Increasing the discrimination of synthetic aperture radar recognition models

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

Abstract. The focus of this work is optimizing recognition models for synthetic aperture radar (SAR) signatures of vehicles to improve the performance of a recognition algorithm under the extended operating conditions of target articulation, occlusion, and configuration variants. The recognition models are based on quasi-invariant local features, scattering center locations, and magnitudes. The approach determines the similarities and differences among the various vehicle models. Methods to penalize similar features or reward dissimilar features are used to increase the distinguishability of the recognition model instances. Extensive experimental recognition results are presented in terms of confusion matrices and receiver operating characteristic (ROC) curves to show the improvements in recognition performance for real SAR signatures of vehicle targets with articulation, configuration variants, and occlusion.
 [DOI: 10.1117/1.1517286]

Subject terms: articulated object recognition; automatic target recognition; object similarity; recognizing configuration variants; recognizing occluded objects.

Paper 010245 received July 16, 2001; revised manuscript received Apr. 26, 2002; accepted for publication Apr. 30, 2002.

1 Introduction

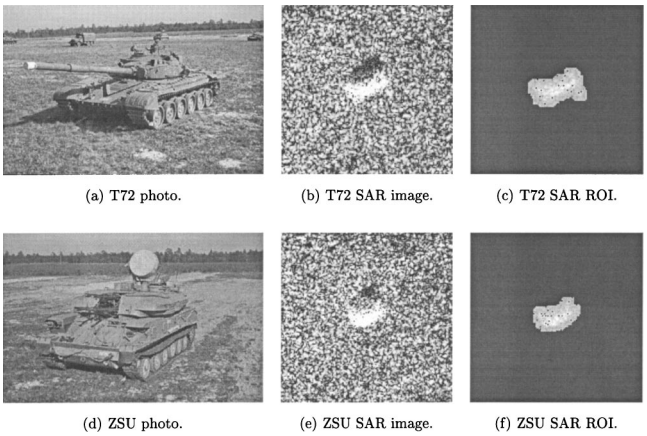
We are concerned with optimizing recognition models of synthetic aperture radar (SAR) signatures of real vehicles to improve the performance of a recognition system. The recognition system starts with real SAR chips of actual military vehicles from the MSTAR public data,[1] and ends with the identification of a specific vehicle type (e.g., a T72 tank). A major challenge is that the vehicles can be in articulated configurations (such as a tank with its turret rotated), have significant external configuration variants (fuel barrels, searchlights, etc.), or they can be partially occluded. The detection theory,[2,3] pattern recognition,[4–6] and neural network[7] approaches to SAR recognition all tend to use global features that are optimized for standard, nonarticulated, nonoccluded configurations. Approaches that rely on global features are not appropriate for recognizing occluded (or articulated) objects, because occlusion (or articulation) changes global features like the object outline and major axis.[8] Our previous work[9–12] relied on local features to successfully recognize articulated and highly occluded objects. We started using invariant locations of SAR scattering centers as features and later developed techniques using quasi-invariant locations and magnitudes of the scattering centers. Other work (Boshra and Bhanu[13]) on predicting the performance of recognition systems introduced the idea that recognition performance depends on the distortion in the test data and the inherent similarity of the object models. We develop an approach that determines the similarities and differences among the object models and uses this *a priori* knowledge to optimize the recognition models to improve the recognition system performance.

The two key contributions of this work are: 1. it quantifies the similarities between object models of SAR scatterer locations and magnitudes; and 2. it develops an approach that successfully uses *a priori* knowledge of the similarities between object models to improve the performance of a SAR recognition system. Using this conceptual approach, we present results on real MSTAR SAR data that shows that by explicitly measuring similarities of model objects and using it appropriately in a recognition system, we can increase target recognition performance under extended operating conditions.

The remainder of the work is organized as follows. The next section gives a description of the basic SAR recognition system. Section 3 describes the approach used to measure model similarity, presents similarity results, and gives example similarity weight functions. Section 4 gives experimental results for various similarity weight functions for the configuration variant cases. It also extends these results to articulated and occluded objects. Finally, conclusions are drawn in Sec. 5.

2 SAR Recognition System

The basic SAR recognition system is an off-line model construction process and a similar on-line recognition process. The approach is designed for SAR and is specifically intended to accommodate recognition of articulated and occluded objects. Standard nonarticulated models of the objects are used to recognize these same objects in nonstandard, articulated, and occluded configurations. An example photograph, MSTAR SAR target image, and extracted target region of interest (ROI), with scattering center locations shown as black dots, is given in Fig. 1 for T72 tank (serial number) #a64 and for ZSU 23/4 antiaircraft gun #d08. The models are a look-up table and the recognition process is an efficient search for positive evidence, using relative loca-



(a) T72 photo. (b) T72 SAR image. (c) T72 SAR ROI.

(d) ZSU photo. (e) ZSU SAR image. (f) ZSU SAR ROI.

