Object Recognition Results
Using MSTAR Synthetic Aperture Radar Data

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

Abstract

This paper outlines an approach and experimental results for Synthetic Aperture Radar (SAR) object recognition using the MSTAR data. With SAR scattering center locations and magnitudes as features, the invariance of these features is shown with object articulation (e.g., rotation of a tank turret) and with external configuration variants. This scatterer location and magnitude quasi-invariance is used as a basis for development of a SAR recognition system that successfully identifies articulated and non-standard configuration vehicles based on non-articulated, standard recognition models. The forced recognition results and pose accuracy are given. The effect of different confusers on the receiver operating characteristic (ROC) curves are illustrated along with ROC curves for configuration variants, articulations and small changes in depression angle. Results are given that show that integrating the results of multiple recognizers can lead to significantly improved performance over the single best recognizer.

1 Introduction

The performance of an automatic system for recognizing objects in SAR imagery is a complex function of sensor operating conditions, background clutter, object configurations and the features and algorithms used for object detection/segmentation, feature extraction and recognition [4]. Model based recognition using MSTAR data has been a very active area of research in the last several years [13]-[15] The focus of this paper is the recognition subsystem itself, starting with SAR chips of real vehicles from the MSTAR public data and ending with the vehicle identification. The specific challenges for the recognition subsystem are to address the need for automated recognition of military vehicles that can have articulated parts (like the turret of a tank), have significant external configuration variants (like fuel barrels, searchlights, etc.), or the vehicles can be partially hidden. Previous recognition methods involving template matching [12] are not useful in some of these cases, because articulation or occlusion changes global features like the object outline and major axis. In order to characterize the performance of the recognition subsystem we approach the problem scientifically from fundamentals. For a given sensor system we characterize the

variance of features with azimuth, for the objects we characterize articulation and configuration invariants and develop a SAR specific recognition system based on the quasi-invariance of SAR scattering center locations and magnitudes. We then characterize the forced recognition performance of this system in terms of recognition rate, pose accuracy, and show the effect of occlusion. The effects of various confuser vehicles on identification performance results are presented as vote space scatter plots and ROC curves for configuration variants. In addition, ROC curves compare the results for configuration variants, for articulated objects and for a small change in depression angle. Finally, the effect of integrating multiple recognizers is shown for the most difficult case, configuration variants of the T72 tank and BMP2 armored personnel carrier (APC) using the most difficult confuser, the BTR70 APC.

2 Scattering Center Characteristics

The typical detailed edge and straight line features of man-made objects in the visual world do not have good counterparts in SAR images at one foot 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, independent of translation and serve as distinguishing features. In addition to the scatterer locations, the magnitudes of the peaks are also features that we use for recognition.

2.1 Variance with object pose

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°) 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 (when we compare scatterer locations in the ground plane of an image, rotated by some azimuth increment, with another image at the resulting azimuth angle). Figure 1 shows an example of the scatterer location invariance (for the 40 strongest
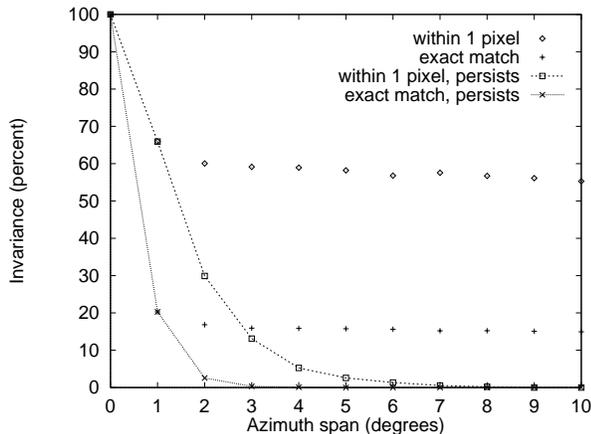
Figure 1: Scatterer location persistence, T72 #132.

