ELSEVIER

Pattern Recognition Letters 23 (2002) 1019–1029

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Model-based recognition of articulated objects

J.S. Ahn, B. Bhanu *

Center for Research in Intelligent Systems, University of California, B232 Bourns Hall, Riverside, CA 92521-0425, USA

Received 17 May 2001; received in revised form 11 September 2001

Abstract

This paper presents a model-based matching technique for recognition of articulated objects (with two parts) and the poses of these parts in synthetic aperture radar (SAR) images. Using articulation invariants as features, the recognition system first hypothesizes the pose of the larger part and then the pose of the smaller part. Geometric reasoning is carried out to correct identification errors. The thresholds for the quality of match are determined dynamically by minimizing the probability of a random match for the recognition system. Results are presented using both occluded synthetic articulated object SAR signatures and actual signatures of articulated objects from the real-world data. The system performance is evaluated with respect to identification performance and accuracy of estimates for the poses of the object parts.

Keywords: Articulated objects; Object recognition; SAR images

1. Introduction

Recognition of objects in synthetic aperture radar (SAR) imagery is an active area of research in pattern recognition and computer vision (Jones and Bhanu, 1999; Novak et al., 1994). In this paper we focus on the problem of recognizing articulated objects (with two parts) and the poses of the articulated parts in SAR imagery. Previous work in the area of recognizing articulated objects in visual imagery (Beinglass and Wolfson, 1991; Hel-Or and Werman, 1994), has used simple models (like scissors and lamps) and has used constraints around a joint to recognize these objects. For SAR images, it is difficult to extract linear features (commonly used in visual images), especially at six inch to a foot resolution. This is because of the unique characteristics of SAR image formation (specular reflection, multiple bounces, low resolution and non-literal nature of the sensor). Previous recognition methods for SAR imagery using templates (Novak et al., 1994) or boundary contours are not suitable for the recognition of articulated objects, because articulation or occlusion changes the object outline and each different articulation configuration requires a different template leading to a combinatorial explosion.

The key contribution of the paper is the recognition of articulated objects and the articulation

* Corresponding author. Tel.: +1-909-787-3954; fax: +1-909-787-3188.

E-mail addresses: ahn@cris.ucr.edu (J.S. Ahn), bhanu@cris.ucr.edu (B. Bhanu).

0167-8655/02/$ - see front matter
PII: S0167-8655(02)00033-8

of the objects in SAR imagery using models based on articulation invariants. In contrast to other work on recognizing articulated objects in SAR images (Bhanu and Jones, 2000; Jones and Bhanu, 1999, 2001), this paper presents an approach that explicitly models both parts and, thus, can determine the poses of *both* parts.

An *end-to-end system* has been developed whose input is a target chip (region-of-interest, ROI) and the final result is the identification of the target and the poses of its parts. The system has been extensively tested using both simulated and publically available real-world SAR image data.

2. System overview

Fig. 1 depicts an end-to-end system for recognizing articulated objects. It shows the matching components including the feedback (loop) used in geometric reasoning from the matching module to the indexing module that generates all the hypotheses.

The basic assumption behind our approach is that more scattering centers (local peak features in the SAR image) are from the larger part than from the smaller one so that the models for the larger part are used first to find the target identification (ID) and the pose of the larger part. As an example, tanks have two parts: body (larger part) and turret (smaller part). For some targets like the M1 tank, occasionally more scattering centers come from the smaller part (turret) than from the larger one (body). If the system succeeds in recognizing the target and its body pose, those scattering centers used for the body part recognition (the positives) are eliminated from the test data. The remaining scattering centers (called the negatives) are supplied to the next stage to recognize the turret part and its pose. If the system fails to recognize the target ID and its body pose, the system tries to recognize the target ID and its turret pose instead. After the recognition of the turret, it tries to recognize the body pose based on the negative scattering centers, which are left over from the turret part recognition.