Fig. 1 Example photo, MSTAR SAR image, and ROI (with peaks) for T72 tank #a64 and ZSU 23/4 antiaircraft gun #d08 (photo not to scale).

tions 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). Because of the specular radar reflections in SAR images, a significant number of features do not typically persist over a few degrees of rotation.[9] Consequently, we model each object at 1-deg azimuth increments. Any local reference point, such as a scattering center location, can be chosen as a basis point to establish a reference coordinate system for building a model of an object at a specific azimuth angle pose. 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, nonstandard target configurations, etc.), we use all N strongest scattering centers in turn as basis points to ensure that a valid basis point is obtained. Thus, to handle articulation and occlusion, the size of the look-up table models (and also the number of relative distances that are considered in the test image during recognition) are increased from N to $N(N-1)/2$. Using a technique like geometric hashing,[14] 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. Since the relative distances are not unique, there can be many of these labels (with different target, pose, etc. values) at each look-up table entry. The basic model construction algorithm is outlined in Fig. 2.

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 to $\pm L\%$. The design parameters N and L are optimized, based on experiments, to produce the best recognition results. Given that the MSTAR targets are ''centered'' in the chips, a ± 5 pixel limit on allowable translations is imposed for computational efficiency and to accommodate some imprecision in the centering algorithm. To accommodate some uncertainty in the scattering center locations, the eight neighbors of the nominal range and cross-range relative location are probed, and the translation results are also accumulated for a 3×3 neighborhood in the translation space. This voting in translation space, in effect, converts the consideration of scatterer pairs back into a group of scatterers at a consistent translation. 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 recognition algorithm actually makes a total of $9N(N-1)/2$ queries of the look-up table to accumulate evidence for the appropriate target type, azimuth angle, and translation. The models (labels with object, azimuth, etc.) associated with a specific look-up table entry are the ''real'' model and other models that happen, by coincidence, to have a scatterer pair with the same (range, cross-range) relative distance. The constraints on magnitude dif-

1. For each model Object do 2

2. For each model Azimuth do 3, 4, 5

3. Obtain the location (R, C) and magnitude (S) of the strongest N scatterers.

4. Order (R, C, S) triples by descending S.

5. For each origin O from 1 to N do 6

6. For each point P from O+1 to N do 7, 8

7. $dR = R_P - R_O$; $dC = C_P - C_O$.

8. At look-up table location dR, dC append to list entry with: Object, Azimuth, R_O, C_O, S_O, S_P.

Fig. 2 Basic model construction algorithm.

1. Obtain from test image the location (R,C) and magnitude (S) of N strongest scatterers.

2. Order (R,C,S) triples by descending S.

3. For each origin O from 1 to N do 4

4. For each point P from O+1 to N do 5, 6

5. $dR = R_P - R_O$; $dC = C_P - C_O$.

6. For DR from dR-1 to dR+1 do 7

7. For DC from dC-1 to dC+1 do 8, 9

8. Look up list of model entries at DR, DC.

9. For each model entry (E) in the list do 10

10. IF |tr $= R_O - R_E|$ < translation_limit and |tc $= C_O - C_E|$ < translation_limit and $|1 - S_{EO}/S_O|$ < magnitude_limit and $|1 - S_{EP}/S_P|$ < magnitude_limit THEN increment accumulator array [Object, Azimuth, tr, tc] by weighted_vote.

11. Query accumulator array for each Object, Azimuth, tr and tc, summing the votes in a 3x3 neighborhood in translation subspace about tr, tc; record the maximum vote_sum and the corresponding Object.

12. IF maximum vote_sum > threshold THEN result is Object ELSE result is "unknown".

Fig. 3 Recognition algorithm.

ferences filter out many of these false matches. In addition, while these collisions may occur at one relative location, the same random object-azimuth pair does not often keep showing up at other relative locations with appropriate scatterer magnitudes and mapping to a consistent 3×3 neighborhood in translation space, while the "correct" object does.

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 criteria for the quality of the recognition result that the votes for the potential winning object exceed some minimum threshold *d*. By varying the decision rule threshold, we obtain a form of receiver operating characteristic (ROC) curve with probability of correct identification (PCI) $= P\{$decide correct modeled object$|$object is true case$\}$, versus probability of false alarm, $P_f = \{$decide any modeled object$|$unknown is true case$\}$. The recognition algorithm is given in Fig. 3.

More formally, a radar image of object c at azimuth pose a consists of N (strongest in magnitude) scatterers, each with a magnitude S_k and range and cross-range locations R_k and C_k (R and C are range and cross-range directions, see Fig. 4), which (for consistency) are ordered by decreasing magnitude, such that $S_k \geqslant S_{k+1}$, where $k=1,...,N$ (and S_{N+1} is not defined). A model M of object c at azimuth a is given by:

$$M(c,a)=\{V_1(c,a),V_2(c,a),\ldots,V_{N(N-1)/2}(c,a)\}, \qquad (1)$$

which is comprised of the set of all pairwise observations, V_i:

$$V_i(c,a)=\{f_1,f_2,\ldots,f_6\}_i, \qquad (2)$$

where $i=1,2,...,N(N-1)/2$, $f_1=R_P-R_O$, $f_2=C_P-C_O$, $f_3=R_O$, $f_4=C_O$, $f_5=S_O$, $f_6=S_P$, and with the individual scatterers in each pair labeled O and P, so that $S_O \geqslant S_P$ for consistency (see Fig. 4).

