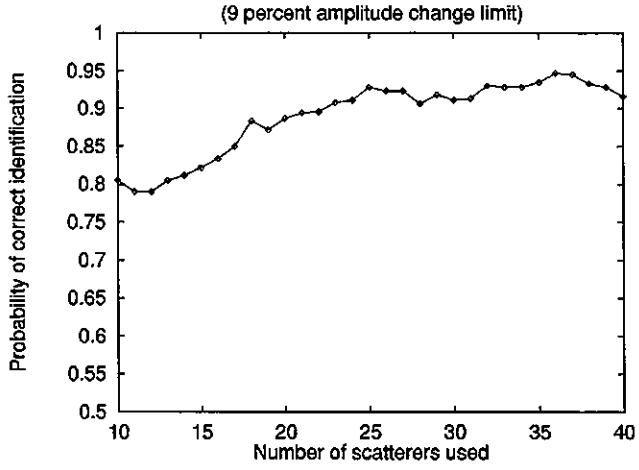


Fig. 10 Effect of number of scattering centers used on recognition of MSTAR configuration differences.

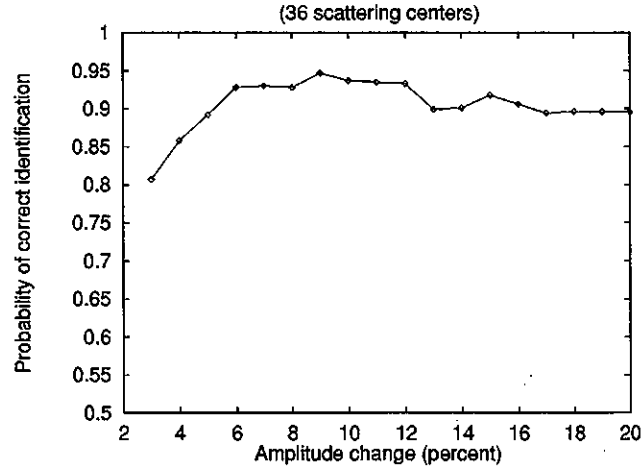


Fig. 11 Effect of amplitude-change tolerance on recognition of MSTAR configuration differences.

Figures 12(a)–12(d) show scatterplot recognition results in BMP2–T72 vote space for configuration variants of the BMP2 and T72 and for various confusers: BTR70 #c71, ZSU 23/4 #d08, and BRDM2 #e71. The 45-deg line in Fig. 12 represents the decision boundary of the simplest decision rule: "The object with the most votes wins." In this forced-recognition case, Fig. 12(a), the overall recognition rate is 94.7%, where 2.3% of the BMP2s and 9.4% of the T72s are on the wrong side of the boundary and are misidentified. In Figs. 12(b)–12(d) the BTR is the most difficult confuser, the BRDM is somewhat less difficult, and the ZSU is easy. For example, Fig. 12(b) shows that 99.6% of the BTR70 confuser false alarms could be eliminated with a 3000-vote threshold, but Fig. 12(a) shows that a 3000-vote threshold would eliminate more than half of the BMP2 and T72 identifications. In contrast, Fig. 12(d) shows that almost all of the ZSU confuser false alarms could be eliminated with a 1000-vote threshold without any reduction in the BMP2 and T72 identifications.

4.2 *Articulation Experiments*

In the articulation experiments, the models are nonarticulated versions of T72 #a64 and ZSU23/4 #d08, and the test data are the articulated versions of these same-serial-number objects and BRDM2 #e71 as a confuser vehicle (all at 30-deg depression angle). The articulated-object recognition results are shown in Table 6 using the 6-D system with a 2100-common-vote decision criterion for an overall 0.927 PCI at a 0.039 Pfa. (The overall forced-recognition rate is 100% over a range from 14 to 40 scattering centers.) Figures 13(a) and 13(b) show 6-D system scatterplot results in ZSU-T72 vote space for articulation of the ZSU 23/4 and T72 tank and for the BRDM2 confuser. Here the results for the ZSU 23/4 and T72 are widely separated, giving 100% forced-recognition results. Figure 13(b) shows that while the BRDM2 is always classified as a T72, a unique threshold of 2000 to 2500 T72 votes will eliminate most if not all of the false alarms at the cost of only a few T72s moved to the "unknown" classification. A common threshold ap-

