# Invariants for the Recognition of Articulated and Occluded Objects in SAR Images

Grinnell Jones III and Bir Bhanu*
College of Engineering, University of California, Riverside, CA 92521
{grinnell, bhanu} @constitution.ucr.edu
URL: http://constitution.ucr.edu

## Abstract

Articulation invariant features are found (and quantified) in Synthetic Aperture Radar (SAR) images of military vehicles. They are used in the development of a SAR recognition engine that successfully identified articulated objects based on non-articulated recognition models. The engine also achieves robust recognition performance with mostly spurious data from noise or highly occluded objects. Performance results are related to the percent of invariant or unoccluded features.

## 1 Introduction

### 1.1 Problem Definition and Scope

Automated object recognition in SAR imagery is a significant problem because recent developments in image collection platforms will soon produce far more imagery (terabytes per day per aircraft) than the declining ranks of 2000 image analysts are capable of handling [10]. The specific challenges of this research are to address the need for automated recognition of military vehicles that can be in articulated configurations (such as: tanks, where the turret can rotate and the SCUD missile launcher, where the missile can erect) and can be partially hidden. Previous recognition methods involving template matching [11] are not useful in these cases, because articulation or occlusion changes global features like the object outline and major axis. In this paper the problem scope is the recognition subsystem itself, starting with SAR chips of target vehicles and ending with the vehicle identification. Because the very high resolution SAR target chips are not openly available, the U.S. Air Force provided the XPATCH high frequency radar signature prediction code [1], which is used to construct 4320 target chips for this research.

### 1.2 Overview of Approach

Our approach to object identification is specifically designed for SAR. The peaks (local maxima) in radar return are related to the physical geometry of the object. The relative locations of these scattering centers are independent of translation and serve as distinguishing features. The specular radar return varies greatly with the uncontrolled target orientation (azimuth). This azimuthal variance is captured by using 360 azimuth models. (The radar depression angle to the target is controllable, or known, and it is fixed at 15° for this study). Useful articulation invariants are found, which permit building non-articulated recognition models and using them to successfully recognize articulated targets. The SAR recognition engine, Figure 1, has an off-line model construction phase and an on-line recognition process. The recognition model is a look-up table that relates the relative distances among the scattering centers (in the radar range and cross range directions) to object type and azimuth. 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 (and azimuth).

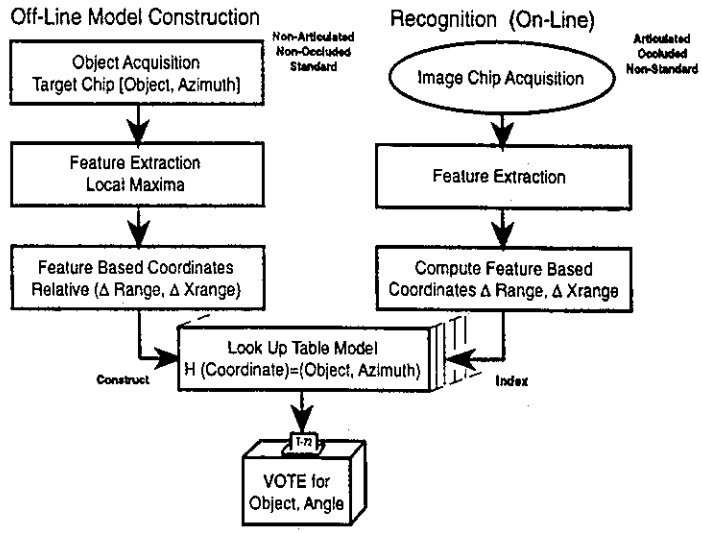



Figure 1: SAR recognition engine architecture.

### 1.3 Related Work and Our Contribution

A comparison of this approach, for the articulated and occluded object recognition problems in SAR, with related work is given in Table 1. Our approach is designed specifically for SAR, but is related to geometric hashing [8]. Scattering center relative positions are used as SAR recognition features. Template matching

*Supported in part by grants MDA972-93-1-0010 and DAAH04-95-1-0488; the contents and information do not reflect positions or policies of the U.S. Government.

Table 1: Related work comparisons.

| Area | Related Approach | This Approach |
|---|---|---|
| Indexing:<br>• transformation<br><br><br>• bin type/size | Geometric Hashing [8]:<br>translation<br>scale<br>rotation<br>real/varies | SAR specific:<br>translation<br>scale fixed in SAR<br>azimuth models<br>integer/fixed |
| SAR Recognition:<br>• azimuth models<br>• reference frame | Template Matching [11]:<br>12 blobs<br>global | Peak Locations:<br>360 constellations<br>local |
| Occlusion:<br>• azimuth models<br>• bin size<br>• near neighbor?<br>• search | Invariant Histogram [7]:<br>36 (10° apart)<br>2-4 feet<br>yes<br>exhaustive ($L_1$ norm) | Recognition Engine:<br>360 (1° apart)<br>6 inches<br>no<br>voting |
| Articulation | Constraint Models [2] [6] | invariants |

[11] is not suitable for recognizing articulated and occluded objects since there will be a combinatorial explosion of the number of templates with varying articulations and occlusions. The SAR recognition engine presented here has sufficient precision to perform both indexing and matching functions, while the invariant histogram technique (that has been applied to recognize occluded objects [7]) has limited performance, is only capable of indexing object models and requires a seperate template matching step. Constrained models of parts and joint articulation, used in optical images [2] [6], are not appropriate for the relatively low resolution, non-literal nature and complex part interactions of SAR images; which are handled by using articulation invariants as discussed in this paper. The major contributions of this paper are:

1. Identifies and quantifies articulation invariants.
2. Demonstrates a SAR recognition engine with robust performance for articulated and occluded objects.
3. Relates performance with invariance of features.
4. Quantifies azimuthal variance.