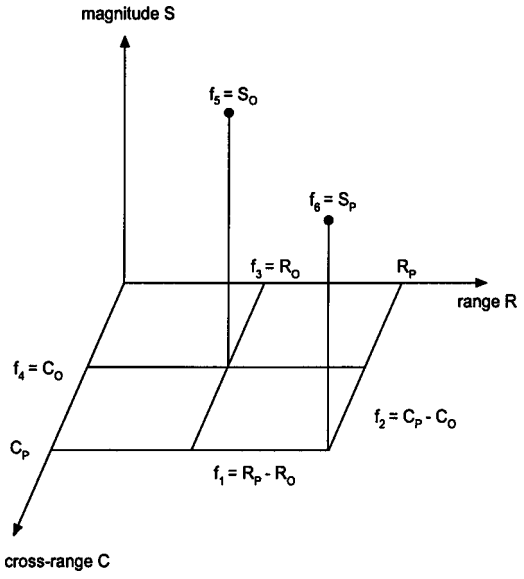


Fig. 4 Observation for a pair of scatterers *O* and *P*.

We define a match, H, as:

$$H(V_i,V_j)=\begin{cases}1 & \text{if } |(f_b)_i-(f_b)_j|\leq\delta_b,\ \forall b=1,...,6\\ 0 & \text{otherwise}\end{cases}, \quad (3)$$

where the match constraints are $\delta_1=\delta_2=0$ pixels, $\delta_3=\delta_4=5$ pixels, and $\delta_5=\delta_6=L$ percent. A subscript t applied to a match denotes that the match, H_t, is associated with the relative translation $t(R,C)=(\Delta f_3,\Delta f_4)$ of the stronger scatterers in the two observations.

The recognition result, T, for some test image (with undetermined class x and azimuth y) is a maximal match that is greater than a threshold, d, given by:

$$T=\begin{cases}[c,a], & \text{if } \arg\max_{c,a,t}\{\Sigma_{l=1}^{9}\Sigma_{k=1}^{N(N-1)/2}\Sigma_{n=1}^{9}H_t^l[V_k^n(x,y),V_m(c,a)]\}>d\\ \text{unknown}, & \text{otherwise}\end{cases}, \quad (4)$$

where $V_m\in M(c,a)\forall m$, such that $|(f_1)_{V_k^n}-(f_1)_{V_m}|=0$ and $|(f_2)_{V_k^n}-(f_2)_{V_m}|=0$. Note that this formulation for V_m avoids an exhaustive search of all the models and can be implemented as a look-up table. The nine observations (denoted by the superscript n in V_k^n) are made to account for location uncertainty by taking the 3×3 neighbors about the nominal values for the relative locations f_1 and f_2 of scatterer pair k in the test image. Similarly, the nine matches (denoted by the superscript l in H_t^l) are computed at the 3×3 neighbors at ±1 pixel about the nominal values for translation, $t(R,C)$, of the scatterers in the test image from the model.

3 Model Similarity Measurement and Weighting

3.1 *Model Similarity*

Model similarity can be measured in terms of collisions, where a collision is an instance when observations of two different objects map into the same location (within some specified region of uncertainty) in feature space, i.e., if $H(V_i,V_j)=1$ and $c_i\neq c_j$. The recognition system described in the preceding section has a 6-D feature space based on the range and cross-range positions and magnitudes of pairs of scatterers [see Eq. (2)]. As noted before [in Eq. (1)], the model of an object at some azimuth, with N scatterers, is represented by $N(N-1)/2$ observations using pairs of scatterers with each pair mapped into the 6-D feature space. While the 6-D feature space could be represented by a simple 6-D array in concept, the large range of potential feature values and high dimensionality make other implementations more practical. The nature of the SAR problem, with discrete pixel values for distances and a large dynamic range for scatterer magnitudes, leads to a natural model implementation, shown previously in Fig. 4, where the relative range and cross-range locations of a scatterer pair are direct indices to a physical 2-D array of lists that contain another 4-D of information and the label with the object and pose. Thus, the model construction algorithm of Fig. 2 does not directly provide collisions in all six dimensions of feature space. To determine if two objects map to the same location in feature space, we need to apply the same constraints that are used in the recognition algorithm [see step 10 of Fig. 3 and Eq. (3)], because the constraints dictate the size of the region or bucket in feature space that is considered the same location.

The general approach to measure the similarity of one model object with respect to several other objects is to first build the look-up table models of the other objects using the normal model construction algorithm of Fig. 2, and then use a modified version of the recognition algorithm of Fig. 3 with the subject model object (at all the modeled azimuths) as the test conditions to obtain a histogram of the number of occurrences of various numbers of collisions. Basically the modified algorithm uses the first 10 steps of Fig. 3, with the consideration of each pair of scatterers as a separate occurrence (starting a new count of collisions at step 5), and if the constraints are satisfied (at step 10), then a collision is counted. The total number of occurrences (and observations) is equal to $AN(N-1)/2$, where A is the number of azimuths modeled (some of the MSTAR data was sequestered, so not all 360 deg were available).