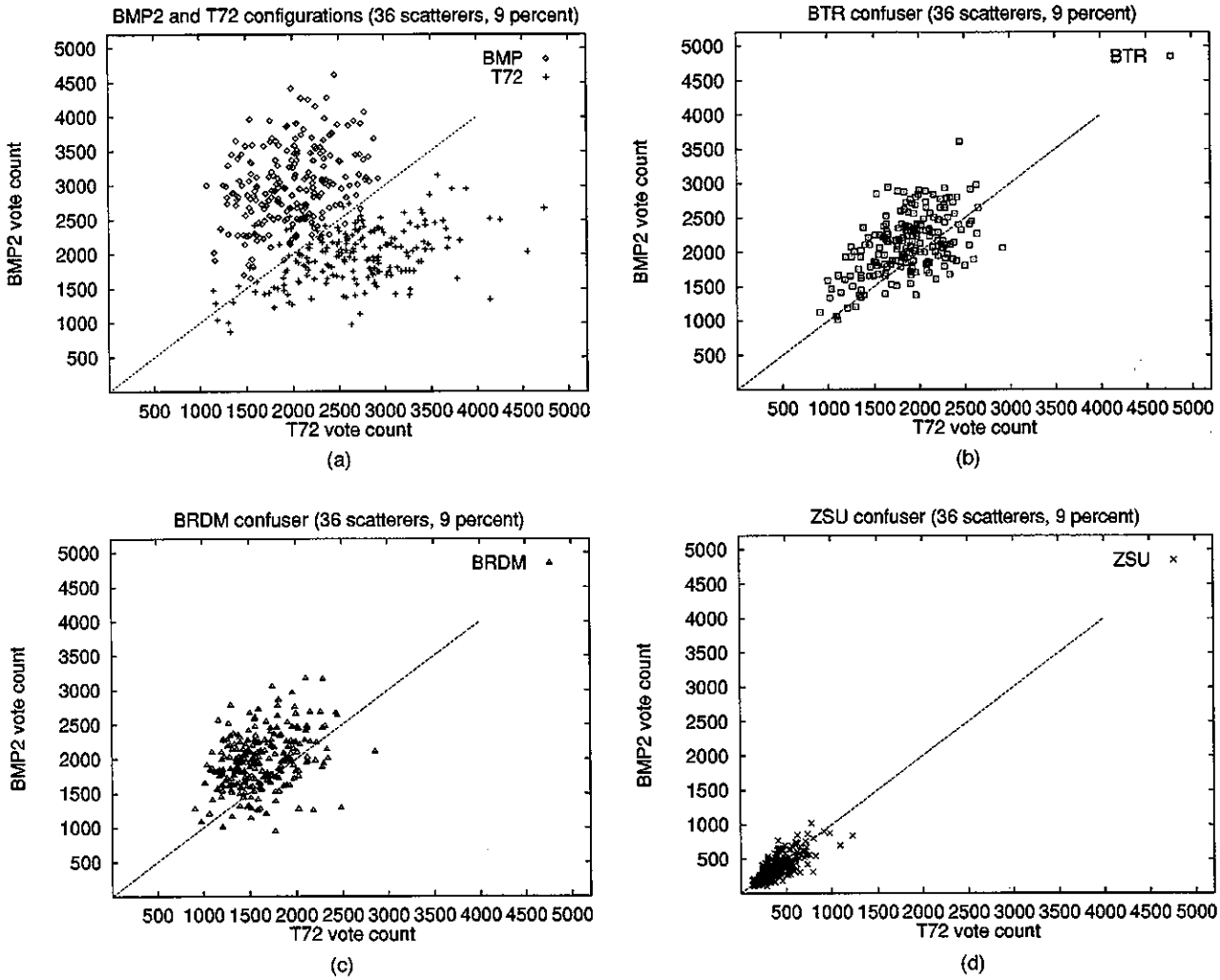


Fig. 12 Scatterplots for 6-D system results with configuration variants: (a) BMP versus T72, (b) BTR confuser, (c) BRDM confuser, (d) ZSU confuser.