## 2 SAR Scattering Centers

The relative locations of peaks in the radar return are characteristic features that are related to the geometry of the object. The typical detailed edge and straight line features of man-made objects in the visual world, do not have good counterparts in SAR images for sub-components of vehicle sized objects at six inch to two foot resolution. The amplitude map, Figure 2, of a typical SAR target (SCUD launcher with missile erect, 18° azimuth, 15° depression) at six inch resolution shows a wealth of peaks corresponding to scattering centers and has no obvious lines or edges within the boundary of the vehicle. The 4320 target chips for the T72, T80, M1a1, FRED tanks and the SCUD missile launcher have a range of 52 to 284 local peaks. The locations of the peaks are related to the aspect and detailed geometry of the object. For example, for the T72 tank model, the strongest returns (that persist for 20° or more in azimuth span)

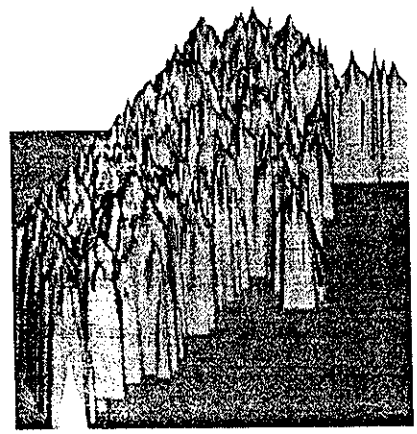
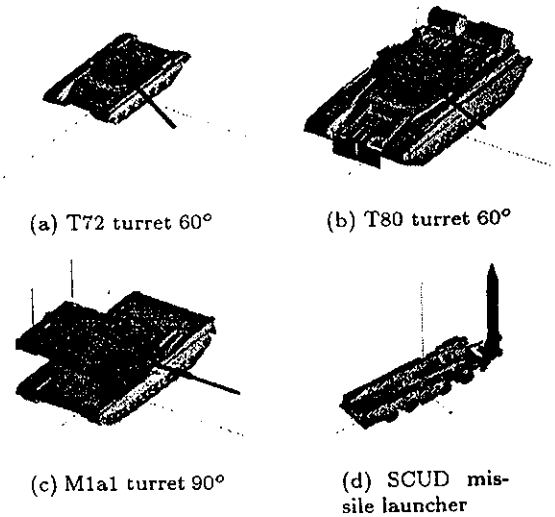


Figure 2: Example SAR image amplitude map.

are from four trihedral corners on the upper rear deck of the tank hull [3]. Figure 3 shows target geometries (model sizes in increasing order: T80, M1a1,T72 and SCUD launcher). The tank turret angle is measured counter-clockwise from the hull forward centerline.

### 2.1 Azimuthal Variance

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, because *significant* numbers of features *do not* typically persist over a few degrees of rotation. Averaging the results for 360 azimuths of the T72 tank, only about one-third of the 50 strongest scattering center locations (in object centered coordinates) remain unchanged (i.e. within an error radius of 1/2 pixel) for a 1° azimuth change (see Figure 4) and less than 5% persist for 10°. These are significantly less than the one foot resolution ISAR results of Dudgeon [5], whose



(a) T72 turret 60° (b) T80 turret 60°

(c) M1a1 turret 90° (d) SCUD missile launcher

Figure 3: Articulated objects (not to scale).

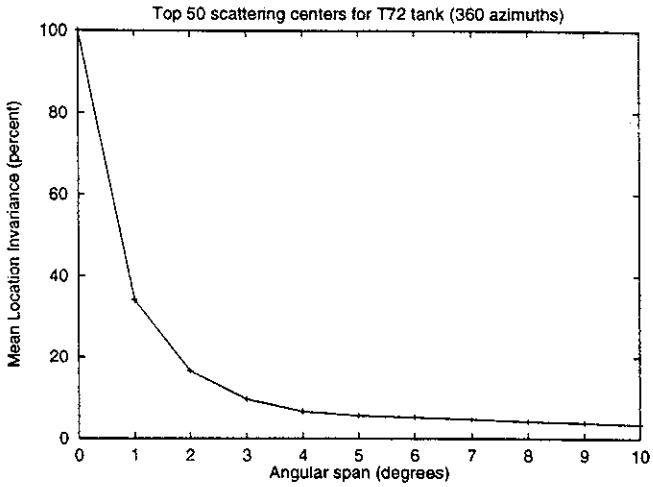
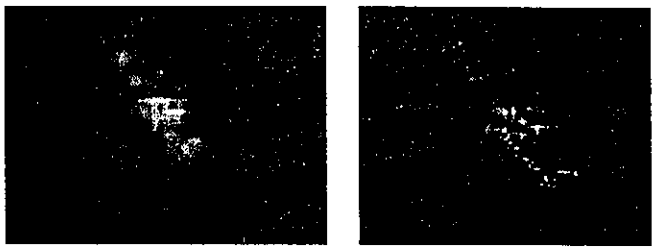



Figure 4: T72 azimuthal invariance.



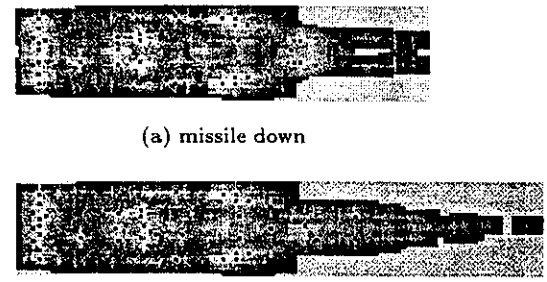
(a) two foot (b) six inch

Figure 5: SAR image resolution examples.

definition of 'persistence' allowed scintillation (i.e., a point was required to be present/absent for 2 consecutive angles, $1^o$ apart, to appear/disappear, thus a feature point would be 'persistent' if it appears and then disappears in images separated by $1^o$). Because of the presence of azimuthal variation and the goal of recognizing articulated and occluded objects, in this research we use 360 azimuth models ($1^o$ intervals), in contrast to others who have used $6^o$ [9] and $10^o$ [7] intervals and 12 models [11].

## 2.2 Image Resolution

A SAR image is formed by collecting backscattered field over a frequency band and over an angular span of incident directions. The resolution and scale of objects are fixed by the operating parameters of the radar beam: frequency, frequency bandwidth and angular span. Six inch resolution X-band images (10.0 GHz center frequency, 1.0 GHz bandwidth, $5.6^o$ angular span) are used to provide a rich feature set to facilitate the task of recognizing articulated and occluded targets. This provides 16 times as many pixels as two foot resolution. The comparison of the two foot resolution target 'blobs' with the six inch resolution constellation of image points is shown in Figure 5 for the FRED tank.