Figure 5 shows example model collision histograms (at $N=39$ and $L=9$) for four MSTAR vehicles (at 15-deg depression angle): BMP2 armored personnel carrier (APC) serial number (#) c21; BTR70 APC #c71; T72 tank #132; and ZSU23/4 antiaircraft gun #d08. Note that the ZSU23/4

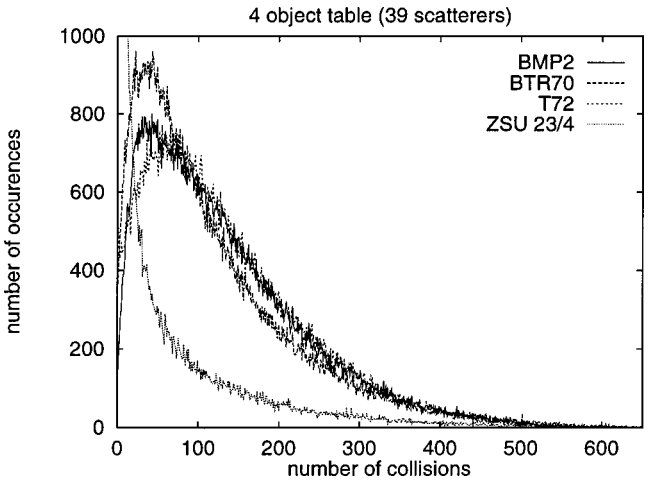


Fig. 5 Example recognition model look-up table collision histograms.

Table 1 Number of collisions for a given percent of the population (example for $N=39$, $L=9$).

Object	Number of collisions									
BMP2	27	46	66	87	110	136	167	209	274	676
BTR70	21	37	53	70	91	116	148	192	266	712
T72	27	48	68	89	111	137	168	209	271	667
ZSU 23/4	0	0	0	0	0	1	3	18	78	760
Population percent	10	20	30	40	50	60	70	80	90	100

has significantly fewer collisions with the other vehicles, because the ZSU23/4 SAR scatterers cover a larger area than the other objects, and thus have fewer collisions.

The similarity of a pair of scatterers of a given object (at a given azimuth) to the other objects modeled can be measured by the number of collisions with other objects in the look-up table. This can be expressed as a relative measure by using the collision histogram. For convenience, the population of collisions for a particular object is mapped into equal partitions (each with 10% of the total number of collisions). As an example, for the collision histograms in Fig. 5, we obtain the results in Table 1, which shows the number of collisions for a given percent of the population. For the BMP2, for example, 27 collisions or less is in the 10% of the population that is the least similar to the other three models (whereas 90% of the BMP2 scatterer pairs have 274 or less collisions).

3.2 *Weighted Voting*

The *a priori* knowledge of the similarities between object models, expressed as the number of collisions for a given percent of the population, can be captured by assigning weighted votes to model entries in the look-up table, based on collisions with other objects. This is accomplished off-line by again using a version of the recognition algorithm to obtain the number of look-up table collisions for a particular observation with a pair of scatterers from a subject model and azimuth, as before, and then based on the number of collisions, determining the population partition (e.g., using Table 1), and finally a given weight function is used to assign a weight label to that instance of the particular model observation entry in the look-up table. Thus, in this approach the model similarities, collisions, and associated weightings are all precomputed and appropriate weightings are stored in the look-up table during the off-line modeling process. The weighted models are obtained using the weighted version of an observation [similar to Eq. (2)] given by:

$$\hat{V}_i(c,a)=\{w,f_1,f_2,\ldots,f_6\}_i, \tag{5}$$

where w is the weight. The weighted version of a match [similar to Eq. (3)], given by:

$$\hat{H}(V_i,\hat{V}_j)=\begin{cases} w & \text{if } |(f_b)_i-(f_b)_j|\leq\delta_b, \ \forall \ b=1,\ldots,6 \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

which can be substituted in Eq. (4) to obtain weighted recognition results. The various weight functions, used in this research to specify w, are shown in Fig. 6, which plots the weight value assigned versus percent of collision population. Function 1 [Fig. 6(a)] applies equal weight to all the values and is later referred to as unweighted. Functions 2 through 4 [Figs. 6(b) through 6(d)], the convex weight functions, penalize the most similar features (in the right tail of the histogram). Function 5 [Fig. 6(e)] with equal steps is linear. Functions 6 and 7 [Figs. 6(f) and 6(g)], which reward uniqueness (the left tail of the histogram), are concave. These weight functions illustrate a range of possibilities, from function 2, which penalizes only the most similar 10% of the population, to function 7, which rewards only the most dissimilar 10%. These seven weight functions are used and comparative performance results are obtained in experiments described in the next section.

4 Experimental Results

4.1 *Configuration Variants*

Our previous results[9] (using a distance-weighted voting technique, where the weight was proportional to the sum of the absolute values of the relative range and cross-range distances between the scatterer pair) showed that for the real vehicles used in the MSTAR data, the differences of configurations for an object type are a more significant challenge for recognition than articulation (where the model and the test data are the same physical object under different conditions). Similarly, the previous results[10] on occluded objects (using an unweighted voting technique) demonstrated significantly better recognition results than the configuration variant cases. For these reasons, in this research we follow a similar approach and optimize the recognition system for the difficult configuration variant cases, and then utilize the same system parameters for the articulation and occlusion cases.