scatterers) as a function of azimuth angle span using T72 tank (serial number) #132, with various location tolerances and definitions of persistence. The 'exact match' cases have the center of the scatterer pixel from the rotated image within the pixel boundaries of a corresponding scatterer. The 'within 1 pixel' cases allow the scatterer location to move into one of the 8 adjacent pixel locations. Note that while only 20% of the scatterer locations are invariant for an 'exact match' at 1° azimuth span , 65% of the scatterer locations are invariant 'within 1 pixel'. The cases labeled 'persists' in Figure 1 enforce the constraint that the scatterer exist for the entire span of angles and very few scatterers continuously persist for even 5°. 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 one foot ISAR results of Dudgeon et al. [5], whose definition of persistence allowed scintillation. Because of the higher scatterer location invariance with 1° azimuth span, we use azimuth models at 1° increments for each target, in contrast to others who have used 5° [11], 10° [7], and 12 models [12].

2.2 Invariance with object configuration

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. Figure 2(a) shows the location invariance of the strongest 40 scattering centers with articulation (turret rotated 315° vs straight) for T72 tank #a64 and similarly, Figure 2(b) shows the invariance for configuration variants: T72 tank #812 vs. #132. In both cases, while the invariance for an 'exact match' of scattering center locations averages less than 20%, it is over 55% for a location match

within a one pixel (3x3) neighborhood tolerance.

Figure 3(a) shows the probability mass functions (PMFs) for percent amplitude change for the strongest 40 articulated vs. non-articulated scattering centers of T72 tank #a64. 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. Figure 3(b) shows similar results for the configuration variants. These results demonstrate that for the quasi-invariant scattering center locations the magnitudes are also quasi-invariant for articulation and configuration variants. We have obtained similar results for a small (2°) change in depression angle.

3 SAR Recognition System

Our invariant-based recognition system uses standard non-articulated models of the objects (at 1° azimuth increments) to recognize the same objects in non-standard and articulated configurations. By modeling the object rotations (at 1° intervals), we have a simplifying special case of geometric hashing [10] with only translation and convenient integer buckets that correspond to radar range/cross-range bins. In this approach the relative positions of the scattering centers in the range and cross-range directions are indices to a look-up 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 2D version [2][8], using only the relative distances and the 'exact' scatterer locations, to the current 6D version that uses more local features and accommodate a 'within 1 pixel' scatterer location uncertainty. (The earlier 2D version gave much poorer results for the MSTAR data than the 6D version [1].) The detailed 6D model construction and recognition algorithms are given in [9].