plied to votes for either the T72 or the ZSU has a higher cost, because many ZSUs are moved to "unknown," as shown in Table 6.

4.3 *Depression-Angle Experiments*

In the depression-angle experiments the models are T72 #132 and BMP2 #c21 at a 15-deg depression angle, and the test data are the same-serial-number objects and the BTR70 #c71 confuser at 17 deg. The confusion matrix, shown in Table 7, for the depression-angle results has an overall 0.822 PCI at 0.10 Pfa, obtained with the 6-D system and a 2800-vote threshold. Figures 14(a) and 14(b) show scatterplot recognition results in BMP2–T72 vote space for the depression-angle experiments. These results show better separation than the configuration-variant results of Fig. 12(a) with a 99.3% forced-recognition rate. The vote counts for depression-angle change in Fig. 14(a) are typically higher (more away from the origin and away from the decision boundary) than in Fig. 12(a) for configuration variants, while the BTR confuser plots are generally similar. Thus, a common vote threshold for the depression-angle cases eliminates false alarms at a lower cost than for the configuration-variant cases.

Table 6 Example MSTAR articulated-object confusion matrix (36 scatterers, ±9% amplitude tolerance, 2100-vote threshold).

MSTAR (public) articulated test targets		Identification results		
		T72	ZSU	Unknown
T72	(315-deg turret)	94	0	4
ZSU	(315-deg turret)	0	84	10
BRDM2	(confuser)	10	0	248

4.4 *ROC-Curve Results*

ROC curves can be generated from the scatterplot data in Figs. 12–14 by varying the vote threshold (typically from 1000 to 4000 in 50-vote increments). Figure 15 shows the significant effect on the configuration-variant recognition ROC curves of using the different ZSU, BRDM, and BTR confusers whose scatterplot results were given in Fig. 12. Excellent results are obtained with the ZSU 23/4 confuser, while the BTR70 is a difficult case. Figure 16 shows the