(a) missile down

(b) missile up

Figure 6: SCUD launcher articulation example.

(a) straight turret (b) $-60^o$ turret

Figure 7: T72 articulation example.

# 3 Articulation Invariants

The existence of articulation invariants in six inch resolution SAR imagery can be seen in Figures 6 and 7, where the locations of scattering centers are indicated by the black squares. In the example of the SCUD launcher, with the radar directed (from the left in Figure 6) at the front (cab end) of the launcher, many of the details from the launcher itself are independent of whether the missile is erect or down. In the similar view of the T72 tank, many of the details from the hull are independent of the turret angle. An example of articulation invariance is shown in Figure 8, which plots the percentage of the strongest 50 scattering centers for the T72 tank that are in exactly the same location with the turret rotated $60^o$ as they are with the turret straight forward, for each of 360 azimuths. The mean $\mu$, standard deviation $\sigma$ and $\mu - \sigma$ values of the average percent articulation invariance (for 360 azimuths) is shown in Table 2 for the individual articulated objects and the overall average. Comparing the cases for 25 and 50 scattering centers, the mean values are similar, but the $\mu - \sigma$ values are consistantly less for the 25 scatterer cases. The smaller average articulation invariance for the M1a1 tank is expected, because the M1 tank has a comparatively much larger turret than the other tanks (see Figure 3).

# 4 SAR Recognition Engine

## 4.1 Local Reference Coordinate System and Translational Invariance

Establishing an appropriate local coordinate reference frame is critical to reliably identifying objects (based on locations of features) in SAR images of ar-

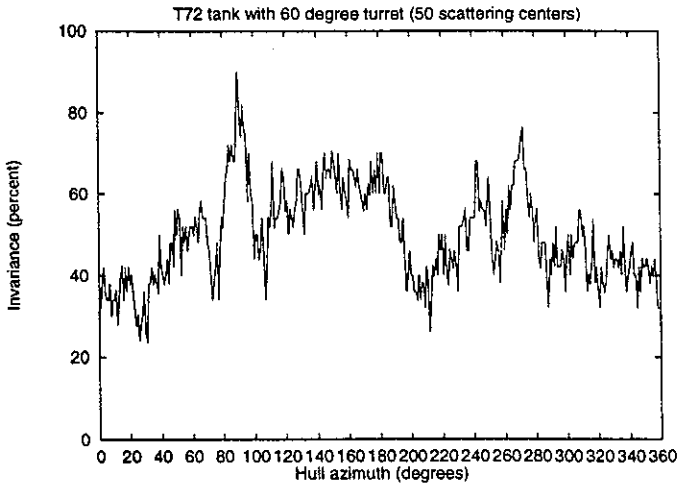



Figure 8: Articulation invariants example.

Table 2: Articulation invariance percentages.

| | 25 Scatterers | | | 50 scatterers | | |
|---|---|---|---|---|---|---|
| | $\mu$ | $\sigma$ | $\mu - \sigma$ | $\mu$ | $\sigma$ | $\mu - \sigma$ |
| SCUD missile up | 53.26 | 12.40 | 40.86 | 53.67 | 9.79 | 43.88 |
| T72 60°turret | 48.76 | 14.88 | 33.88 | 49.94 | 11.71 | 38.23 |
| 90°turret | 46.52 | 14.96 | 31.56 | 48.00 | 12.03 | 35.97 |
| M1a1 60°turret | 37.37 | 14.56 | 22.81 | 37.96 | 9.32 | 28.64 |
| 90°turret | 38.33 | 13.83 | 24.50 | 37.66 | 9.06 | 28.60 |
| T80 60°turret | 57.70 | 14.69 | 43.01 | 54.17 | 8.02 | 46.15 |
| 90°turret | 57.82 | 13.35 | 44.47 | 53.84 | 7.94 | 45.90 |
| average | 48.54 | 14.10 | 34.94 | 47.89 | 9.70 | 38.19 |

ticulated and occluded objects. The object articulation and occlusion problems require the use of a local coordinate system; global coordinates and global constraints do not work, as illustrated in Figures 6 and 7 where the center of mass and the principal axes change with articulation. In the geometry of a SAR sensor the 'squint angle', the angle between the flight path (cross-range direction) and the radar beam (range direction), can be known and fixed at 90°. Given the SAR squint angle, the image range and cross-range directions are known and any local reference point chosen, such as a scattering center location, establishes a reference coordinate system. (The scattering centers are local maxima in the radar return signal.) The relative distance and direction of the 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 the examples shown in Figures 6, 7 and 10 - 13 range is to the right ($x$ axis), cross-range is up ($y$ axis). The recognition engine takes advantage of this natural system for SAR, where a single basis point performs the translational transformation and fixes the coordinate system to a 'local' origin. For ideal data, picking the location of the strongest scattering center as the reference point is sufficient. However, for potentially corrupted data where any feature point could be spurious or missing (due to the effects of noise, articulation, occlusion, non-standard configurations, etc.), the process needs to continue with other scattering centers as the reference point to ensure a valid feature point is obtained as the origin. Although this implementation used all the scattering center locations in turn as reference points, heuristics could be applied to use fewer reference points for increased efficiency.

## 4.2 Recognition System Architecture

The basic system architecture, Figure 1, is an off-line model construction process and a similar on-line recognition process. The approach is based on local features and local reference coordinate systems. A systematic method is employed for constructing recognition models of objects that are not articulated, the models are stored in a look-up table and then local image features are used to index into the look-up table of models and recognize the same objects in articulated configurations. The image features used are the positions of the scattering centers (local maxima in the signal strength). The number of scattering centers used is a design parameter that is optimized based on experimental results. The positions of the scattering centers are expressed in relative distances in the known SAR range and cross-range coordinates. The model construction technique extracts these relative distances of the scattering centers from the non-articulated training data for all 360 azimuths for each target type. An example relative distance distribution for the SCUD launcher with the missile up is shown in Figure 9 (with the distances shifted by 154 range and 99 cross range to make the values non-negative). The model database is basically a table that relates these distances to object labels (target type and azimuth). The bounds of the table indices (and the shift amounts) are dictated by the relative distances between scattering centers of the largest target (the SCUD launcher with the missile up, establishes a 308 range by 198 cross-range table and the 154 range, 99 cross range shift for all the data). In the recognition phase, the relative positions of scattering centers are obtained from the articulated (or occluded) test data. These relative distances are indices into a look-up table which provides the associated object label(s) that are used to accumulate evidence for target identification. The process is repeated with different scattering centers as reference points, providing multiple 'looks' at the database. The target type and azimuth angle pair with the most 'votes' is chosen as the answer.

