Geometrical and Magnitude Invariants for Recognition of Articulated and Non-standard Objects in MSTAR Images

Grinnell Jones III, Bir Bhanu and Jun Guo*
College of Engineering, University of California, Riverside, CA 92521
{grinnell, bhanu, jguo} @vislab.ucr.edu
URL: www.vislab.ucr.edu

Abstract

Using SAR scattering center locations and magnitudes as features, invariances with articulation (i.e. turret rotation for the T72 tank and ZSU 23/4 gun), with configuration variants (e.g. fuel barrels, searchlights, wire cables, etc.) and with a depression angle change are shown for real SAR images obtained from the MSTAR public data. This location and magnitude quasi-invariance is used as a basis for an innovative SAR recognition engine that successfully identified real articulated and non-standard configuration vehicles based on non-articulated, standard recognition models. Identification performance results are given as confusion matrices and ROC curves for articulated objects, for configuration variants, and for a small change in depression angle with the MSTAR data. The recognition rate is related to the percent of location and magnitude invariant scattering centers.

1 Introduction

In this paper we are concerned with the problem of recognizing articulated vehicles and actual vehicle configuration variants in real SAR images. Our previous work in this area (see [1][2]) used the location invariance of scattering centers and was primarily based on simulated SAR data.

The key contributions of this paper are:

1. Demonstrates that quasi-invariant scattering center *locations* exist and that their *magnitudes* are also quasi-invariant for (a) articulation; (b) configuration variants; and (c) a depression angle change for actual vehicles in real SAR data.

2. Develops a new recognition engine based on scattering center location and magnitude features that achieves significant vehicle recognition performance for articulation, configuration variants and depression angle changes with real SAR data.

*Supported in part by DARPA/AFOSR grant F49620-97-1-0184; the contents and information do not reflect positions or policies of the U.S. Government.

2 Scatterer Location Invariance

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. A photo image of a typical MSTAR target, T72 tank serial number (#) a64, is shown in Figure 1. Regions of interest (ROI) are found in the MSTAR SAR target chips (as outlined in [1]) and the scattering centers are extracted from the SAR magnitude data (within the boundary contour of the ROI) by finding local eight-neighbor maxima. Example SAR images and the regions of interest (ROI), with the locations of the scattering centers superimposed, are shown in Figure 2 for baseline and articulated versions of the T72. Because the object and ROI are not registered, we express the scattering center location invariance with respect to articulation, configuration differences or depression angle changes as the maximum number of corresponding scattering centers (whose locations match within a stated tolerance) for the optimum integer pixel translation.

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

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

where $l = 0$ for an 'exact' match and $l = 1$ for a match 'within one pixel'.

The scatterer location invariance, L_n, of n scatterers, expressed as a percentage of matching points, is

Figure 1: T72 tank #a64.

given by:

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

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

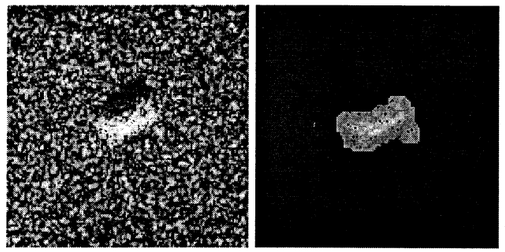
Figure 3 shows the location invariance, L_{40}, of the strongest 40 scattering centers with articulation for MSTAR T72 #a64 (at a 30° depression angle) as a function of the hull azimuth. The average invariance is 17.2% for an exact match of scattering centers and 57.8% for a location match within a one pixel (3x3 neighborhood) tolerance. Similarly, Figure 4 shows the percent of the strongest 40 scattering center locations that are invariant for T72 #812 vs. #132 (at a 15° depression angle). Figure 5 shows the percent scattering center location invariance for T72 #132 at 17^0 vs. 15° depression angles. The mean and standard deviation for percent location invariance (for 40 scatterers, and depression angle ϕ) are shown in Table 1 for articulated versions of the T72 and ZSU23/4, for configuration variants of the T72 and BMP2 and for depression angle changes with the T72 and BMP2.