3. Invariant feature extraction and building part models

Fig. 2 shows examples of XPATCH (Andersh et al., 1994) generated synthetic SAR images (at six inch resolution, a fixed depression angle of 15° and squint angle of 90°) for a T72 tank with different turret articulations. The SAR scattering centers are local peaks in the amplitude of the radar return (eight-neighbor maxima) and in Fig. 2 their range and cross-range locations are marked as black squares. Fig. 2 illustrates that some of the scattering locations are invariant to the turret articulations. The invariant scattering centers are

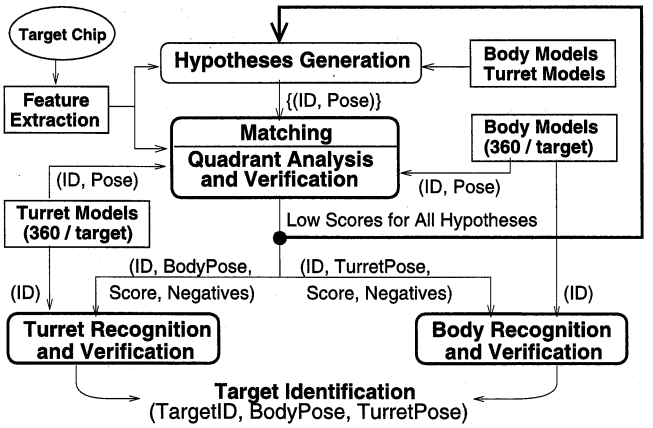

Fig. 1. An end-to-end system for recognizing articulated targets.

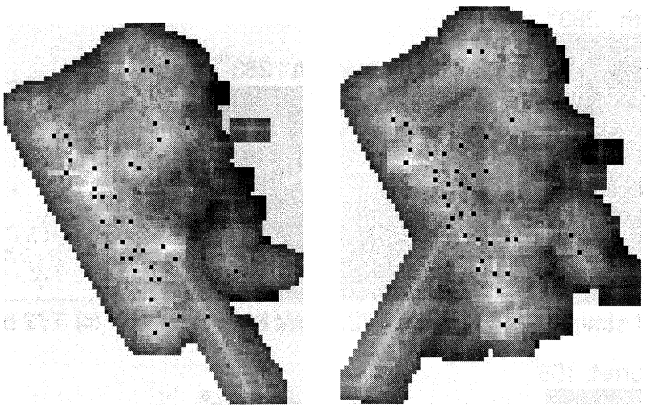

Fig. 2. SAR images of T72 (body at 240°, 40 scattering centers marked): (a) turret at 240°; (b) turret at 300°.

defined as those that remain in exactly the same positions as a result of articulation. Details for real SAR MSTAR images are given in Section 5.2.

Table 1 shows the percentage of articulation variance for each target with turret 30°, 60° and 90° rotated versus turret straight. The articulation variances for XPATCH generated M1, T72, and T80 tanks are 63%, 52%, and 47%, respectively. For model building and experimentation, we use the invariants to build body models and a subset of variants (caused by turret articulation or inter-reflections) to build turret models. We do not fix the number of scattering centers to be extracted from the target chips because we want to get as many features as possible from both body and turret.

To build a body model (at some azimuth), we use two different turret articulations with the body at the same azimuth. Fig. 3 shows examples of model building for the body and turret using XPATCH data. Both (a) and (b) show the T72 tank with a body pose at 283° but different turret poses. Using scattering centers from both (a) and (b), the body model (positives) of the T72 tank at 283° is shown in (c). Similarly, both (d) and (e) show the M1 tank with a turret pose at 105° but different body poses. The corresponding turret model for the M1 tank at 105° is shown in (f).

4. Hypotheses generation and verification

In the recognition phase, a set of hypotheses is generated using a geometric hashing technique based on Jones and Bhanu (1999). The system then finds the best data/model correspondence using a quadrant analysis technique which transforms the scattering centers from the model coordinate system to the image coordinate system. In this transformation, only translation is considered because rotation is handled by 360 models for every single degree of azimuth (note that there is no

Table 1
Average and standard deviation for articulation variance (against 360 non-articulated cases)

Articulation	M1		T72		T80	
	μ (%)	σ	μ (%)	σ	μ (%)	σ
Turret 30°	64	9.26	52	12.05	47	8.06
Turret 60°	63	9.37	51	11.75	46	8.02
Turret 90°	63	9.09	52	12.03	47	7.99
Average	63	9.24	52	11.94	47	8.02