## 4.3 Algorithms and Complexity

A single test data or model (object, azimuth angle) instance $i$, with $M$ features is represented by relative range ($r$) and cross range ($c$) distance distributions given by:

$$a(r,c) = \begin{cases} k \\ 0 \end{cases} \quad (1)$$

where $k = 1$ for the models and $\sum k \leq M(M-1)/2$ (duplicates are removed), while for the test data $k \geq 1$ and $\sum k = M(M-1)/2$ (duplicates are allowed). In the recognition phase an object, azimuth model instance $i$ gets votes, $V_i$, as the intersection of the test data distance distribution and the distance distribu-

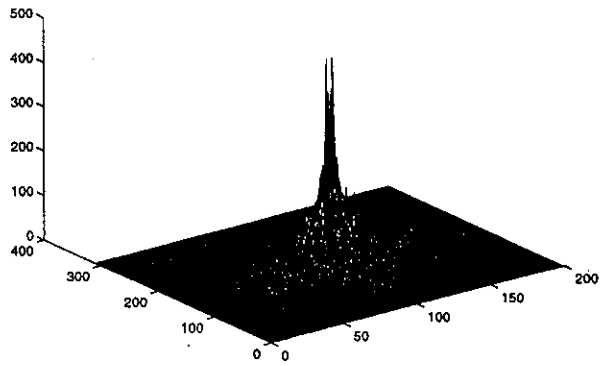



Figure 9: Relative distance distribution for SCUD launcher with the missile up.

tion $a_i$ for model $i$:

$$V_i = \sum_r \sum_c a_{\text{test}}(r,c) a_i(r,c). \quad (2)$$

The rule for object identification $I$ is given by:

$$I = \max\{V_i\} \text{ for all models } (i = 1 \text{ to } 360 \text{ N}); \quad (3)$$

where N is the number of objects. The recognition engine implements equations ( 1) to ( 3) as a look-up table and a decision rule.

The basic steps of the off-line model construction algorithm are:

1. For each of N Objects do 2:
2. For each of 360 Azimuth angles do 3,4:
3. Obtain the (row, column) locations of the strongest M peaks from image(Object, Azimuth),
4. For each of M Origins do 5:
5. For each Point from Origin + 1 to M do 6,7:
6. Find the relative Range and Cross-range positions of the Point from the Origin,
7. Add a new table entry at (Range, Cross-range) with Object, Azimuth label (add to a list if table already occupied).

The model construction algorithm complexity is 360 N M (M-1)/2, where N is the number of target types and M is the number of peaks used. The on-line recognition algorithm steps are:

1. Obtain the (row, column) locations of the strongest M peaks from test image.
2. For each of M Origins do 3:
3. For each Point from Origin + 1 to M do 4,5,6:
4. Find the relative Range and Cross-range positions of the Point from the Origin,
5. Look up the list of table entries at (Range, Cross-range),
6. Traverse the list: reading labels and incrementing Object, Azimuth label accumulators.
7. At completion, select the Object, Azimuth label with the largest accumulated total.

The recognition algorithm makes M (M - 1)/2 queries to the lookup table, where M is the number of peaks used. The only models associated with a lookup table entry are the "real" model and any other models that happen, by coincidence, to have a feature pair with the same relative distances.

## 5 Results

The XPATCH radar signature prediction code [1] is used to generate target chips at 360 azimuth angles (at a 15° depression angle, 90° squint angle) from the CAD models of the various objects. For 5 objects, the non-articulated set used for model construction is 1800 target chips. There are 7 sets of articulated test data (SCUD Launcher with the missile up, and the T72, M1a1, and T80 tanks with 60° and 90° turret angles) for an additional 2520 target chips. The scattering center locations are obtained at local maxima in the signal amplitude (where amplitude is greater than the surrounding eight neighbors, if equal reject unless last in raster scan order). Examples, at various azimuths, of the object geometry, SAR image and (strongest 50) scattering center locations are shown for both non-articulated and articulated cases of the T72 (Figure 10), M1a1 (Figure 11), T80 (Figure 12), and the SCUD launcher (Figure 13). (Figures 10 - 13 are not to scale and the image is displayed at 8 intensity levels, the scattering centers at 256 levels).

### 5.1 Articulated Objects

#### 5.1.1 Identification and Pose (4 Object Table)

A 4 object recognition table for the SCUD missile launcher, T72, M1a1 and T80 tanks is constructed from 1440 non-articulated target chips using the model construction algorithm given above. The experimental results of 2520 trials with articulated objects for the recognition engine using 50 scattering centers and this 4 object table are shown as a confusion matrix in Table 3. The overall performance is a 93.14% probability of correct identification (PCI). The azimuth accuracy is shown in Table 4, where 'e' is an exact match and 'c' indicates a match within ±5°. The azimuth results are reported for the hull angle. In the case of the M1a1 tank, decreased azimuth accuracy results when identifications are based on the (relatively large) turret rather than the hull.