In contrast to many model-based approaches [6] we are not 'searching' all the models; instead we are doing table look-ups 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 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 non-standard configurations. Thus, we model and use all the scattering center locations in turn as the origin, so the size of the look-up table models and the number of nominal relative distances considered in a test image is $n(n-1)/2$, 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° azimuth increments for each target type. The relative distances are the indices to a lookup table and, in the 2D version, each entry in the table is a list of labels that give the appropriate object type and azimuth. In the 6D version the model look-up table labels contain four
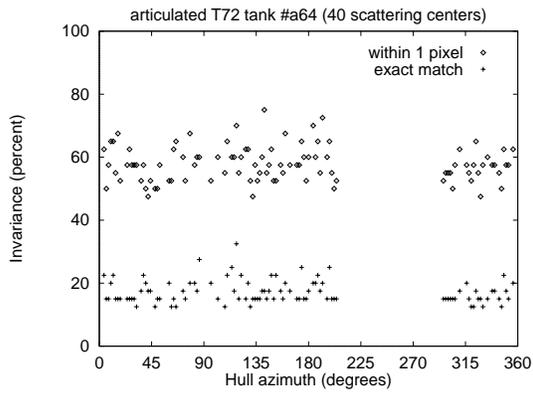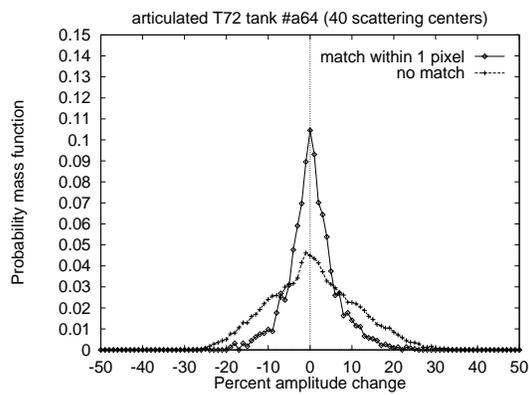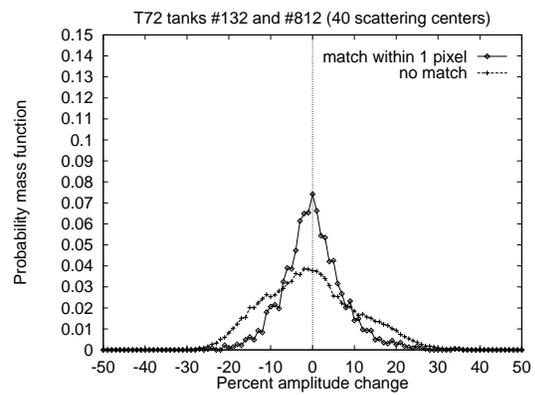
(a) Articulation.
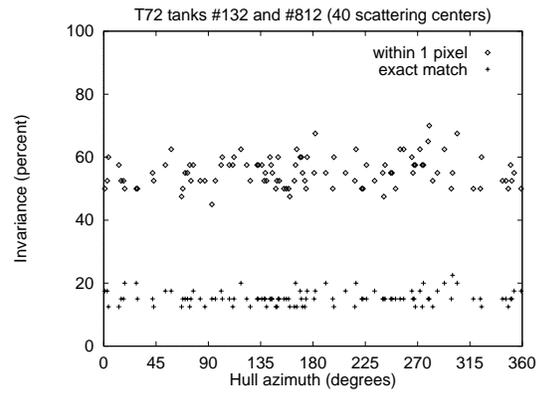
(b) Configuration.

Figure 2: T72 tank scatterer location invariance.

(a) Articulation.

(b) Configuration.

Figure 3: T72 tank scatterer magnitude invariance.

additional features: range and cross-range position of the 'origin' and the magnitudes 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 look-up table. In the 2D version of the recognition algorithm each query of the look-up table may directly generate votes for one or more potential candidate solutions. The 2D version accumulates votes in a 2D 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 due to articulation, configuration differences, or noise.

The 6D version of the recognition algorithm extends the basic 2D algorithm by adding additional features as constraints and accommodates a 'within 1 pixel' scatterer location uncertainty. In the 6D version, the comparison of the test data pair of scatterers with the model look-up table result(s) also provides information on the range and cross-range translation and the percent magnitude changes for the two scatterers. Limits on allowable values for translations and magnitude changes are used as constraints to reduce the number of false matches. Votes are accumulated in a 4D space: object, azimuth, range and cross-range translation. 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 results are accumulated for a 3x3 neighborhood in the translation subspace.

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. For the most difficult forced recognition case, configuration variants of the T72 and BMP2, Figure 4 shows the effect of the number of scattering centers used on the probability of correct identification (PCI) and Figure 5 shows the effect of varying the amplitude change limit. The optimum 6D system parameters for this case are using 36 scatterers, a translation limit of ± 5 pixels and a percent magnitude change of less than $\pm 9\%$. These same parameter settings are also used in the articulation and depression angle change results given in subsequent sections. (In other work [3], adaptive learning methods can be used to optimize the system parameters.)

To handle identification with 'unknown' objects, we introduce a criteria 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 with PCI vs. probability of false alarm (Pfa).