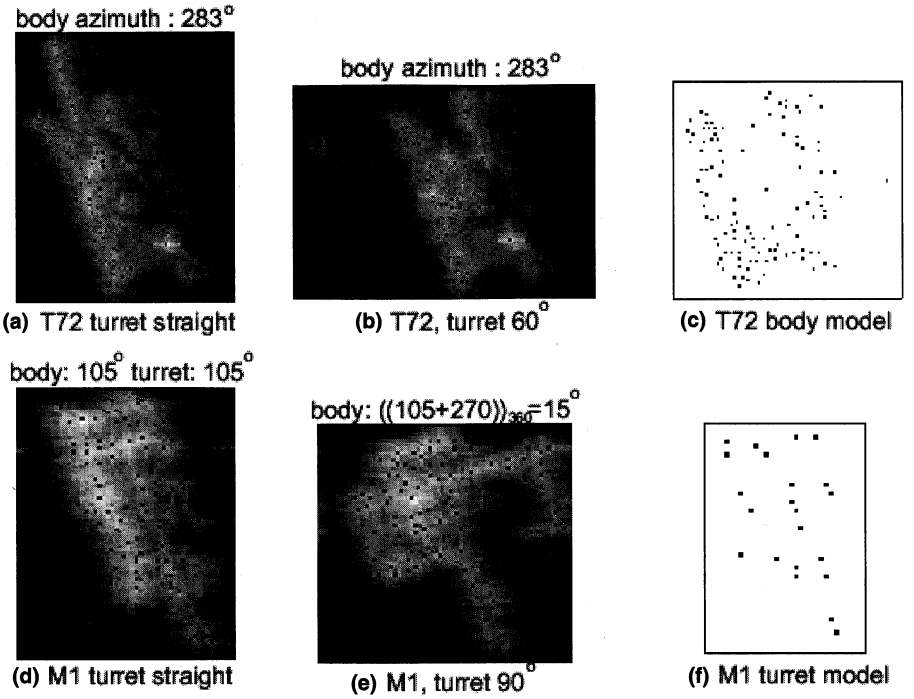

Fig. 3. T72 body model at 283° and M1 turret model at 105°: (a) body at 283°, turret straight; (b) body at 283°, turret at 60°; (c) body model at 283°; (d) body at 105°, turret pose at 105°; (e) body at 15°, turret pose at 105°; (f) turret model at 105°.

scaling involved in SAR image formation). This quadrant analysis technique allows positional error, ϵ_p, within ± 1 pixel. Finally, the system chooses the top ten hypotheses from the quadrant analysis, which have the highest matching scores, then verifies them using threshold determined dynamically based on the number of model features, the number of data features, and the area of transformation space. If matching scores are low for all hypotheses, geometric reasoning is performed.

4.1. Quadrant analysis

Each entry in the transformation space represents a transformation $\mathcal{T}$ and the value represents the number of correspondences between model and data scattering centers for exact matching positions. Since the feature extraction module uses eight-neighbor comparison, the closest two scattering centers (points) are two pixels apart from each other. In order to allow correspondence between points that are one pixel apart, the quadrant analysis routine generates a new transformation space by adding all values at the four corners of each quadrant. A new transformation space, $\mathcal{Q}$, is constructed as follows:

$$\mathcal{Q}(x,y) = \mathcal{T}(x,y) + \mathcal{T}(x+1,y) + \mathcal{T}(x,y+1) + \mathcal{T}(x+1,y+1).$$

Fig. 4 shows an example of quadrant analysis.