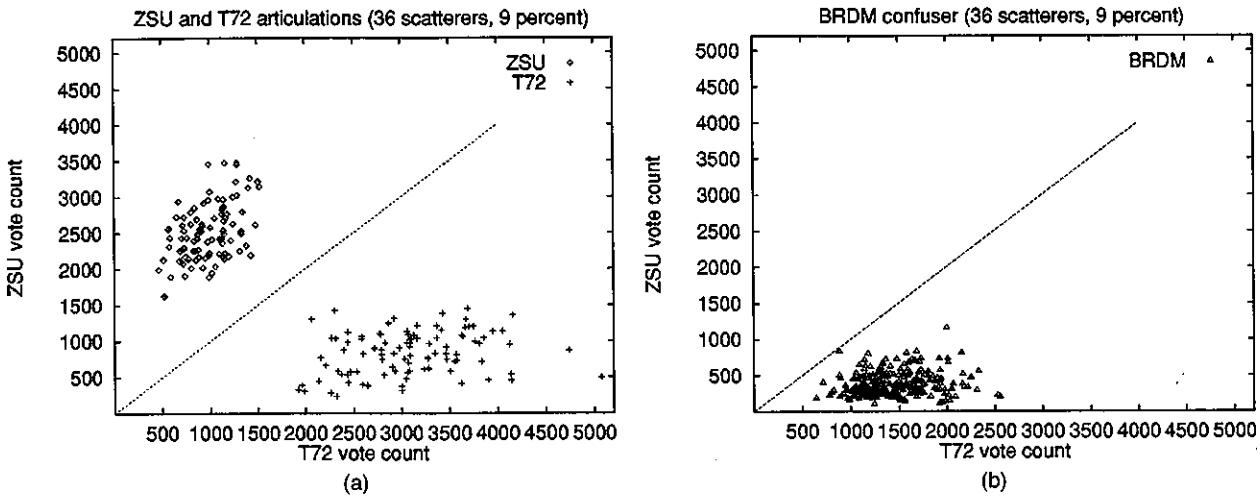


Fig. 13 Scatterplots for 6-D system results with articulation: (a) ZSU versus T72, (b) BRDM confuser.

ROC-curve recognition results for the articulation, depression-angle change, and configuration-variants cases, all with the 6-D system using the same operating parameters. The ROC curves in Fig. 16 show that the differences in configuration of an object type are a more difficult challenge for the recognition system than small depression-angle changes, since both were generated using the BTR confuser. The excellent results for the articulation case are basically due to the dissimilarity of the ZSU, T72, and BRDM as seen in Fig. 13.

Figure 17 shows ROC curves for the same MSTAR T72 and BMP2 configuration variants with the BTR confuser using the 2-D, 6-D, and 8-D recognition systems. Each of the systems was optimized for the forced-recognition configuration-variant case: the 2-D system at 20 scatterers; the 6-D system at 36 scatterers; the 8-D system at 50 scatterers (with a $\pm 30\%$ shape-factor change limit). Both the 6-D and 8-D system results are a substantial improvement over the earlier 2-D system results. While Fig. 17 shows that the 8-D system gave worse results than the 6-D system in the region below 0.1 Pfa, reoptimizing the operating parameters (e.g., using 45 scatterers) gives the 8-D system better results in the region below 0.1 Pfa at the cost of a slightly reduced forced-recognition rate.

The ROC curves in Figs. 15–17 were based on simple vote thresholds and did not utilize any knowledge about the characteristics of the confusers shown in Figs. 12(b)–12(d), 13(b), and 14(b). Figure 18 shows the effect of various decision rules on the ROC curves for recognition of configuration variants with the 8-D system and the BTR confuser. The lower curve in Fig. 18 is the original simple vote-threshold case (the same as in Fig. 17 for the 8-D system). The upper curve, with almost perfect results, is an upper bound using the tightest decision-rule boundary that can be drawn around the BTR confuser results in BMP-T72 vote space. The middle curve is an example decision-rule boundary based on a stepwise continuous envelope around the confuser results. Thus, Fig. 18 illustrates the potential benefits available from utilizing knowledge of the confuser characteristics in optimizing decision rules for improved performance.

Table 7 Example confusion matrix for MSTAR depression-angle changes (36 scatterers, $\pm 9\%$ amplitude tolerance, 2800-vote threshold).

MSTAR (public) depression-angle 17-deg test targets		Identification results (15-deg models)		
		BMP2	T72	Unknown
BMP2	[#c21]	110	0	28
T72	[#132]	0	117	21
BTR70	(confuser)	15	8	207

5 Conclusions

While less than 20% of the SAR scattering-center locations exactly match under object articulation, configuration differences, and small depression-angle changes, a significant percentage (56.4% to 61.4%) of these locations are quasi-invariant within a 3×3 pixel tolerance. The magnitudes of these quasiinvariant scatterers (expressed as a radar cross-section) typically change by less than $\pm 10\%$. The positions and magnitudes of pairs of these quasiinvariant scatterers (and the peak shape factors) can be used in an efficient recognition system to achieve good recognition results with real SAR data for object articulation, configuration differences, and small depression-angle changes. While these three problems are similar, the differences among configurations of an object type are a more significant challenge for recognition than articulation and depression-angle changes (where the model and test data are the same physical object under different conditions). The confuser vehicle used has a significant effect on the ROC curve results; however, knowledge about the characteristics of confuser vehicles can be used to optimize decision rules for improved performance. These recognition results with real SAR data are a substantial improvement over the performance of the earlier recognition approach[9,10] that required an exact match of scatterer locations and only used the relative distance information. Future work to explicitly incorporate additional information on the unique-