3 Scatterer Magnitude Invariance

Using a scaled scatterer amplitude (S), expressed as a radar cross section in square meters, given by $S = 100 + 10\log_{10}(i^2 + q^2)$, where i and q are the components of the complex radar return, we define a percent amplitude change (A_{jk}) as: $A_{jk} = 100(S_j - S_k)/S_j$. (This form allows a larger variation for the stronger signal returns.) A location and magnitude match $Q_{jk}(t)$ is given by:

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

where l_A is the percent amplitude change tolerance. The scatterer magnitude and location invariance (I_n),



(a) baseline image. (b) baseline ROI.

(c) articulated image. (d) articulated ROI.

Figure 2: MSTAR SAR images and ROIs (with peaks shown as +) for T72 tank #a64 at 56° azimuth.

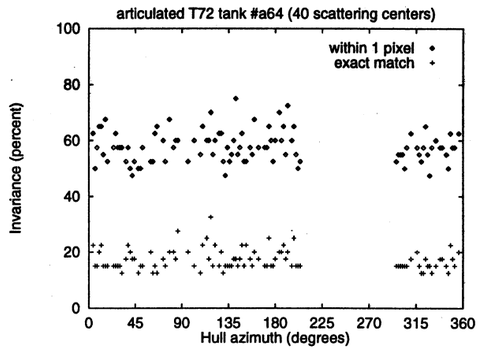


Figure 3: Articulated T72 scatterer location invariance.

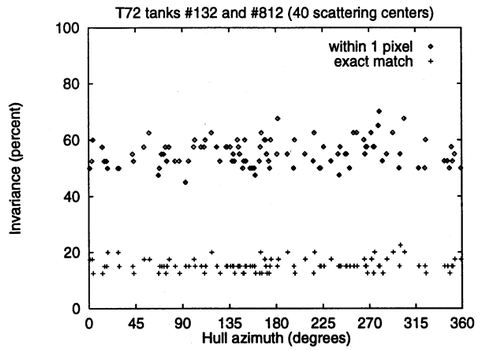


Figure 4: Example T72 scatterer location invariance with configuration.

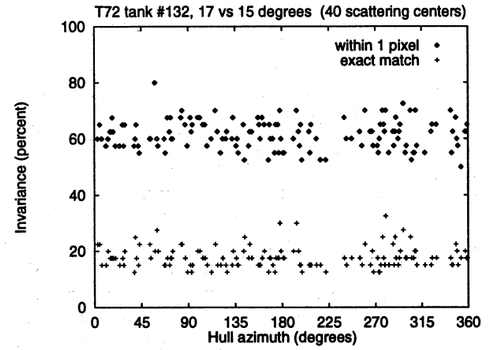


Figure 5: Example T72 scatterer location invariance with depression angle.

Table 1: Scatterer percent location invariance for MSTAR targets with articulation, configuration variants and depression angle changes.

	dep. angle	exact match invariance mean	exact match invariance s. d.	within 1 pixel invariance mean	within 1 pixel invariance s. d.
articulation:					
T72 #a64	30	17.17	1.47	57.83	2.23
ZSU #d08	30	15.69	0.91	55.05	1.72
average		16.45		56.47	
configuration variants:					
T72: #812 vs #132	15	15.34	0.89	55.34	1.91
#s7 vs #132	15	15.40	0.83	56.68	1.95
BMP2: #9563 vs #c21	15	16.34	0.84	58.52	1.97
#9566 vs #c21	15	16.17	0.99	57.93	1.97
average		15.83		57.15	
depression angle:					
T72 #132	17-15	17.76	1.52	61.55	2.05
BMP2 #c21	17-15	17.19	1.23	61.31	2.11
average		17.47		61.43	