4.2. Dynamic selection of matching threshold

Grimson and Huttenlocher (1991) have derived an expression for a matching threshold using a statistical occupancy model. The main assumption underlying this model is that the extraneous features in an image will be uniformly randomly distributed with respect to a given object model. We use this approach with restrictions and modifications that are required for SAR images. In particular, we are concerned with scattering centers as point features in a SAR image and only the

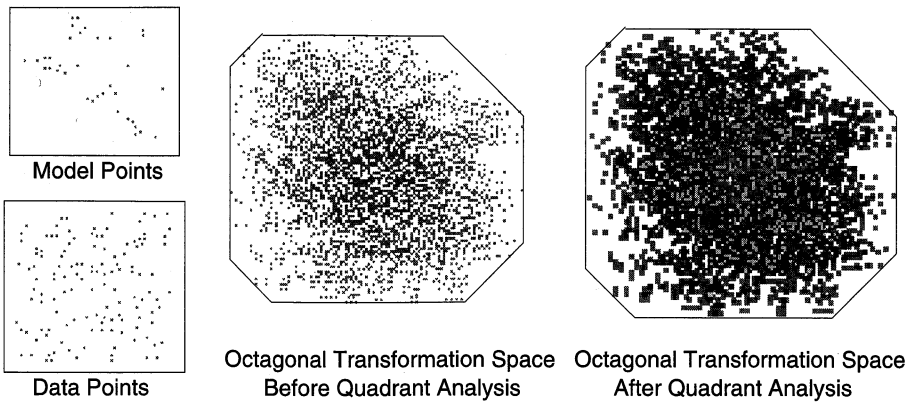

Fig. 4. An example of new transformation space with convex octagonal bounding box.

translation transformation between the model and data features.

Given a correspondence between data and model points for a fraction f ($f \epsilon [0,1]$) of the m model points, what is the relation between f and the probability δ that correspondences can occur at random? Once f and δ are known, the matching threshold can be dynamically selected for each input. In order to characterize the probability of a false match of a model to an image, statistical occupancy models are used. Let l be the number of random feature correspondences. Assuming the acceptable positional error to be within $\pm\epsilon_p$ in both x and y directions, which is a square of dimensions $2\epsilon_p \times 2\epsilon_p$, the size of the range of feasible transformations V_{jJ} for single data–model pairing (j,J) is $c_{jJ} = 4\epsilon_p^2$. Let the sum of the sizes of all the transformation space over the total size of the transformation space be λ. Then,

$$\lambda = \frac{sm4\epsilon_p^2}{A} = sm\bar{c},$$

where s and m are the number of scattering centers for data and model, respectively, A is the convex octagonal area (the bounding octagon) of the transformation space, and $\bar{c}$ is the average normalized size which is $4\epsilon_p^2/A$.

Given n cells and r events, what is the probability, p_k, that a given cell contains exactly k events? Two widely used models for the probability are *Maxwell–Boltzmann* and *Bose–Einstein* models. Maxwell–Boltzmann model assumes that the events are uniformly randomly distributed, such that all n^r possible placements of the r events in the n cells are equally probable. Bose–Einstein model, an alternative model, assumes that each *distinguishable distribution* of events across cells has an equal probability of occurrence.

To select a particular occupancy model for articulated object recognition, we generate test data with random scattering centers inside the bounding box whose size is the average size of all test SAR chips. We compute the cumulative empirical distribution as follows:

(1) For each pair of model and data points, compute a point in the transformation space. Enter an event into the cell containing this point; (2) Do the quadrant analysis to get a new transformation space; (3) Over all cells inside the convex octagonal hull of the events, count the number of events in each cell (the occupancy numbers), and tally the number of cells with each occupancy number; (4) For each entry in the tally, normalize the entry by the total number of cells, thus producing the empirical distribution of the number of events per cell; (5) Sum the normalized values to obtain the cumulative empirical distribution.

Fig. 5 shows the fitness comparison of the two occupancy models for the empirical distribution generated using the model and data shown in Fig. 4. This cumulative empirical distribution is supplied to the Kolmogorov–Smirnov test (K–S test), which measures the maximal difference between the empirical distribution $F(x)$ and some hypothesized