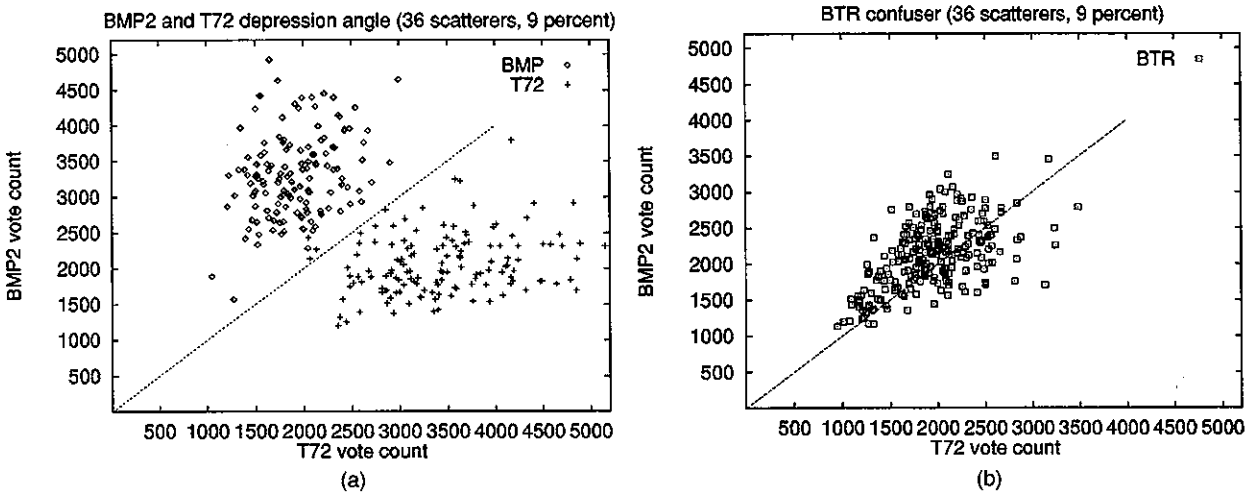


Fig. 14 Scatterplots for 6-D system results with depression-angle change: (a) BMP versus T72, (b) BTR confuser.

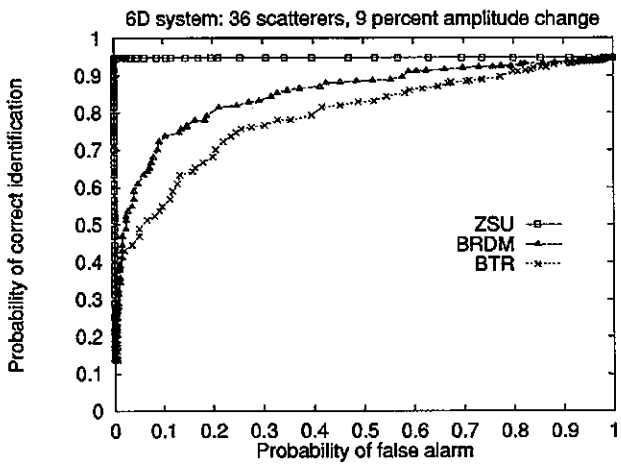


Fig. 15 Effect of confusers on configuration-variant ROC curve.