4 Experimental Results

4.1 Forced recognition results and pose accuracy

In the articulation experiments the models are non-articulated versions of T72 tank #a64 and ZSU23/4 #d08 (a radar and anti-aircraft gun turret on a tracked vehicle) and the test data are the articulated
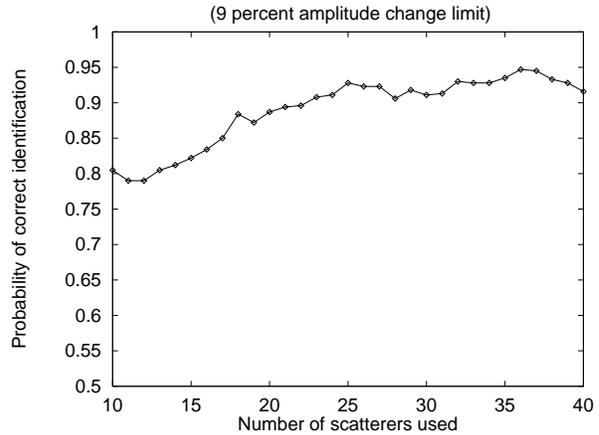

Figure 4: Effect of number of scattering centers used on recognition of MSTAR configuration differences.
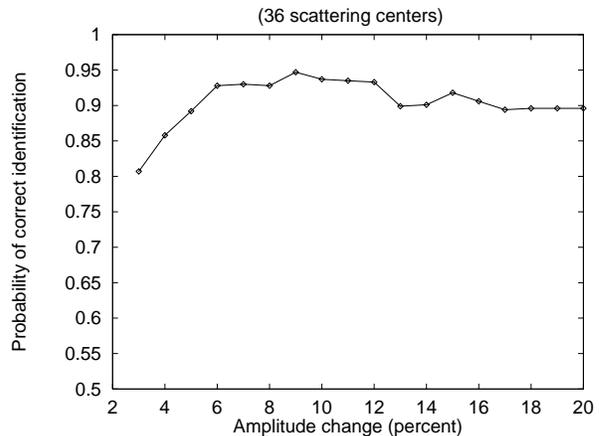

Figure 5: Effect of amplitude change tolerance on recognition of MSTAR configuration differences.

versions of these same serial number objects, all at a 30° depression angle. In the depression angle experiments the models are T72 tank #132 and BMP2 APC #c21 at a 15° depression angle and the test data are the same serial number objects at 17°. In both these experiments the test object and the model are the same physical object under different conditions and a PCI of 1.00 was achieved for forced recognition with articulation and 0.993 was obtained with the 2° depression angle change. In the more difficult configuration variant experiments, a single configuration of the T72 (#132) and BMP2 (#C21) vehicles are used as the models and the test data are two other variants of each vehicle type (T72 tanks #812 and #s7, as well as BMP2 APCs #9563 and #9566), all at a 15° depression angle. For the configuration variant cases a forced recognition rate of 94.7% is achieved with the 6D recognition system (which is a great improvement over the directly comparable 68.4% rate for the original 2D version of the system given in [2]).

Figure 6 illustrates the pose accuracy of the forced recognition configuration variant results. The top curve shows that 99% of the time the correct pose was achieved within $\pm 15^0$ (with a 180° front vs back direction ambiguity), while the correct object and pose were achieved 94% of the time with the directional ambiguity and 89% of the time with no ambiguity. Thus, the differences between the top and middle curves are the misidentifications, between the middle and the bottom are the cases where the direction is wrong by 180°.

Figure 7 is an example of the effect of object occlusion on forced recognition of the BMP2, BTR70, T72 and ZSU23/4, all at 15° depression angle. Since there is no real SAR data with occluded objects available for unrestricted use, occluded test data is simulated by starting with a given number of the strongest scattering centers and then removing the appropriate number of scattering centers encountered in order, starting in one of four perpendicular directions corresponding to the radar range and cross-range axes. Then the same number of scattering centers (with random magnitudes) are added back at *random locations* within the original bounding box of the chip. This keeps the number of scatterers constant and acts as a surrogate for some potential occluding object. Our approach, using simulated occlusion provides an enormous amount of data with varying known amounts of occlusion for carefully controlled experiments. This simulated occlusion and noise generate "invalid" scattering centers in a controlled manner that is similar to the uncontrolled variations generated by object articulation or configuration variations. Plots of recognition rate versus percent invariance with articulation or configuration variants are typically a mirror image of occlusion plots like Figure 7.