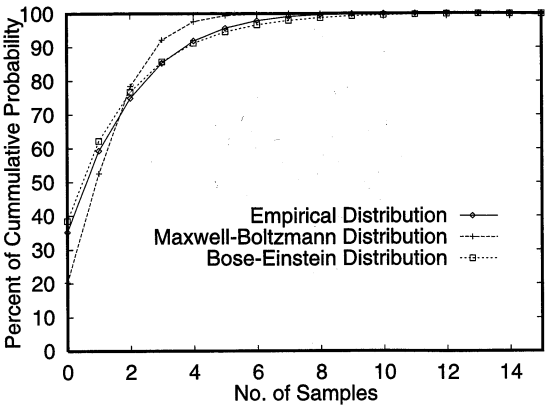

Fig. 5. Comparison for the fitness of two occupancy models.

distribution, $F_a(x)$, (e.g., Bose–Einstein model or Maxwell–Boltzmann model). All 360 models of both body and turret parts for each target are tested against the data. Table 2 shows the results of the K–S test. For $\alpha = 0.05$ level, D_n's (the maximum difference between two cumulative distribution functions) for the Maxwell–Boltzmann model are large enough to reject hypothesis H_0 ($H_0 : F(x) = F_a(x)$) while D_n's for the Bose–Einstein model are not large enough to reject H_0. Accordingly, we choose the Bose–Einstein model for the occupancy model.

From $Pr\{v \geqslant l\}$, probability that l or more of the spaces intersect at random, we can determine the fraction of model features f_0 such that the probability of mf_0 features being matched at random is less than some predefined level δ. Since δ is a function of the noise in the data measurements, and the uncertainty in position $4\epsilon_p^2$, we have

$$f_0 \geqslant \frac{\log(\frac{A}{4\epsilon_p^2})}{m\log(1+\frac{1}{msc})}.$$

In the verification stage, we consider the number of correspondences from the quadrant analysis, the number of data features, the number of model features, and the size of the transformation space. Based on this information, the verification stage calculates the matching score along with the formal threshold.

4.3. Geometric reasoning

In our approach, we assume that a target has two parts one of which is larger than the other. For example, the body part is larger than the turret part and the body part has more articulation invariant points than the turret part. But there are some exceptions like the M1 tank which has a relatively large turret compared to the size of the body. Even though the body part is still larger than the turret part, there are more articulation invariant points on the turret than on the body for some configurations. In this case, the recognition of the body part will fail because of the lack of articulation invariant points from the body. For this case we try to recognize the turret part first.

Fig. 6 shows the improvement obtained through geometric reasoning ('with-loop' results). The identification does not consider the pose of each part. For the body and turret pose recognition, the correct pose is within $\pm 5°$ accuracy. The improvement is significant for the identification and turret pose recognition cases.

5. Results

The geometric hashing technique of the recognition engine (Jones and Bhanu, 1999) efficiently accumulates the evidences of invariance, but does

Table 2
D_n for Kolmogorov–Smirnov test (total 360 models/target)

	Body		Turret	
	Bose–Einstein (%)	Maxwell–Boltzmann (%)	Bose–Einstein (%)	Maxwell–Boltzmann (%)
M1	4.13	13.20	3.31	10.99
T72	12.52	32.41	3.64	9.38
T80	4.31	15.97	4.97	5.22
Average	6.90	20.53	3.97	8.53

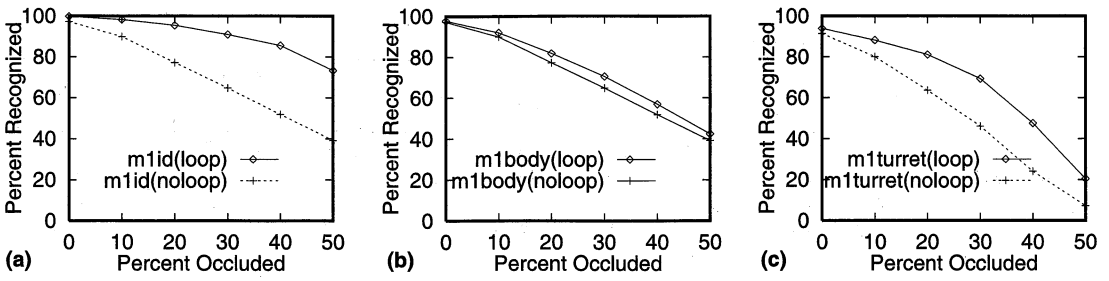

Fig. 6. Effect of geometric reasoning for M1 tank: (a) identification; (b) body pose recognition; (c) turret pose recognition.