Data. In these (15-deg depression angle) configuration variant experiments, the two object model cases use T72 tank #132 and BMP2 APC #C21 as models, while the four object model cases add BTR70 APC #c71 and ZSU23/4 gun #d08. The test data are two other variants of the T72 (#812, #s7) and two variants of the BMP (#9563, #9566). In addition, BRDM2 APC #e71 is used as an unknown confuser vehicle.

Results. The forced recognition results for MSTAR configuration variants are shown in Fig. 7 for both two-object and four-object look-up table models using various weight functions (defined earlier in Fig. 6). These results use the optimal parameters (N,L) for each weight function and table size. For the two object cases, function 3 gives the

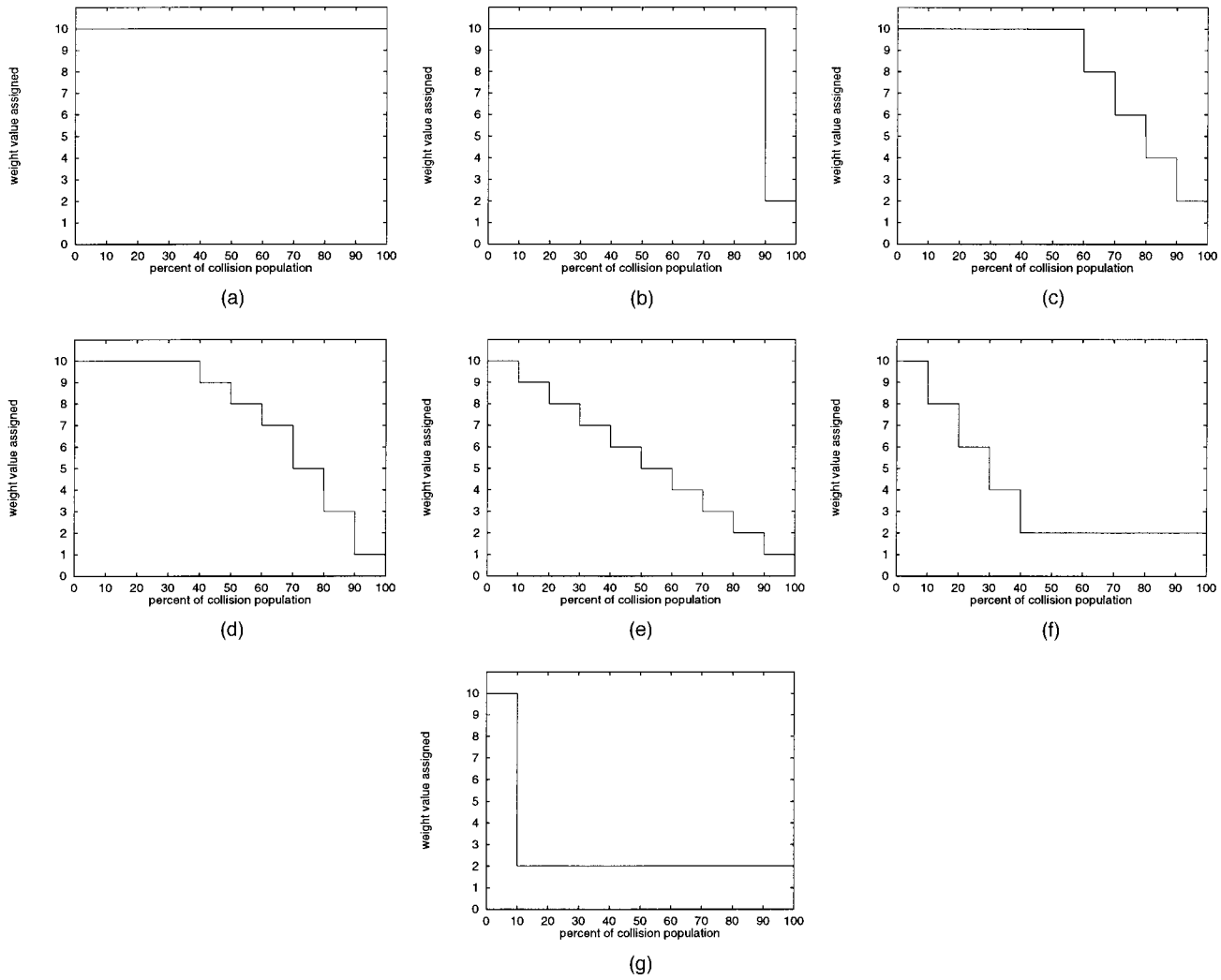


Fig. 6 Table weighting functions: (a) function 1, (b) function 2, (c) function 3, (d) function 4, (e) function 5, (f) function 6, and (g) function 7.