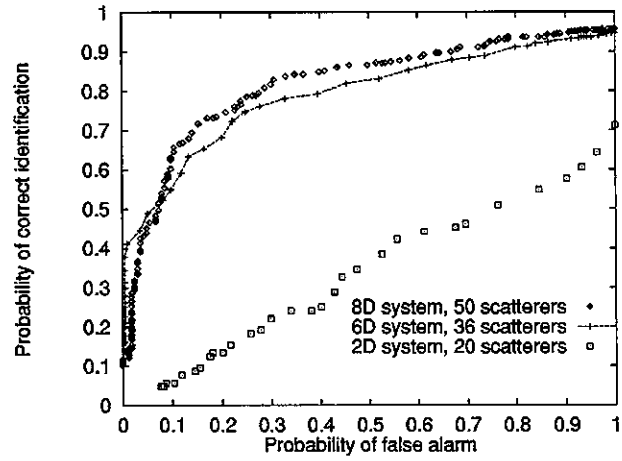


Fig. 17 ROC curves for configuration variants with 2-D, 6-D, and 8-D systems.

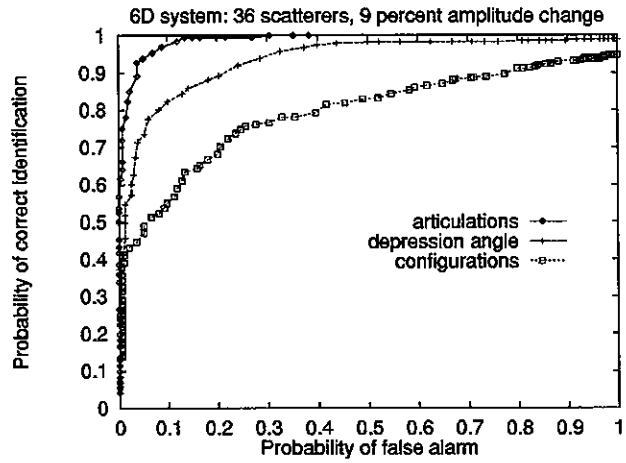


Fig. 16 ROC curves for articulation, depression-angle, and configuration variants.

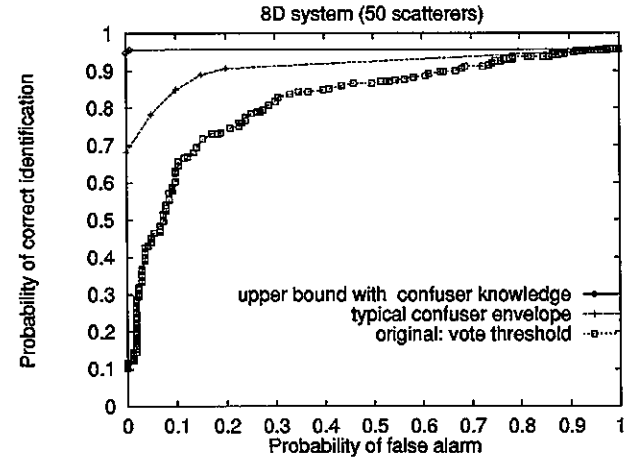


Fig. 18 Effect of decision rules on configuration ROC curves for 8-D system.

ness and persistence of these features as well as additional features should lead to further performance improvements and accommodate combined cases such as configuration-angle variants along with depression-angle changes. In addition, it is possible to optimize the parameters of the SAR recognition system using learning techniques.[18]

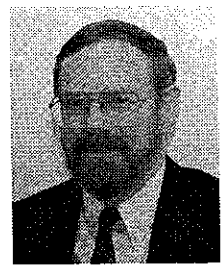
Acknowledgments

This work was supported by DARPA/AFOSR grant F49620-97-1-0184. The contents and information do not necessarily reflect the position or policy of the U.S. Government.

