Increasing the Discrimination of SAR Recognition Models

Bir Bhanu and Grinnell Jones III

Center for Research in Intelligent Systems University of California, Riverside, California, 92521 USA

ABSTRACT

The focus of this paper 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 MSTAR vehicle targets with articulation, configuration variants and occlusion.

Keywords: articulated object recognition, automatic target recognition, object similarity, recognizing configuration variants, recognizing occluded objects, synthetic aperture radar

1. INTRODUCTION

In this paper 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 data¹² 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, 3,4 pattern recognition 10,11,13 and neural network⁵ approaches to SAR recognition all tend to use global features that are optimized for standard, non-articulated, non-occluded 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.¹⁴ Our previous work^{1,6,7,8} 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 quasiinvariant locations and magnitudes of the scattering centers. Other work, by Boshra and Bhanu on predicting the performance of recognition systems,² introduced the idea that recognition performance depends on the distortion in the test data and the inherent similarity of the object models. In this paper we develop an approach that determines the similarities and differences among the object models and uses this apriori knowledge to optimize the recognition models to improve the recognition system performance.

The key contributions of this paper are:

- 1. Quantifies the similarities between object models of SAR scatterer locations and magnitudes.
- 2. Develops an approach that successfully uses apriori knowledge of the similarities between object models to improve the performance of a SAR recognition system.

The remainder of the paper 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. Section 5 extends these results to articulated and occluded objects. Finally, conclusions are drawn in Section 6.

Further author information -Email: {bhanu, or grinnell} @cris.ucr.edu URL: http://www.cris.ucr.edu

- 1. For each model Object do 2 2. For each model Azimuth do 3. 4. 5
- 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 .

Figure 1. Basic model construction algorithm.

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 non-articulated models of the objects are used to recognize these same objects in non-standard, articulated and occluded configurations. The models are a look-up table and the recognition process is an efficient search for *positive evidence*, using relative locations of the scattering centers in the test image to access the look-up table and generate votes for the appropriate object (and azimuth pose).

The relative locations and magnitudes of the N strongest SAR scattering centers (local maxima in the radar return signal) are used as characteristic features (where N, the number of scattering centers used, is a design parameter). Because of the specular radar reflections in SAR images, a significant number of features do not typically persist over a few degrees of rotation.¹ Consequently, we model each object at 1° 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, non-standard target configurations, etc.), we use all N strongest scattering centers in turn as basis points to ensure that a valid basis point is obtained. 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,⁹ 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 lookup table entry. The basic model construction algorithm is outlined in Figure 1.

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 the MSTAR targets are 'centered' in the chips, a ± 5 pixel limit on allowable translations is imposed for computational efficiency.) To accommodate some uncertainty in the scattering center locations, the eight-neighbors of the nominal range and cross-range relative location are also probed and the translation results are accumulated for a 3x3 neighborhood in the translation subspace. This voting in 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.

Obtain from test image the location (R, C) and magnitude (S) of N strongest scatterers.
Order (R, C, S) triples by descending S.
For each origin O from 1 to N do 4
For each point P from O+1 to N do 5, 6
dR = R_P - R_O; dC = C_P - C_O.
For DR from dR-1 to dR+1 do 7
For DC from dC-1 to dC+1 do 8, 9
Look up list of model entries at DR, DC.
For each model entry E in the list do 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.
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.
IF maximum vote_sum > threshold THEN result is Object ELSE result is "unknown".

Figure 2. Recognition algorithm.

these false matches. In addition, while these collisions may occur at one relative location, the same random objectazimuth pair doesn't often keep showing up at other relative locations with appropriate scatterer magnitudes and mapping to a consistent 3x3 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 v_{min} . 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 object|object is true}, vs. probability of false alarm, P_f = {decide any object|unknown is true}. The recognition algorithm is given in Figure 2.

3. MODEL SIMILARITY MEASUREMENT AND WEIGHTING

Model similarity can be measured in terms of *collisions*, where a collision is an instance when two different objects map into the same location in feature space. The recognition system described in the preceding section has a 6 dimensional (6D) feature space based on the range and cross range positions and magnitudes of pairs of scatterers. As noted before, the model of an object at some azimuth pose, with N scatterers, is represented by N(N-1)/2 pairs of scatterers with each pair mapped into the 6D feature space. While the 6D feature space could be represented by a simple 6D 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 Figure 1, where the relative range and crossrange locations of a scatterer pair are direct indices to a physical 2D array of lists that contain another 4D of information and the label of the object and pose. Thus, the model construction algorithm of Figure 1 does not directly provide collisions in all 6 dimensions of feature space. In order to determine if two objects map to the same location in feature space we need to apply the same constraints as are used in the recognition algorithm (see step 10 of Figure 2), because the constraints dictate the size of the region or bucket in feature space that is considered "the same".

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 the normal model construction algorithm of Figure 1, then use a modified version of the recognition algorithm of Figure 2 with the subject model object (at all the modelled 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 Figure 2, with the consideration of each pair of scatterers as a seperate 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 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° were available).



Figure 3. 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	
Percent	10	20	30	40	50	60	70	80	90	100	

Figure 3 shows example model collision histograms (at N = 39 and L = 9) for four MSTAR vehicles (at 15° depression angle): BMP2 armored personnel carrier (APC) serial number (#) c21; BTR70 APC #c71; T72 tank #132 and ZSU23/4 anti-aircraft gun #d08. Note that the ZSU23/4 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 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 Figure 3 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).

The apriori 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 occurrence with a pair of scatterers from a subject model and azimuth, as before, and then based on the number of collisions determine the population partition (e.g., using the inverse of Table 1) and finally a given weight function is used to assign a weight label to that instance of the particular model, azimuth scatterer pair 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.



Figure 4. Table weighting functions.