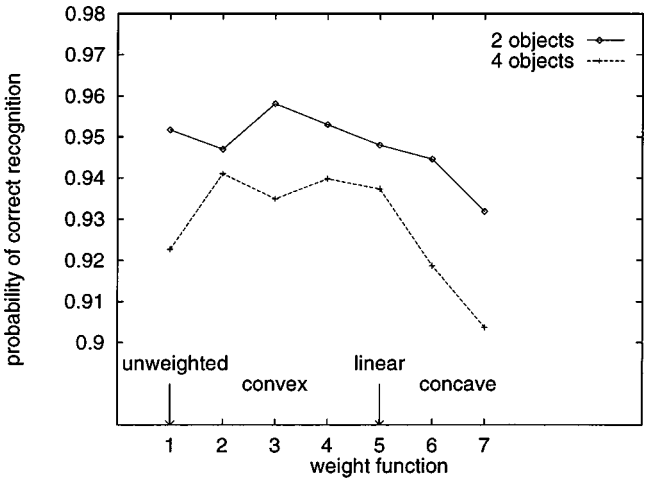


Fig. 7 Effect of table size and weighting function on forced recognition of MSTAR configuration variants.

best results, a recognition rate of 95.81%, compared to the unweighted case of 95.17%. For the four object cases, the convex and linear weighting functions all provide better forced recognition performance than the unweighted case. The concave weighting functions result in worse performance than the unweighted case. The best four object result is 94.17% for function 2, compared to the unweighted case of 92.27%. Thus, increasing the number of objects modeled from two to four reduces the forced recognition rate by 2.9% (95.17 to 92.27) for the unweighted case, while using model similarity information in the optimum weight function reduces that loss to 1% (95.17 to 94.17).

Table 2 shows example confusion matrices that illustrate the effect of going from a two-object recognition system to a four-object model recognition system for the MSTAR configuration variant data. In both cases the system parameters (N,L) are optimized for forced recognition [two objects at (38, 11) and four at (38, 12)], both are unweighted cases (constant weight of 10), and both are for $d=1700$. (At least 1700 votes, with a weight of 10, is equivalent to 19 or more scatterers that matched.) Comparing the two-object results on the left of Table 2 with the four-object

Table 2 Effect of two and four models on MSTAR configuration variant confusion matrices (unweighted, $d=1700$).

test targets [serial number]		Identification results (configuration modeled)			Identification results (configuration modeled)				
		BMP2 (#C21)	T72 (#132)	Unknown	BMP2 (#C21)	T72 (#132)	BTR70 (#C71)	ZSU23/4 (#d08)	Unknown
BMP2	[#9563,9566]	189	3	25	189	2	8	0	18
T72	[#812,s7]	8	131	58	11	138	1	0	47
BRDM2	(confuser)	28	4	214	27	5	47	0	167

results on the right, we observe that basically a large number of confusers and a few targets move from the Unknown column to the additional models. Thus, while the recognition results are similar for two and four models (PCI $=0.773$ and 0.790, respectively) there are increased false alarms ($P_f=0.13$ and 0.32, respectively), which would move the knee of the ROC curve to the right.

Table 3 shows an example MSTAR configuration variant four-object confusion matrix for weight function 4. The system parameters (37, 9) are optimized for forced recognition with weight function 4 and a d of 1100 is chosen to yield a PCI of 0.776, which is similar to the results shown in Table 2. (At least 1100 votes, with an average weight for function 4 of 7.3, is equivalent to 18 or more scatterers matched.) Comparing the earlier four-object unweighted results, shown on the right of Table 2, with the weighted results of Table 3, we observe that half the misidentifications (11 of 22) are moved to the unknown column. This reduction in misidentifications shows that the model weighting approach is increasing the distinguishability of the modeled objects. This reduction in misidentifications does not show up directly in the ROC curve results, which treat the off-diagonal target misidentifications the same as the misses, where a target is called unknown (i.e., both are cases where the target was not correctly identified). However, the weight function (which effectively reduces the average weighting) allows a similar PCI to be achieved with a lower vote threshold (1100 votes versus 1700 votes) and results in fewer false alarms. Thus, the lower P_f of 0.276 for the weighted case versus 0.321 for the unweighted case would move the ROC curve to the left.

ROC curves are generated for the four-object configuration variant cases by using the optimum parameters for the forced recognition case and varying the vote threshold. Figure 8 shows that the ROC curves for the convex and linear weight functions provide generally better performance than the unweighted case. In addition, Fig. 9 shows that the concave weight functions give worse performance than the unweighted case (except for the region where PCI<0.5, $P_f<0.05$). The convex weight functions penalize the most common features and so are not much affected by noise (due to configuration differences or other confuser vehicles). On the other hand, the concave weight functions reward (very strongly reward in function 7) the relatively unique features, which makes them susceptible to conditions where noise is strongly rewarded.

4.2 *Articulation*

Data. 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).

Results. Figure 10 shows the ROC curves, with excellent articulated object recognition results for both the weight function 2 and the unweighted cases. Since weight function 2, with $N=39$ and $L=9$, gives the optimum ROC results for the 2 object (T72, BMP2) configuration experiments and the optimum unweighted parameters are $N=38$ and $L=11$, these same parameters are used for the articulation experiments.