#### 5.1.2 Number of Scattering Centers Used

The number of scattering center locations used is a design parameter that can be tuned to optimize performance of the recognition engine. For the objects and

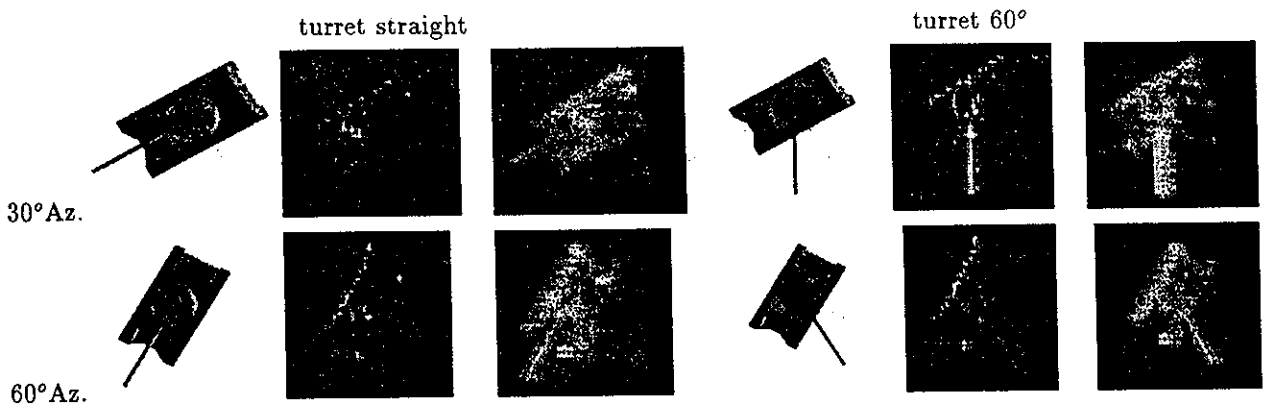



Figure 10: Examples of T72 tank geometry, SAR image and scattering centers for 30° and 60° hull azimuths.

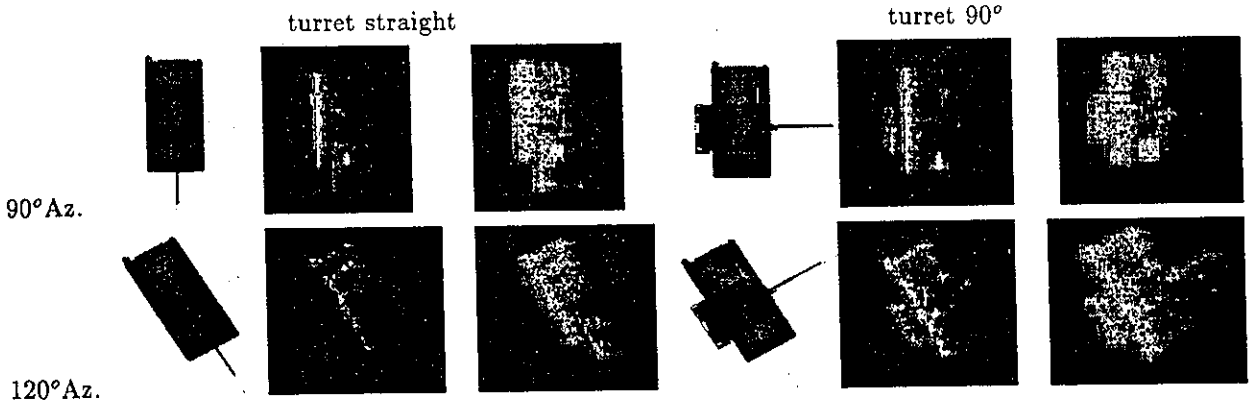



Figure 11: Examples of M1a1 tank geometry, SAR image and scattering centers for 90° and 120° hull azimuths.

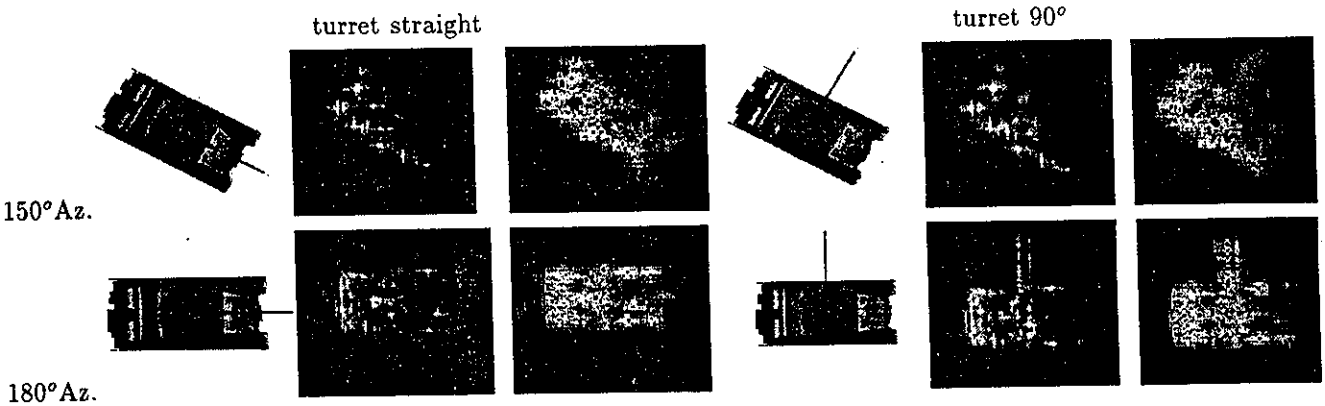



Figure 12: Examples of T80 tank geometry, SAR image and scattering centers for 150° and 180° hull azimuths.

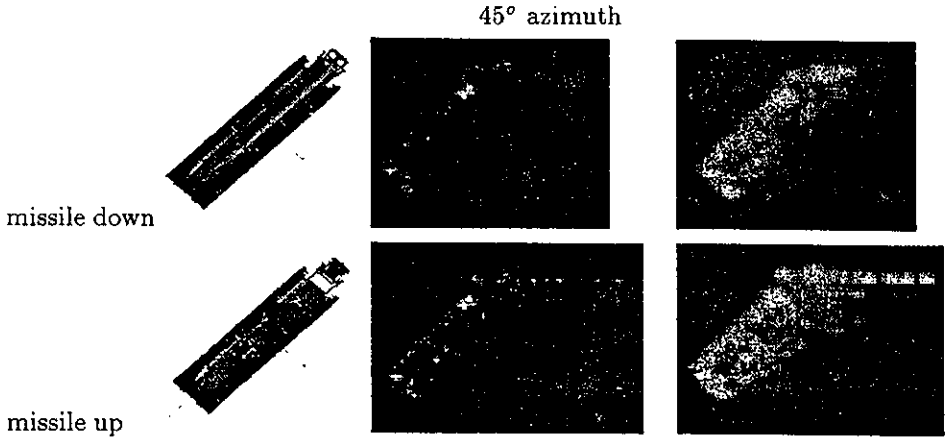



Figure 13: Examples of SCUD Launcher geometry, SAR image and scattering centers.