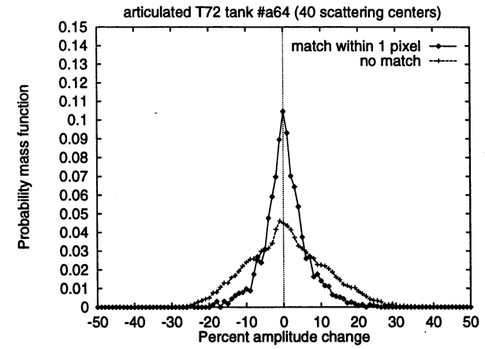


Figure 6: T72 scatterer percent amplitude change with articulation.

expressed as a percentage of n scatterers, is given by:

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

Figure 6 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 a one pixel tolerance and for all the combinations of scatterers whose locations do not match. For the cases with locations that matched within 1 pixel, the percent amplitude change mean and standard deviation are 0.51 and 5.91, while the non-matching cases are 0.75 and 10.44 respectively. The crossover points of the two curves are at -5 and +6 percent. Similarly, Figure 7 shows the PMFs for percent amplitude change for the strongest 40 scattering centers of T72 #812 vs. #132 (at a 15° depression angle), while Figure 8 shows this for 17° vs. 15° degrees depression angle (for T72 #132). The mean and standard deviation for these matching and non-matching scatterers and the crossover points for the PMFs are given in Table 2. Table 3 shows the mean and standard deviation for the percent location and magnitude invariance (within a 1 pixel location tolerance and an amplitude change tolerance of l_A) of the strongest 40 scatterers for these same articulation, configuration difference and depression angle change cases.

4 SAR Recognition Engine

Our new 6D model-based SAR recognition engine (the original 2D version is described in detail in [2]) uses standard non-articulated models of the objects

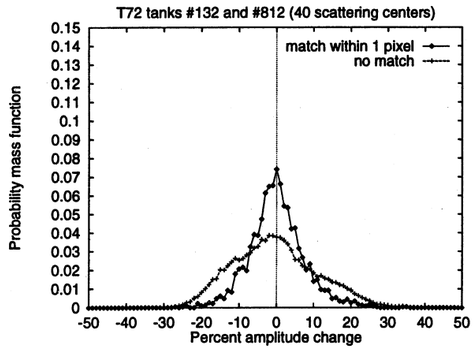


Figure 7: Example T72 scatterer percent amplitude change with configuration.

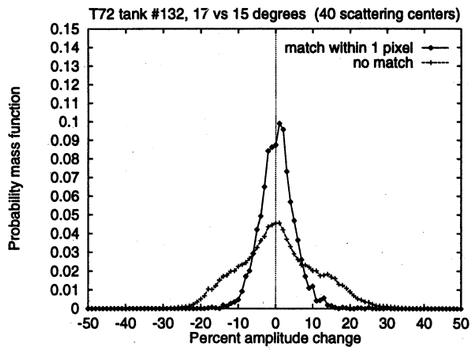


Figure 8: Example T72 scatterer percent amplitude change with depression angle.

Table 2: Scatterer percent amplitude change.

	within 1 pixel		no match		
	mean	s. d.	mean	s. d.	x-over
articulation:					
T72 #a64	.51	5.91	.75	10.44	-5/+6
ZSU #d08	.06	7.44	.08	11.37	±9
configuration variants:					
T72: #812 vs #132	.15	7.29	-.38	11.12	±8
#s7 vs #132	.48	6.69	2.20	11.15	±9
BMP2: #9563 vs #c21	.35	5.72	.94	10.88	-8/+9
#9566 vs #c21	.48	6.20	.56	10.68	-7/+8
depression angle:					
T72 #132	.43	4.66	.84	10.45	-7/+8
BMP2 #c21	.37	4.65	1.53	10.91	-7/+8

Table 3: Scatterer percent location and magnitude invariance (for locations within one pixel and amplitude tolerance l_A).