4.2 Confusers and ROC results

Figure 8(a) - (d) show scatter plot recognition results in BMP2–T72 vote space for configuration variants of the tracked BMP2 APC and the tracked T72 tank and for various confuser vehicles: the wheeled BTR70 APC #c71, the tracked ZSU23/4 gun #d08
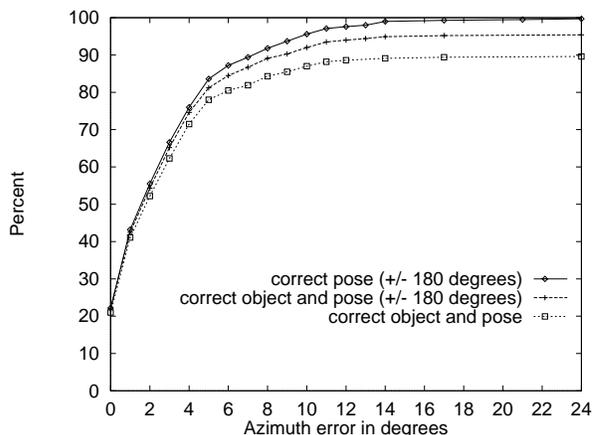

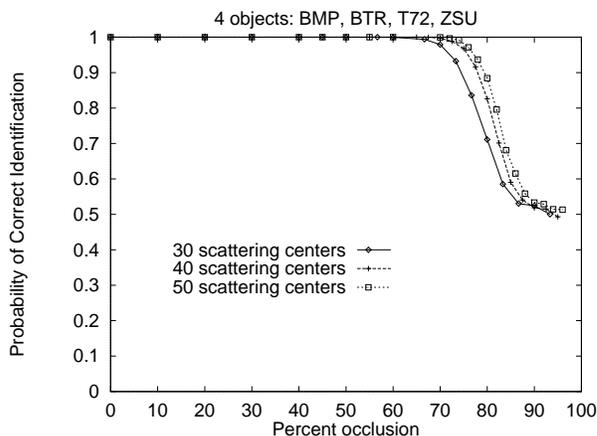Figure 6: Forced recognition pose accuracy for configuration variants.

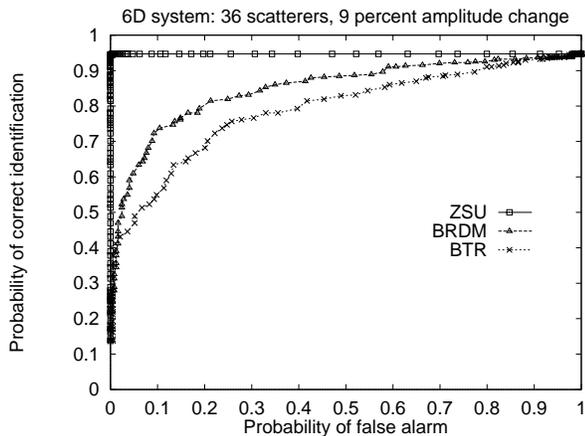Figure 7: Effect of occlusion on forced recognition results.

Figure 9: Effect of confusers on configuration varian t R OC cur ve.