References

1. T. Ross, S. Worell, V. Velten, J. Mossing, and M. Bryant, "Standard SAR ATR evaluation experiments using the MSTAR public release data set," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 566–573 (1998).
2. S. Stanhope, E. Keydel, W. Williams, V. Rajlich, and R. Sieron, "The use of the mean squared error matching metric in a model-based automatic target recognition system," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 360–368 (1998).
3. R. Bhatnagar, R. Dilsavor, M. Minardi, and D. Pitts, "Intra-class variability in ATR systems," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 383–395 (1998).
4. L. Kaplan, R. Murenzi, E. Asika, and K. Namuduri, "Effect of signal-to-clutter ratio on template-based ATR," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 408–419 (1998).
5. J. Mossing and T. Ross "An evaluation of SAR ATR algorithm performance sensitivity to MSTAR extended operating conditions," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 554–565 (1998).
6. D. Kottke, P. Fiore, K. Brown, and J. Fwu, "A design for HMM-based SAR ATR," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 541–551 (1998).
7. Q. Pham, T. Brosnan, M. Smith, and R. Mersereau "An efficient end-to-end feature based system for SAR ATR," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 519–529 (1998).
8. B. Bhanu and G. Jones III, "Performance characterization of a model-based SAR target recognition system using invariants," in *Algorithms for Synthetic Aperture Radar Imagery IV, Proc. SPIE* **3070**, 305–321 (1997).
9. B. Bhanu, G. Jones III, and J. Ahn "Recognizing articulated objects and object articulation in SAR images," in *Algorithms for Synthetic Aperture Radar Imagery V, Proc. SPIE* **3370**, 493–505 (1998).
10. G. Jones III and B. Bhanu, "Recognition of articulated and occluded objects," *IEEE Trans. Pattern. Anal. Mach. Intell.* **21**(7), 603–613 (1999).
11. *Khoros Pro v2.2 User's Guide*, Addison Wesley Longman Inc. (1998).
12. D. Dudgeon, R. Lacoss, C. Lazott, and J. Verly, "Use of persistent scatterers for model-based recognition," in *Algorithms for Synthetic Aperture Radar Imagery, Proc. SPIE* **2230**, 356–368 (1994).
13. L. Novak, S. Halversen, G. Owirka, and M. Hiett, "Effects of polarization and resolution on SAR ATR," *IEEE Trans. Aerosp. Electron. Syst.* **33(1)**, 102–115 (1997).
14. K. Ikeuchi, T. Shakunga, M. Wheeler, and T. Yamazaki, "Invariant histograms and deformable template matching for SAR target recognition," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 100–105 (1996).
15. 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," in *Automatic Target Recognition III, Proc. SPIE* **1960**, 57–71 (1993).
16. Y. Lamden and H. Wolfson. "Geometric hashing: a general and efficient model-based recognition scheme," in *Proc. IEEE Computer Soc. 2nd Int. Conf. on Computer Vision*, pp. 238–249 (1988).
17. W. E. L. Grimson, *Object Recognition by Computer: The Role of Geometric Constraints*, The MIT Press (1990).
18. B. Bhanu, Y. Lin, G. Jones III, and J. Peng, "Adaptive target recognition using reinforcement learning," in *Proc. DARPA Image Understanding Workshop*, pp. 1177–1180, Monterey, CA (1998).

Bir Bhanu received the SM and EE degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge; the 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 the BS degree (with honors) in electronics engineering from the Institute of Technology, BHU, Varanasi, India, and the ME degree (with distinction) in electronics engineering from Birla Institute of Technology and Science, Pilani, India. Since 1991, he has been a professor of electrical engineering and computer science and director of the Visualization and Intelligent Systems Laboratory, and recently he has been 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, MN. He has been on the faculty of the Department of Computer Science at the University of Utah, Salt Lake City, and has worked at Ford Aerospace and Communications Corporation, INRIA-France, and IBM San Jose Research Laboratory. 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 books on *Computational Learning for Adaptive Computer Vision* (Kluwer, forthcoming), *Genetic Learning for Adaptive Image Segmentation* (Kluwer, 1994), and *Qualitative Motion Understanding* (Kluwer, 1992). He received an 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 the guest editor of several IEEE transactions and journals, and on the editorial board of various journals. He holds 10 U.S. and international patents and has produced over 200 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 ACM, AAAI, Sigma Xi, the Pattern Recognition Society, and SPIE.



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 databases, and systems applications.