	l_A	mean	s. d.
articulation:			
T72 #a64	±9	53.47	2.63
ZSU #d08	±9	47.98	2.22
average		50.78	
configuration variants:			
T72: #812 vs #132	±9	48.40	2.42
#s7 vs #132	±9	50.69	2.44
BMP2: #9563 vs #c21	±9	54.38	2.34
#9566 vs #c21	±9	53.00	2.51
average		51.68	
depression angle:			
T72 #132	±7	56.15	2.38
BMP2 #c21	±7	55.66	2.53
average		55.91	

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 M scatterers.
4. Order (R, C, S) triples by descending S.
5. For each origin O from 1 to M do 6
6. For each point P from O+1 to M 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 9: Model construction algorithm

(at 1^o azimuth increments) to recognize the same objects in non-standard, articulated and occluded configurations. The model construction algorithm is outlined in Figure 9 and the recognition algorithm is given in Figure 10. Using a technique like geometric hashing, [3] the relative positions of the scattering centers in the range (R) and cross-range (C) directions are the (initial 2D) indices to a look-up table of labels that give the associated target type/pose and the remaining 4D features: range and cross-range position of the 'origin' and the magnitudes of the two scatterers. (The 'origin' is the strongest of a pair of scatterers, the other is a 'point'.) In comparing the test data with the model, the additional 4D information provides results on the range and cross-range translation and the percent magnitude changes for the two scattering centers. The number of scattering centers used, the limits on allowable translations and the limits on allowable magnitude changes are design parameters that are optimized, based on experiments, to produce the best forced recognition results.

1. Obtain from test image the location (R, C) and magnitude (S) of M strongest scatterers.
2. Order (R, C, S) triples by descending S.
3. For each origin O from 1 to M do 4
4. For each point P from O+1 to M 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, 10
8. weighted_vote = |DR| + |DC|.
9. Look up list of model entries at DR, DC.
10. For each model entry E in the list do 11
11. 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.
12. 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.
13. IF maximum vote_sum > threshold
THEN result is Object ELSE result is "unknown".

Figure 10: Recognition algorithm

The recognition process is an efficient search for positive evidence, using relative locations of scattering centers to access the look-up table and generate votes for the appropriate object, azimuth, range and cross range translation. A (city-block) weighted voting method is used to reduce the impact of the more common small relative distances. 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. The process is repeated with different scattering centers as reference points, providing multiple 'looks' at the model database to handle spurious scatterers that arise due to articulation, occlusion or configuration differences.

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 (typically from 1000 to 4000 votes in 50 vote increments) we obtain a form of Receiver Operating Characteristic (ROC) curve with probability of correct identification (PCI) vs. probability of false alarm (PFA).

Table 4: Example MSTAR articulated object confusion matrix (38 scatterers, ±9% amplitude tolerance, 2100 vote threshold).

MSTAR (Public) articulated test targets		Identification results		
		T72	ZSU	unknown
T72	315^0 turret	98	0	0
ZSU	315^o turret	0	92	2
BRDM2	(confuser)	32	0	222

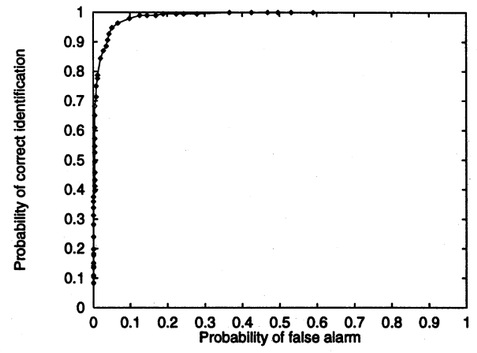


Figure 11: Receiver Operating Characteristics for recognizing MSTAR articulation.

5 Recognition Results

5.1 Articulated Object Results

Table 4 shows recognition results for articulated versions of the T72 #a64 and ZSU23/4 #d08 at 30^o depression angle, using the non-articulated versions of these same serial number objects as the models and BRDM2 #e71 as an "unknown" confuser vehicle. These results, 0.990 PCI at 0.126 PFA, are obtained using a 2100 vote decision criterion with 38 scatterers and with a ±9 percent amplitude change limit. (The overall forced recognition rate is 100% over a range from 14 to 40 scattering centers.) For the conditions in Table 4, varying the vote threshold results in the ROC curve shown in Figure 11.