Table 3 Example of the MSTAR configuration variant confusion matrix for weight function 4 ($d=1100$).

test targets [serial number]		Identification results (configuration modeled)				
		BMP2 (#C21)	T72 (#132)	BTR70 (#C71)	ZSU23/4 (#d08)	Unknown
BMP2	[#9563,9566]	179	6	1	0	32
T72	[#812,s7]	4	143	0	0	50
BRDM2	(confuser)	30	6	32	0	178

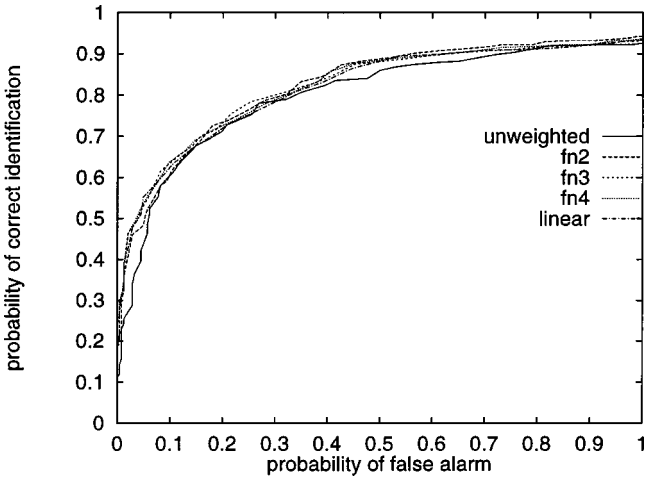


Fig. 8 MSTAR configuration variant ROCs for beneficial weight functions (four objects).

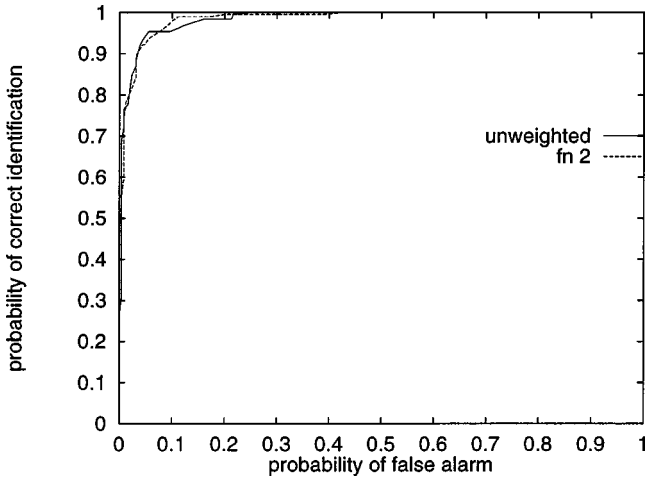


Fig. 10 Articulation recognition results.

4.3 *Occlusion*

Data. The occlusion experiments use the same four models as the configuration variant experiments: T72 tank #132, BMP2 APC #C21, BTR70 APC #c71, and ZSU23/4 gun #d08 (all at 15-deg depression angles). Since there is no real SAR data with occluded objects available to the general public, the occluded test data in this work are simulated by starting with a given number of the strongest scattering centers in target chips of these same four objects, and then removing the appropriate number of scattering centers encountered in order from 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 is the same technique used in Ref. 10; it keeps the number of scatterers constant and acts as a surrogate for some potential occluding object. In our previous work on occluded objects,[10] the confuser vehicle was occluded. However, while the target may be occluded, the confuser vehicle may not necessarily be occluded in the practical case. Hence, in this research the BRDM2 APC (#e71) is an unoccluded confuser vehicle, which is a more difficult case.

Results. Figure 11 shows the effect of occlusion on ROC curves for weight function 2, with $N=40$ and $L=9$ (while $N=40$ is not optimum, it yields occlusion in 5% increments). Here with the unoccluded confuser, excellent recognition results are achieved for less than 45% occlusion, compared with the prior 70% occlusion with an occluded confuser.[10]

5 Conclusions

The similarities between object models can be effectively quantified using histograms of collisions in feature space. This *a priori* knowledge of object similarity can be successfully used to improve the performance of SAR target recognition. The approach can increase the distinguishability of the modeled objects, reduce misidentifications, and result in decreased false alarms. In the most difficult configuration variant cases, the convex and linear weight func-

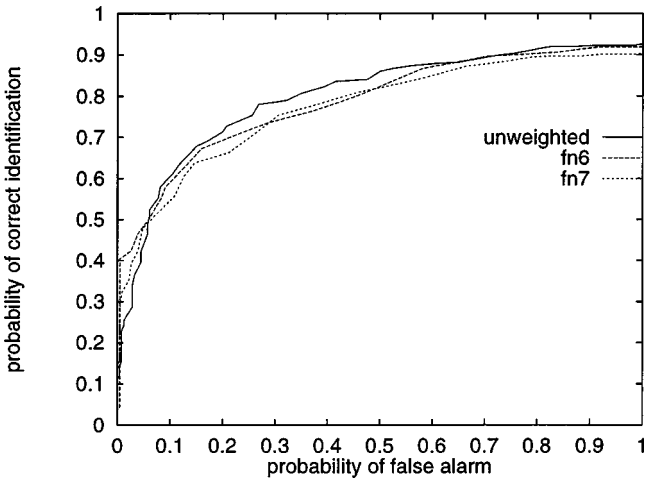


Fig. 9 MSTAR configuration variant ROCs for concave weight functions (four objects).