Figure 10: R OC curves for articulation, depression angle and configuration variants.

and the wheeled BRDM2 APC #e71. The 45° line in Figure 8 represents the decision boundary of the simplest decision rule: "the object with the most votes wins". In this forced recognition case, Figure 8(a), the ov erall 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 Figures 8(b) - (d) the BTR70 APC is the most difficult confuser, the BRDM2 APC is somewhat less difficult and the ZSU23/4 gun is easy. F or example, Figure 8(b) illustrates that 99.6% of the BTR70 confuser false alarms could be eliminated with a 3000 vote threshold, but Figure 8(a) shows that a 3000 vote threshold would eliminate more than half of the BMP2 and T72 identifications. In con trast, Figure 8(d) shows that almost all of the ZSU23/4 confuser false alarms could be eliminated with a 1000 vote threshold without any reduction in the BMP2 and T72 identifications.

R OC curves can be generated from the scatter plot data in Figure 8 b y v arying the vote threshold (from 1000 to 4000 in 50 v ote incremen ts). Figure 9 shows the significant effect on the configuration variant recognition ROC curves of using the different ZSU, BRDM and BTR confusers whose scatter plot results are given in Figure 8. Excellent results are obtained with the ZSU23/4 gun confuser, while the BTR70 APC is a difficult case.

Figure 10 shows the ROC curve recognition results for the articulation, depression angle change and configuration varian ts cases, all with the 6D system using the same operating parameters. The ROC curves in Figure 10 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 are generated using the BTR confuser. The excellent results for the articulation case are basically due to the dissimilarity of the ZSU23/4 gun, T72 tank and BRDM2 APC.

4.3   Integration of multiple recognizers

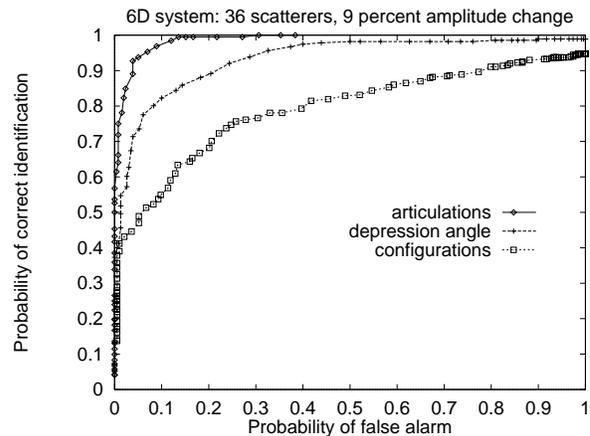Instead of tuning the parameters of a single recognizer to achiev e the optimum forced recognition per-

formance, the results of multiple recognition systems with different parameters can be in tegrated to produce a b etter o v erall result. Figure 11 shows the results of integrating 9 recognizers (operating in a 3x3 neighborhood about the optimum parameters of 36 scatterers and a 9 percent maximum percent magnitude change) for the most difficult case of configuration varian ts. The upper R OC curve, labeled "high confidence", represents the combined 90.3% of the T72 tank and BMP2 APC target cases and 70.1% of the BTR70 APC confuser cases where all 9 of the recognizers agreed. The second curve, labeled "combined results", represents the ov erall a v erage of all the multiple recognizer results (five or more of nine recognizers agree). It is a significant improvement ov er the best single system result, previously shown as the configurations result in Figure 10 as well as the BTR result in Figure 9 and now replotted and labeled as "best single result" in Figure 11. For example, at a 15% false alarm rate the combined result of multiple recognizers is a 79% recognition rate compared to a 65% rate for the best single recognizer. The "medium confidence" curve represents the 8.4% of the targets and 23.7% of the BTR70 APC confusers where 6 to 8 of the 9 recognizers agreed. The "low confidence" curve represents 1.3% of the targets and 6.2% of the BTR70 APC confusers where only 5 of 9 recognizers agreed.