not show the one-to-one correspondence between model and data. Our algorithm developed here, based on the 'positive–negative' feature analysis, finds the best correspondence between two given sets of points, divides the data into a set of 'positives' that match and 'negatives' that do not match, and finds the distance transform. This is needed ultimately to determine the pose of the body and turret. Given two sets of points, one from the hypothesized model and the other from the input data, we transform the first set to find the maximum number of corresponding points. In this transform, only translation is considered because the rotation and scaling are taken care of by the design of the recognition system and the characteristics of the SAR sensor. The algorithm basically uses each model and data point in turn as a trial basis for the transform and finds the transform with the maximum number of corresponding points. The data points which have corresponding points to the model points, are called *positives*, and the rest of the points, are called *negatives*.

5.1. Results using XPATCH data

In building synthetic models, we use three targets: T72 tank, M1 tank and T80 tank. Each target has four different configurations which are achieved by rotating the turret by 0°, 30°, 60°, and 90° relative to the tank body centerline. For each articulation configuration, we generate 360 SAR images (one for each degree in azimuth) of each target at a constant radar depression angle of 15°. So, the total number of images generated is 4320 (3 targets × 4 articulations × 360). From each image, we extract scattering centers from their signal returns as point features of the model. We are able to perform six sets of experiments by using the (4 choose 2) combinations of two articulation configurations of each target to build both the turret and body models, while the remaining two articulation configurations (2160 images) are used as a test data set. The results are averaged over the six experiments to obtain the final performance.

In addition to the articulated test data as described above, we also generated occluded test data (with upto 50% occlusion) to further evaluate the performance of the technique. The occluded data is simulated by starting with a given number of strongest scattering centers and then removing the appropriate number of scattering centers encountered in order, starting in one of the 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). This simulates the spatial correlation of missing features in occluded objects. Then the same number of scattering centers (with random magnitude, within the overall range of target scatterers) are added back at random locations within the original bounding box of the chip. This synthetic noise keeps the number of scatterers constant and acts as a surrogate for some potential occluding object.

Table 3 shows the object identification confusion matrix for M1, T72 and T80 tanks for various occlusion rates. The test data for each (non-zero) percent of occlusion is 4320 XPATCH images: 3 targets × 4 articulations (turret 0°, 30°, 60°, and 90° relative to the tank body) × 360 azimuths. The NIL column shows the cases where no hypothesis has a high enough matching score (determined by

Table 3
Confusion matrices for occluded articulated objects

Ocl. (%)	M1 tank				T72 tank				T80 tank			
	M1	T72	T80	Nil	M1	T72	T80	Nil	M1	T72	T80	Nil
0	99.75	0.23	0.00	0.02	0.00	100.0	0.00	0.00	0.00	0.00	100.0	0.00
10	98.31	1.64	0.02	0.02	0.00	100.0	0.00	0.00	0.00	0.00	100.0	0.00
20	95.60	4.17	0.02	0.21	0.00	99.81	0.00	0.19	0.23	0.51	99.24	0.02
30	91.46	7.57	0.09	0.88	0.12	98.38	0.02	1.48	1.76	3.03	94.93	0.28
40	86.46	11.04	0.44	2.06	1.69	92.94	0.09	5.28	5.37	5.25	88.82	0.56
50	74.75	17.96	0.62	6.67	4.75	83.52	0.14	11.60	11.11	8.24	79.12	1.53

Total test data used for each (non-zero) % of occlusion = 4320 images.

the dynamic threshold algorithm) to be accepted as valid.