Table 3: Confusion matrix for articulated identification results (4 object table, 50 scattering centers).

| articulated test targets | Non-articulated models | | | |
|---|---|---|---|---|
| | SCUD down | T72 | M1a1 | T80 |
| | | 0° turret | | |
| SCUD missile up | 360 | | | |
| T72 60° turret | | 335 | 7 | 18 |
| 90° turret | | 327 | 8 | 25 |
| M1a1 60° turret | | 1 | 300 | 59 |
| 90° turret | | 2 | 305 | 53 |
| T80 60° turret | | | | 360 |
| 90° turret | | | | 360 |

Table 4: Confusion matrix for articulated pose accuracy results (e = exact pose, c = pose within ±5° ).

| articulated test targets | Non-articulated models | | | |
|---|---|---|---|---|
| | SCUD down | T72 | M1a1 | T80 |
| | | 0° turret | | |
| SCUD missile up | 360e | | | |
| T72 60° turret | | 333e | 1c | 2c |
| 90° turret | | 323e | | |
| M1a1 60° turret | | | 261c, 254e | 4c, 1e |
| 90° turret | | 2c | 272c, 261e | 4c, 1e |
| T80 60° turret | | | | 356c, 355e |
| 90° turret | | | | 360e |

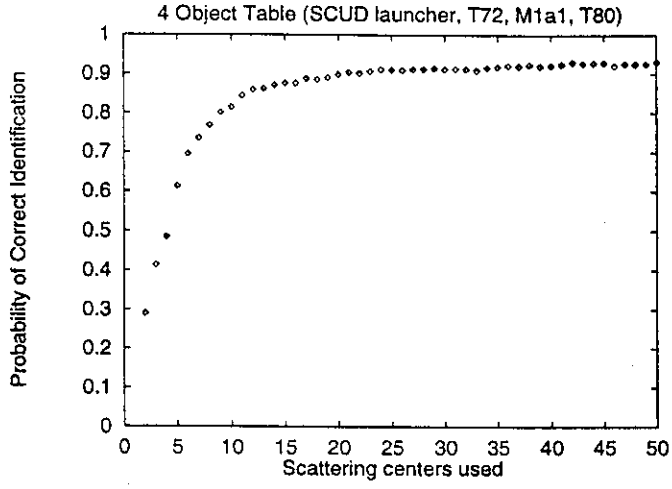



Figure 14: Effect of the number of scattering centers on articulated recognition rate.

articulations used in Tables 3 and 4, the plot of overall PCI vs number of scattering centers is shown in Figure 14 (each point is the result of 2520 trials). The maximum performance is achieved at 50 scattering centers (93.14%), but virtually the same performance could be found at 42 scattering centers (93.10%). A more optimal system with 35 scattering centers achieves similar performance, 92% PCI, with slightly less than half the storage and twice the speed of 50 scattering centers.

### 5.1.3 Articulation Invariance

The detailed recognition results can be related to the articulation invariance of articulated objects. The recognition failures for the T72 tank with the turret at 90° are plotted on the curve of percent invariance vs azimuth in Figure 15. These results show that recognition failures generally occur for azimuths where the percent invariance is low. Figure 16 shows how the PCI varies with the percent invariance. The points at low invariance values are misleading, because they

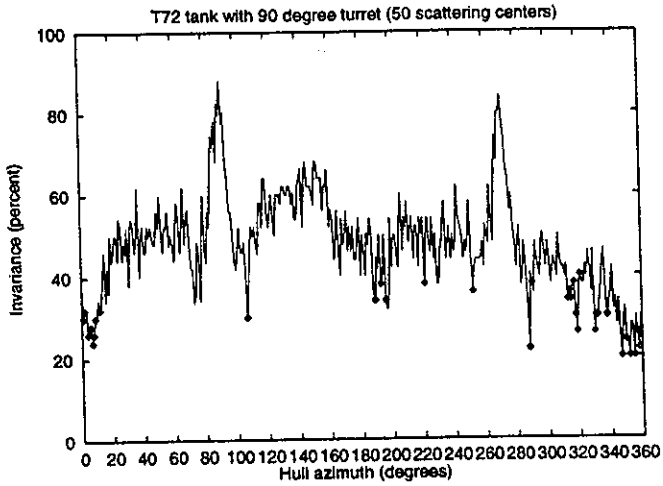



Figure 15: T72 tank (turret 90°) recognition failure plot on articulation invariance curve.

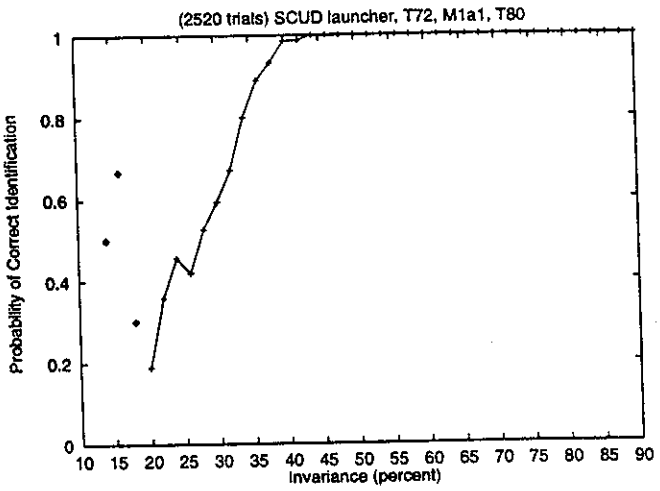



Figure 16: Recognition rate and articulation invariance (50 scatterers, average of 4 objects).

are due to a few correct identifications for the M1a1 tank, where the invariance (measured with respect to the hull) is low, yet a correct identification is made from features on the large turret. The overall recognition engine performance is almost perfect for invariance values greater than 40 percent (ie. down to 60 percent spurious data).