5   Conclusions and F utureWork

The large v ariance in SAR scattering center location with object pose (azimuth) can be successfully captured b y modeling objects at small (e.g., 1°) azimuth increments. The locations and magnitudes of many scatterers are quasi-invariant with object configuration variations, articulations and small changes in radar depression angle. A model-based recognition system, using inexact match of the local features scatterer location and magnitude, can successfully handle difficult conditions with object configuration varian ts, articulation and occlusion with significan t forced recognition rates and excellent pose accuracy results. Some of the confuser vehicles are sufficiently similar to
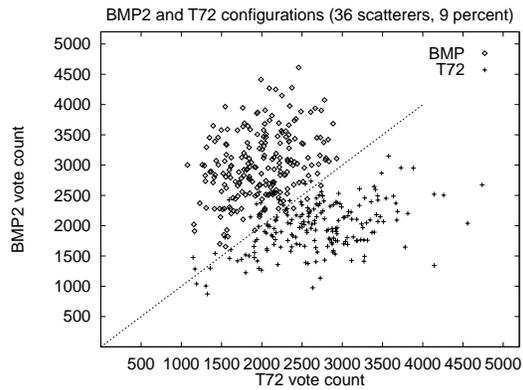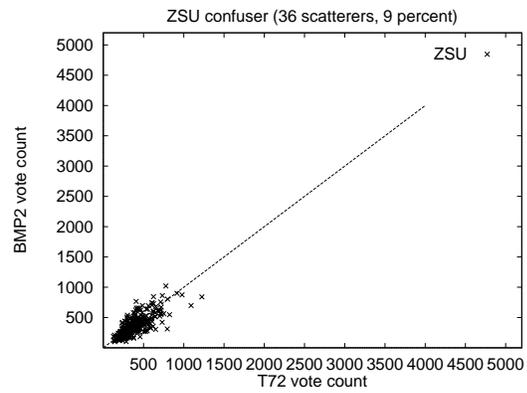
(a) BMP2 vs. T72.
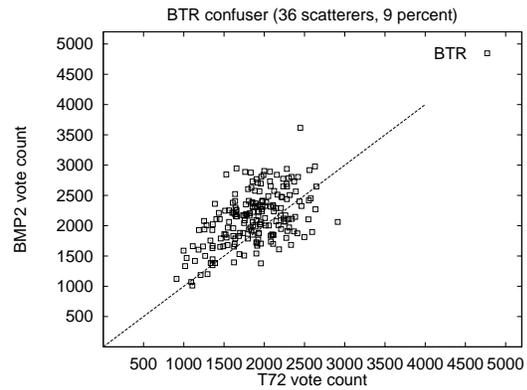
(b) BTR70 confuser.
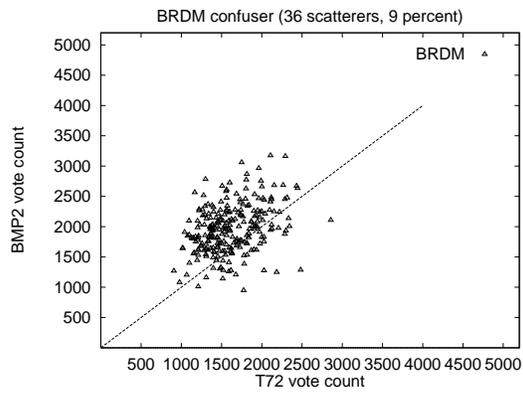
(c) BRDM2 confuser.

(d) ZSU23/4 confuser.

Figure 8: Scatter plots for 6D system results with configuration varian ts.
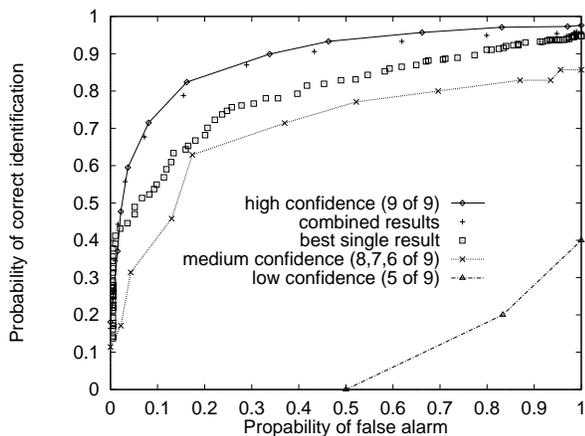
Figure 11: Effect of integrating multiple recognizers on ROC curves for configuration varian ts.