5.2 Configuration Variant Results

Table 5 shows a typical forced recognition confusion matrix for configuration variants in the MSTAR data at 15^o depression angle, using a single configuration as the model (BMP2 #C21 and T72 #132) and testing with two other variants of each vehicle type. Although more extensive T72 configuration variant data is available, only two configurations are used so that

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

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

the amount of test data for the T72 and BMP2 is comparable and the results are not artificially biased toward recognizing the T72. The optimum forced recognition result is an overall rate of 94.7%, obtained at 36 scattering centers with a translation limit of ±5 pixels and a percent magnitude change of less than ±9 percent. (The 94.7% rate for this 6D recognition engine is directly comparable to the 68.4% rate for the prior 2D version of the recognition engine given in [1].) The effect on PCI of the number of scattering centers used is shown in Figure 12 (for $l_A = 9$) and Figure 13 shows the effect of varying the amplitude change limit (for 36 scattering centers). Using the BTR70 #c71 as an "unknown" confuser, for the optimum conditions given above, results in the ROC curve shown in Figure 14.

The effect of scatterer location and magnitude invariance on the forced recognition rate for configuration differences of the T72 and BMP2 is shown in Figure 15, based on the 22 failures in 415 tests shown in Table 5. (The two cases of perfect recognition below 35% are each a single instance with no failure.) These results with actual SAR data for the 6D recognition engine show over 90% recognition for location and magnitude invariance (within one pixel and an amplitude change tolerance of 9%) values down to 41.7%, compared to the prior results [1] for the 2D engine with simulated SAR data where the recognition rate drops sharply below 40% invariance for an exact match of locations.

5.3 Depression Angle Change Results

Table 6 shows the confusion matrix for recognizing T72 (#132) and BMP2 (#c21) at a 17° depression angle using models with the same serial numbers at a 15^0 depression angle and using BTR70 #c71 as an "unknown" confuser. These results, 0.855 PCI at 0.129 PFA, are obtained using a threshold of 2100 votes with 34 scatterers and with an amplitude change limit of ±7%. (The forced recognition rate is greater than

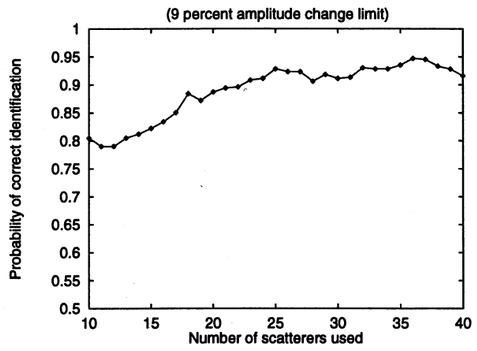


Figure 12: Effect of number of scattering centers used on recognition of MSTAR configuration differences.

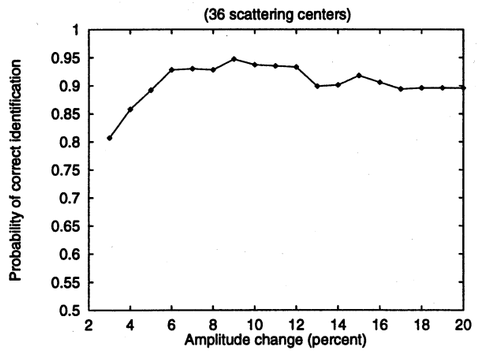


Figure 13: Effect of amplitude change tolerance on recognition of MSTAR configuration differences.