### 5.1.4 Identification (5 Object Table)

Table 5 shows results with a 5 object recognition table (50 scattering centers for each model), with the non-articulated FRED tank (which looks similar to the M1a1 tank, see Figure 17) added as a "confuser" in tests against the same 2520 articulated cases. In only four cases was a test object confused with the FRED tank: three times a T72 tank with 60° turret was now misidentified as a FRED tank, once a T72 tank with

Table 5: Confusion matrix for articulated identification results (5 object table, 50 scattering centers).

| articulated test targets | Non-articulated models | | | | |
|---|---|---|---|---|---|
| | SCUD down | T72 | M1a1 | T80 | FRED |
| | | 0° turret | | | |
| SCUD missile up | 360 | | | | |
| T72 60° turret | | 332 | 7 | 18 | 3 |
| 90° turret | | 327 | 7 | 25 | 1 |
| M1a1 60° turret | | 1 | 300 | 59 | |
| 90° turret | | 2 | 305 | 53 | |
| T80 60° turret | | | | 360 | |
| 90° turret | | | | 360 | |

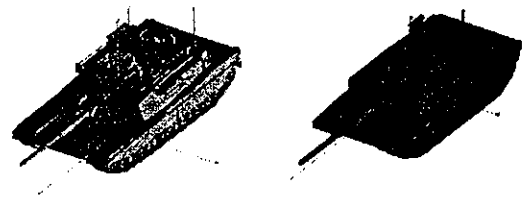


Figure 17: M1a1 (left) and FRED (right) tanks.

90° turret, that had been misidentified as an M1a1 tank, was now misidentified as a FRED tank. The overall PCI for the 5 object table (with 50 scattering center models) was 93.02% versus 93.14% for the 4 object table.

## 5.2 Occluded Objects

The 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 $d_i$ (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. Each data set is 5760 test cases (4 objects $X$ 4 directions $X$ 360 azimuths). For the non-articulated occluded tests (the objects are the T72, M1a1, and T80 tanks with turret at 0° and the SCUD launcher with the missile down) there are 51 data sets (for 10, 30 and 50 scattering centers with 10 to 90% occlusion in 10% steps and the same for 20 and 40 scatterers plus 55, 65, and 75% occlusion) for a total of 293,760 test cases. Actually, only 50 data sets with a total of 288,000 test cases are used, because the data set of 10 scattering centers with 90% occlusion has less than two valid scattering centers for each case. For the articulated occluded tests the same tanks are used with a 90° turret and the missile is erect, but there are only 9 data sets (for 20 scattering centers with 10 to 90% occlusion) for a total of 51,840 test cases.

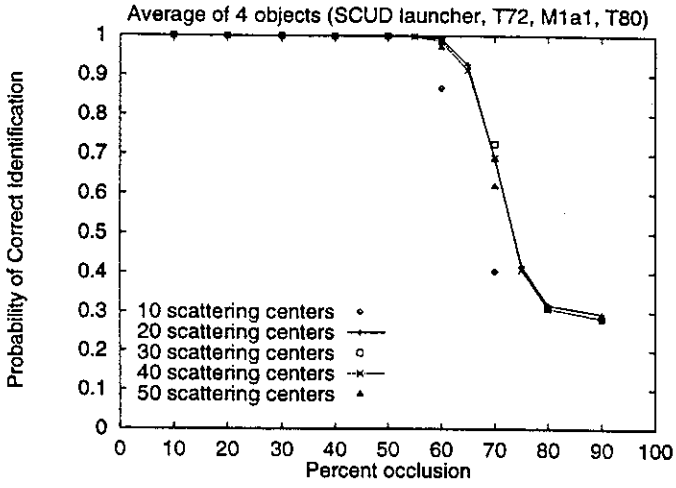



Figure 18: Recognition rate and occlusion percent.

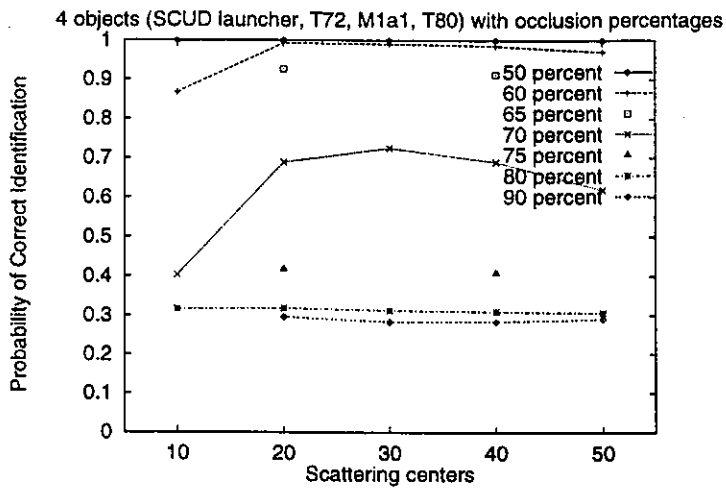



Figure 19: Effect of number of scatterers on occluded recognition rate.

### 5.2.1 Non-articulated Occluded Objects

The performance of the recognition engine with non-articulated occluded objects is shown in Figure 18 in terms of the probability of correct identification (PCI) as a function of percent occlusion with the 'number of scattering centers used' as a parameter. The results of 288,000 test cases are shown, where each point for a specific percent occlusion and number of scattering centers is the average PCI for all 4 occlusion directions, the 4 objects and the 360 azimuths (5760 test cases). The overall recognition engine performance is almost perfect for up to 60% occlusion. (This corresponds to results shown in Figure 16 for articulation invariance of 40% and above.) By 80 to 90% occlusion, the results are not much better than the 0.25 PCI one would expect by chance from the 4 possible objects. These performance results are replotted as Figure 19 to illustrate the effect of the number of scattering centers used on the recognition rate for the highly occluded cases. This indicates that optimal performance is in the region of 20 to 40 scattering centers.

### 5.2.2 Articulated Occluded Objects

Figure 20 shows the average and individual test object performance of the recognition engine (using 20 scattering centers) as a function of percent occlusion with 4 different articulated objects. The results of 51,840 test cases are shown. The overall performance for these articulated objects with 30% occlusion is a 0.698 PCI, almost the 0.70 system level goal [4] of the Moving and Stationary Target Acquisition and Recognition (MSTAR) program. The results are consistent with the average unoccluded articulated results for 20 scatterers, shown previously in Figure 14, which would be a 0.899 PCI at a "0%" occlusion in the occluded articulated results shown here in Figure 20. Figure 21 compares the performance results of the articulated and occluded articulated objects for cases with the same number of valid scatterers (i.e. 'scatterers used' in the unoccluded cases or 'unoccluded scatterers' in the occluded cases). In the occluded data the valid points are 'clustered' in a neighborhood which gets smaller as the occlusion increases (and the number of valid scatterers decreases). These relatively worse results for the naturally 'clustered' occluded articulated data, compared with the more widely distributed unoccluded articulated data (for the same number of valid scattering centers), illustrate the importance of the relatively rare long distances.