the objects of interest (e.g., BTR70 APC vs T72 tank and BMP2 APC) that they pose a muc h more severe challenge than other confusers (such as the ZSU23/4 gun). The physical differences in the configuration variant cases are a more difficult challenge than the articulation and depression angle cases which involv e the same physical object. Combining the results of multiple recognizers can give significantly improved performance over the single best recognizer.

The current w ork, which implicitly relies on the dissimilarity betw een different objects and on the invariance of the same object to changing conditions, can be extended to explicitly determine and utilize these measures for increased recognition performance. One approach is to explicitly discount model similarity betw een objects, as measured by feature space collisions. The intuition is that "ambiguous" features should be discounted. Another approach is to explicitly promote features that are invariant with changing conditions (e.g., configurations, articulations or small pose changes). Here the intuition is that the invariant features are more "reliable" and should be promoted. We plan to explore these ideas in the future.

Acknowledgments

References

[1] B. Bhanu and G. Jones III. "Recognizing T ar- get Variations and Articulations in SAR Images," *Optical Engineering*, in press.

[2] B. Bhanu, G. Jones III and J. Ahn. "Recognizing Articulated Objects and Object Articulation in SAR Images," *SPIE Proceedings: A lgorithms for Synthetic Apertur e R adar Imagery V*, vol 3370, pp. 493-505, Orlando, Fl, April 1998.

[3] B. Bhanu, Y. Lin, G. Jones and J. Peng. "Adaptive T arget Recognition," *Machine Vision and Application*, in press.

[4] M. Boshra and B. Bhanu. "Bounding SAR ATR P erformance Based on Model Similarity," *SPIE Pr oceedings: A lgorithms for Synthetic Apertur e Radar Imagery VI*, v ol 3721, pp. 716-729, Orlando, Fl, April 1999.

[5] D. Dudgeon, R. Lacoss, C. Lazott and J. Verly. "Use of persistent scatterers for model-based recognition," *SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery*, Vol. 2230, pp. 356-368, Orlando, FL, April 1994.

[6] W. E. L. Grimson. *Object R ecognition by Computer: The R ole of Geometric Constraints* The MIT Press, 1990.

[7] K. Ik euc hi, T. Shakunga, M. Wheeler and T. Y amazaki. "In v arian t histograms and deformable template matching for SAR target recognition," *Pr oc. IEEE Conf. on Computer Vision and Pattern Recognition,* pp. 100-105, June 1996.

[8] G. Jones III and B. Bhanu. "Recognition of Articulated and Occluded Objects," *IEEE T ansactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 7, pp. 603-613, July 1999.

[9] G. Jones III and B. Bhanu. "Recognizing Articulated Objects in SAR Images," *Pattern Recognition*, in press.

[10] Y. Lamden and H. Wolfson. "Geometric hashing: A general and efficient model-based recognition scheme," *Pr oc. International Confer ence on Computer Vision*, pp. 238-249, December 1988.

[11] L. Novak, S. Halversen, G. Owirka and M. Hiett. "Effects of P olarization and resolution on SAR ATR," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 33, No. 1, pp 102-115, January 1997.

[12] 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 Pr oceedings: A utomatic T arget R ecognition III*, V ol. 1960, pp 57-71, Orlando, FL, April 1993.

[13] E. Zelnio, editor. *SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery IV*, V ol. 3070, Orlando, FL, April 1997.

[14] E. Zelnio, editor. *SPIE Proceedings: Algorithms for Synthetic Apertur e Radar Imagery V*, V ol. 3370, Orlando, FL, April 1998.

[15] E. Zelnio, editor. *SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery VI*, V ol. 3721, Orlando, FL, April 1999.