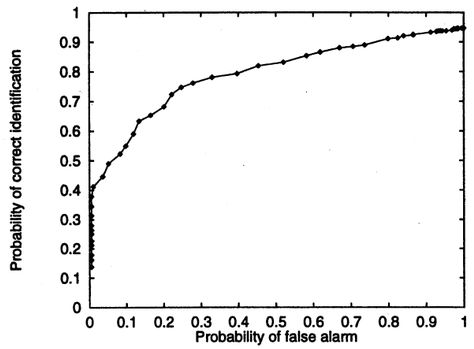


Figure 14: Receiver Operating Characteristics for recognizing MSTAR configuration differences.

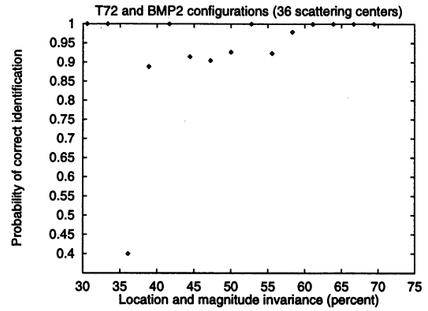


Figure 15: Effect of location and magnitude invariance on forced recognition for MSTAR configuration differences.

Table 6: Example confusion matrix for MSTAR depression angle changes (34 scatterers, ±7% amplitude tolerance, 2100 vote threshold).

MSTAR (Public) depression angle 17° test targets		Identification results [15° models]		
		BMP2	T72	unknown
BMP2	[#c21]	117	0	21
T72	[#132]	1	119	18
BTR70	(confuser)	10	20	202

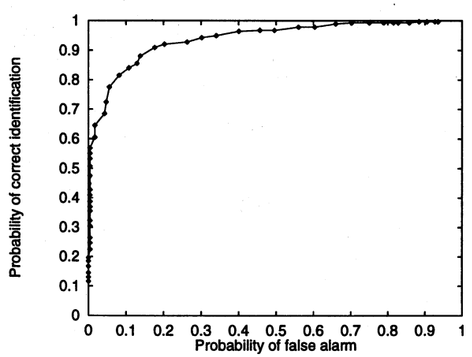


Figure 16: Receiver Operating Characteristics for recognizing MSTAR depression angle changes.

90% over a range of from 13 to at least 40 scatterers with the best forced recognition, 99.6%, at 34 scatterers with a ±7% amplitude change limit.) The ROC curve for these depression angle changes is shown as Figure 16.

6 Conclusions and Future Work

A significant percentage (56.5 - 61.4%) of the SAR scattering center locations are quasi-invariant (within a 3x3 pixel tolerance) for object articulation, configuration differences and small depression angle changes. The magnitudes of these quasi-invariant scatterers (expressed as a radar cross section) typically change by less than ±10%. The positions and magnitudes of pairs of these quasi-invariant scatterers are used in a 6D recognition engine to achieve good recognition results with real SAR data for object articulation, configuration differences and small depression angle changes. While these three problems are similar, the differences among configurations of an object type are a more significant challenge for recognition than articulation and depression angle changes, where the model and test data are the same physical object under different conditions (as seen by comparing the ROC curves in Figure 14 with Figures 11 and 16). These recognition results are a substantial improvement over the performance of the earlier 2D recognition approach with real SAR data. Future work to incorporate additional features in the recognition engine should lead to further performance improvements and accomodate combined cases such as configuration variants along with depression angle changes.

References

[1] B. Bhanu, G. Jones and J. Ahn. "Recognizing articulated objects and object articulation in synthetic aperture radar imagery," *SPIE Proceedings: Algorithms for Synthetic Aperture Radar Imagery V*, Orlando, FL, April 1998.