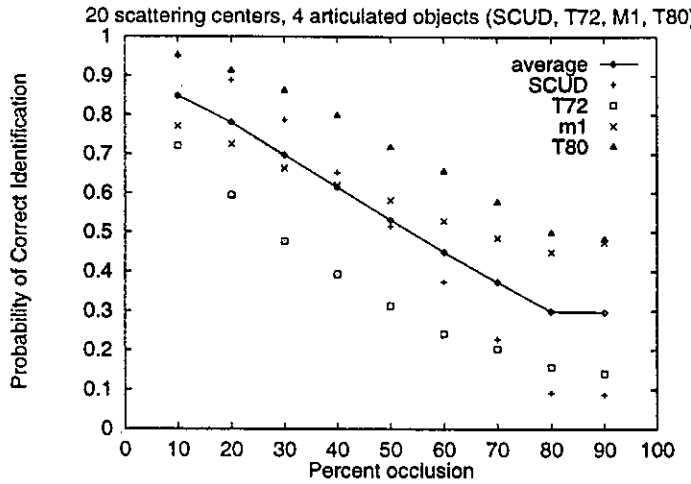



Figure 20: Articulated object recognition rate and occlusion percent.

## 6 Conclusions

The XPATCH generated, six inch resolution SAR imagery has great azimuthal variation that can be successfully captured by using 360 azimuth models for a given depression angle. Useful articulation invariant features are found in SAR images of military vehicles. The feasibility of a new concept for a SAR recognition engine to identify articulated and occluded objects based on non-articulated recognition models is

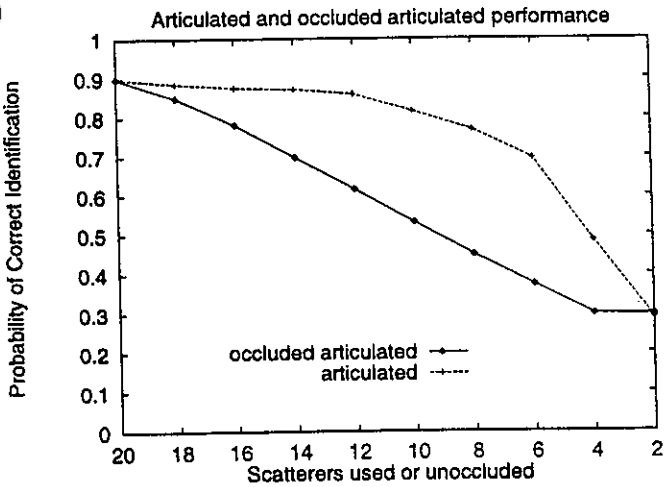



Figure 21: Articulated object and occluded articulated object performance results.

demonstrated. The performance of the recognition engine can be predicted by the percent articulation invariance (or percent unoccluded) when comparing the scattering center locations of the articulated (or occluded) test images with the non-articulated model scattering center locations. Limited experiments show that scaling to model more objects provides similar results, although performance will degrade depending on the number of coincidental similarities found in the radar signatures of the objects. Our results indicate the importance of the relatively rare long distances and suggest an explanation why this approach, which can use long distances (if available), could have an advantage over others [7] that are restricted to a "neighborhood".

Use of real SAR images of actual vehicles (vs XPATCH simulations from CAD models) would change the performance and detailed implementation of the design, but not the basic conceptual approach. Our approach of articulation invariance simply treats the articulated region as a "don't care", which applies to both real and simulated data. If real SAR images are more (or less) persistent in azimuth than XPATCH, then the recognition engine would need fewer (or more) azimuth models. Real SAR images and target chips are likely to have more noise than the ideal models and test data produced by XPATCH, however larger sets of noisy data can be used to produce useful recognition models. The noisy test data is manifest as some percentage of spurious data, which is similar to what was used to generate the occluded test data and the actual recognition results should suffer as indicated in Figures 18 and 20 on the Probability of Correct Identification vs percent occlusion curves (with the corresponding source of the invalid scattering centers being noise rather than 'occlusion').

## Acknowledgment

The authors would like to thank Thomas Ferryman for his helpful suggestions.

## References

[1] D. Andersch, S. Lee, H. Ling, and C. Yu. "XPATCH: A high frequency electromagnetic scattering prediction code using shooting and bouncing ray," *Proceedings of Ground Target Modeling and Validation Conference*, pp.498-507, Aug.1994.

[2] A. Beinglass and H. Wolfson. "Articulated object recognition, or: How to generalize the generalized Hough transform," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp.461-466, June 1991.

[3] B. Bhanu, G. Jones, J. Ahn, M. Li, and J. Yi. "Recognition of articulated objects in SAR images," *Proc. ARPA Image Understanding Workshop*, pp 1237-1250, Palm Springs CA, Feb 1996.

[4] T. Burns. "Moving and Stationary Target Acquisition and Recognition," *DARPA Image Understanding Technology Program Reviews*, Ft. Belvoir, VA, Sep.1996.

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

[6] Y. Hel-Or and M. Werman. "Recognition and localization of articulated objects," *Proc. IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, pp. 116-123, Austin, TX, Nov 11-12, 1994.

[7] K. Ikeuchi, T. Shakunga, M. Wheeler, and T. Yamazaki. "Invariant histograms and deformable template matching for SAR target recognition," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 100-105, June 1996.

[8] 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.

[9] S. Raney. "Automatic radar air to ground target acquisition program (Aragtap)." *ARPA BAA95-03: Moving and Stationary Target Acquisition and Recognition Algorithm Development, Briefing to Industry*, October 21, 1994.

[10] T. Strat. "Image understanding technology programs," *DARPA Image Understanding Technology Program Reviews*, Ft. Belvoir, VA, Sep.1996.

[11] 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 Target Recognition III*, vol 1960, pp. 57-71, Orlando, FL, April 1993.