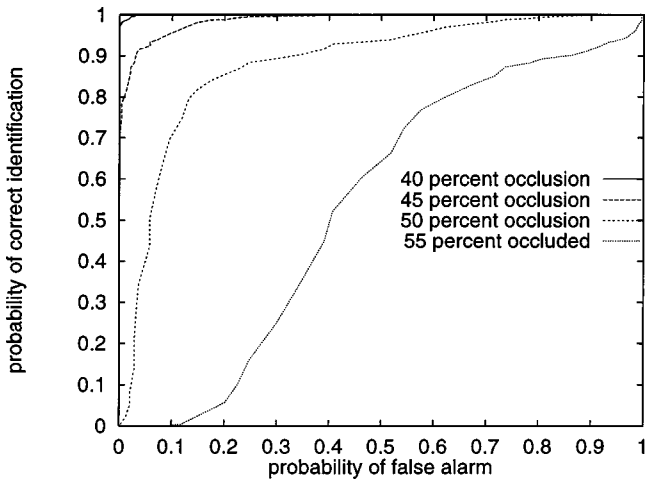


Fig. 11 Effect of occlusion on receiver operating characteristics.

tions, which penalize the most common features, give better performance than the concave weight functions, which strongly reward relatively unique features. The experimentally determined optimum weight function reduces the impact of scaling from two to four models from a 2.9% reduction in forced recognition rate to a 1.0% reduction. The same approach (and parameters) also provide excellent recognition results for articulated objects and up to 45% for occluded objects. While the current work is directed at similarities between different object models, in the future an analogous approach could be applied to determine similarities among variants of the same object to develop a ''class model'' of the object that incorporates the common features.

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. Worrell, V. Velten, J. Mossing, and M. Bryant, ''Standard SAR ATR evaluation experiments using the MSTAR public release data set,'' *Proc. SPIE* **3370**, 566–573 (Apr. 1998).
2. D. Carlson, B. Kumar, R. Mitchell, and M. Hoffelder, ''Optimal tradeoff distance classifier correlation filters (OTDCCFs) for synthetic aperture radar automatic target recognition,'' *Proc. SPIE* **3070**, 110–120 (Apr. 1997).
3. D. Casasent and R. Shenoy, ''Synthetic aperture radar detection and clutter rejection MINACE filters,'' *Pattern Recogn.* **30**(1), 151–162 (Jan 1997).
4. R. Meth and R. Chellappa, ''Automatic classification of targets in synthetic aperture radar imagery using topographic features,'' *Proc. SPIE* **2757**, 186–193 (Apr. 1996).
5. T. Ryan and B. Egaas, ''SAR target indexing with hierarchical distance transforms,'' *Proc. SPIE* **2757**, 243–252 (Apr. 1996).
6. 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,'' *Proc. SPIE* **1960**, 57–71 (Apr. 1993).
7. D. Casasent and R. Shenoy, ''Feature space trajectory for distorted-object classification and pose estimation in SAR,'' *Opt. Eng.* **36**(10), 2719–2728 (Oct. 1997).
8. J. H. Yi, B. Bhanu, and M. Li, ''Target indexing in SAR images using scattering centers and the Hausdorff distance,'' *Pattern Recogn. Lett.* **17**, 1191–1198 (1996).
9. B. Bhanu and G. Jones, ''Recognizing target variations and articulations in synthetic aperture radar images,'' *Opt. Eng.* **39**(3), 712–723 (Mar. 2000).
10. G. Jones and B. Bhanu, ''Recognizing occluded objects in SAR images,'' *IEEE Trans. Aerosp. Electron. Syst.* **37**(1), 316–328 (Jan. 2001).
11. G. Jones and B. Bhanu, ''Recognizing articulated targets in SAR images,'' *Pattern Recogn.* **34**(2), 469–485 (Feb. 2001).
12. G. Jones and B. Bhanu, ''Recognition of articulated and occluded objects,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(7), 603–613 (July 1999).
13. M. Boshra and B. Bhanu, ''Predicting an upper bound on performance of target recognition in SAR images,'' *IEEE Trans. Aerosp. Electron. Syst.* **37**(3), 876–888 (July 2001).
14. Y. Lamdan and H. Wolfson, ''Geometric hashing: A general and efficient model-based recognition scheme,'' *Proc. Intl. Conf. Computer Vision*, pp. 238–249, IEEE (Dec. 1988).

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. Currently he is the Director of the Center for Research in Intelligent Systems (CRIS) at the University of California, Riverside, where he has been a Professor and Director of Visualization and Intelligent Systems Laboratory (VISLab) since 1991. 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 (forthcoming), *Genetic Learning for Adaptive Image Segmentation* (Kluwer 1994), and *Qualitative Motion Understanding* (Kluwer 1992). He has received two outstanding paper awards from the Pattern Recognition Society and has received industrial awards for technical excellence, outstanding contributions and team efforts. He holds 10 U.S. and international patents and more than 200 reviewed technical publications in the areas of his interest. He is a Fellow of IEEE, AAAS, and IAPR, and a member of ACM, AAAI, and SPIE.

Grinnell Jones III is a National Merit Scholar and 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 25 years in development engineering, missile operations, and acquisition management with the U.S. Air Force, he has been conducting computer science research developing systems and algorithms for automatic target recognition using synthetic aperture radar imagery for the past seven years with the University of California, Riverside. His research interests include object recognition, computer vision, machine learning, image and video databases, and systems applications.