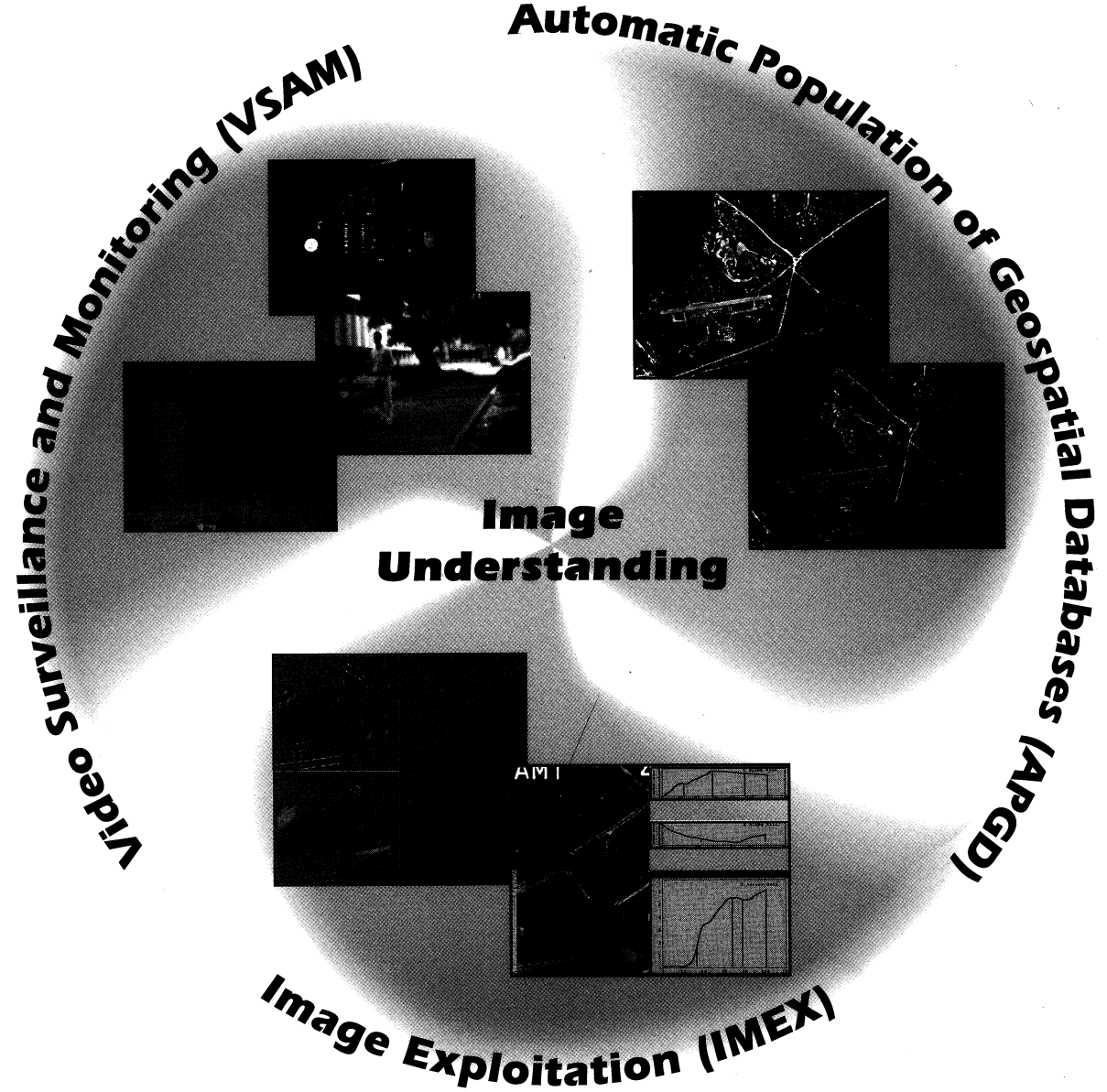
[2] G. Jones and B. Bhanu. "Invariants for the recognition of articulated and occluded objects in SAR images," *Proc. DARPA 1997 Image Understanding Workshop*, pp 1135-1144, New Orleans, LA, May 1997.

[3] 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.

Proceedings Volume II

1998 Image Understanding Workshop

Hyatt Regency Monterey
November 20–23
Monterey, California



Hosted by SRI International

Edited by
George E. Lukes

Sponsored by
Defense Advanced Research Projects Agency
Information Systems Office