Figure 4 shows the various weight functions that are used in this research. Function 1 applies equal weight to all the values and is later referred to as unweighted. Functions 2-4, the convex weight functions, penalize the most similar features (in the right tail of the histogram). Function 5, with equal steps is linear. While functions 6-7, 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%.

4. RESULTS FOR CONFIGURATION VARIANTS

Our previous results¹ (using a distance weighted voting technique) 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 results⁶ 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 other cases.

In these (15° 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. The forced recognition results for MSTAR configuration variants are shown in Figure 5 for both two object and four object look up table models using various weight functions (defined earlier in Figure 4). These results use the optimal parameters (N,L) for each weight function and table size. For the two object cases, function 3 gives the 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 - 92.27) for the unweighted case, while using model similarity information in the optimum weight function reduces that loss to 1% (95.17 - 94.17).

Table 2 shows example confusion matricies 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 (2 objects at 38,11 and 4 at 38,12), both are unweighted cases



Figure 5. Effect of table size and weighting function on forced recognition of MSTAR configuration variants.

Table 2. Effect of 2 and 4 models on MSTAR configuration variant confusion matrices (unweighted, $V_{min} = 1700$)

		Iden	tification	results	Identification results					
		(config	guration n	nodeled)	(configuration modeled)					
test	t targets	BMP2	T72	Unknown	BMP2	T72	BTR70	ZSU23/4	Unknown	
[serial number]		(#C21)	(#132)		(#C21)	(#132)	(#C71)	(#d08)		
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	

(constant weight of 10), and both are for $V_{min} = 1700$. (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 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 2 and 4 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 V_{min} 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 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 vs. 1700 votes) and results in fewer false alarms. Thus, the lower P_f of 0.276 for the weighted case, vs. 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 6 shows that the ROC curves for the convex

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

Table 3. Example MSTAR configuration variant confusion matrix for weight function 4 ($V_{min} = 1100$)



Figure 6. MSTAR configuration variant ROC curves for beneficial weight functions (4 objects).

and linear weight functions provide generally better performance than the unwighted case. In addition, Figure 7 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.

5. ARTICULATION AND OCCLUSION RESULTS

In the articulation experiments the models are non-articulated 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° depression angle). 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 unwighted parameters are N = 38 and L = 11, these same parameters are used for the articulation experiments. Figure 8 shows the ROC curves, with excellent articulated object recognition results for both the weight function 2 and the unweighted cases.

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° depression angle). Since there is no real SAR data with occluded objects available to the general public, the occluded test data in this paper is 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



Figure 7. MSTAR configuration variant ROC for concave weight functions (4 objects).



Figure 8. Articulation recognition results.



Figure 9. Effect of occlusion on receiver operating characteristics.

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 in⁶; it keeps the number of scatterers constant and acts as a surrogate for some potential occluding object. In our previous work on occluded objects,⁶ 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. Figure 9 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 percent occlusion, compared with the prior 70 percent occlusion with an occluded confuser.⁶

6. CONCLUSIONS

The similarities between object models can be effectively quantified using histograms of collisions in feature space. This apriori 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 functions, which penalize the most common features, give better performance than the concave weight functions, which strongly reward relatively unique features. Here the experimentally determined optimum weight function reduces the impact of scaling from 2 to 4 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% 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.

7. 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. B. Bhanu and G. Jones. "Recognizing target variations and articulations in synthetic aperture radar images," *Optical Engineering*, Vol. 39 No. 3, pp. 712-723, March 2000.
- 2. M. Boshra and B. Bhanu. "Predicitng Performance of Object Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 9, pp. 956-969, September 2000.
- 3. D. Carlson, B. Kumar, R Mitchell and M. Hoffelder. "Optimal trade-off distance classifier correlation filters (OTDCCFs) for synthetic aperture radar automatic target recognition," *SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery IV*, Vol. 3070, pp. 110-120, April 1997.
- 4. D. Casasent and R. Shenoy. "Synthetic aperture radar detection and clutter rejection MINACE filters," *Pattern Recognition*, Vol. 30, No. 1, pp. 151-162, Jan 1997.
- 5. D. Casasent and R. Shenoy. "Feature space trajectory for distorted-object classification and pose estimation in SAR," *Optical Engineering*, Vol. 36, pp. 2719-2728, Oct. 1997.
- 6. G. Jones and B. Bhanu. "Recognizing occluded objects in SAR images," *IEEE Trans. on Aerospace and Electronic Systems*, in press 2001.
- 7. G. Jones and B. Bhanu. "Recognizing articulated targets in SAR images," *Pattern Recognition*, Vol. 34, No. 2, February 2001.
- 8. G. Jones and B. Bhanu, "Recognition of articulated and occluded objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 7, pp 603-613, July 1999.
- 9. Y. Lamden and H. Wolfson. "Geometric hashing: A general and efficient model-based recognition scheme," Proc. International Conference on Computer Vision, pp. 238-249, December 1988.
- 10. R. Meth and R. Chellappa. "Automatic classification of targets in synthetic aperture radar imagery using topographic features," SPIE Proceedings: Algorithms for SAR Imagery III, Vol. 2757, pp. 186-193, April 1996.
- 11. T. Ryan and B. Egaas. "SAR target indexing with hierarchical distance transforms," SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery III, Vol. 2757, pp. 243-252, April 1996.
- 12. T. Ross, S. Worrell, V. Velten, J. Mossing, and M. Bryant. "Standard SAR ATR Evaluation Experiments using the MSTAR Public Release Data Set," SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery V, Vol. 3370, pp. 566-573, April 1998.
- 13. 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 Proceedings: Automatic Object Recognition III*, Vol. 1960, pp. 57-71, April 1993.
- 14. 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. 1191-1198, 1996.