In Figs. 7(a)–(f), *average*, *M*1, *T*72, *T*80 curves represent the experimental results of the average and each individual target for percent correct recognition as a function of percent occlusion. The results in Figs. 7(a)–(c) are the identification, body pose and turret pose recognition rates without the geometric reasoning step. The corresponding Figs. 7(d)–(f) show the improvements achieved with geometric reasoning (the M1 curves). The identification performance degrades gracefully as the occlusion rate increases. The recognition of T72 and T80 are similar, while the M1's recognition rate decreases faster as the occlusion rate increases. In the body and turret pose recognition results, correct pose recognition is within $\pm 5°$. The low pose recognition rate of the M1 tank body is due to the fact that the relative size of the turret to the body is large. This characteristic is reflected in the pose recognition of the turret part, where the M1 turret pose recognition is much better than the other two tanks. As the occlusion rate increases, the turret pose recognition drops rapidly beyond 30% occlusion. This is expected because the turret part is in the middle of the image, in general, and

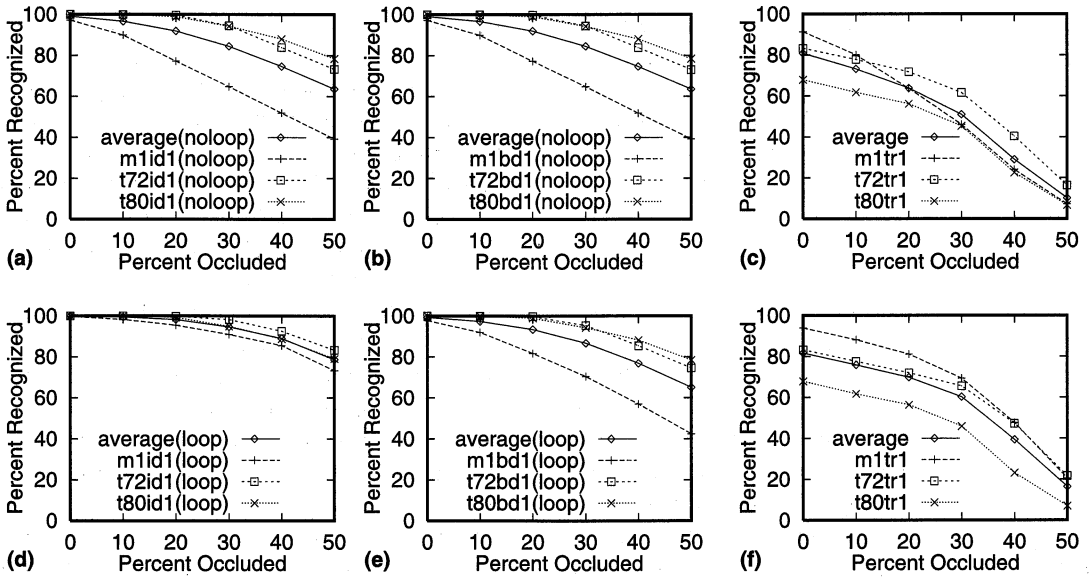

Fig. 7. Recognition results without (a, b, c) and with (d, e, f) geometric reasoning: (a) identification; (b) body pose recognition; (c) turret pose recognition; (d) identification; (e) body pose recognition; (f) turret pose recognition.

will have fewer valid point features as the occlusion rate increases. Note that in addition to recognition of the objects we obtain the poses of both parts.

5.2. Articulated MSTAR results

Real SAR images (at one foot resolution) of only two articulated objects (a T72 tank and a ZSU 23/4 anti-aircraft gun) with only two articulations (turret straight and at 315°) are publicly available in the MSTAR (Ross et al., 1998) data set. The baseline non-articulated set has 288 different azimuthal views for each target and the articulated set has 118 azimuths for each target. With this real data, we are able to correctly recognize the articulated objects (with the turret rotated) based on models with the turret straight.

Example SAR images and ROI, with scattering centers superimposed, are shown in Figs. 8 and 9 for baseline and articulated versions of the T72 and ZSU. The ROIs were found in the MSTAR target chips by reducing the speckle noise using the Crimmins algorithm in (Khoros, 1998), thresholding at the mean plus two standard deviations, dilating to fill small gaps, eroding to have one large region and little regions, discarding the small regions with a size filter and dilating to expand the extracted ROI. The scattering centers were extracted from the MSTAR magnitude data (within the boundary contour of the ROI) by finding local eight-neighbor maxima. In these experiments we obtained an average of 19 scattering centers for T72 tank and 17 for the ZSU-23/4. For the MSTAR data the matching process is divided into three steps.

In the first step, we use the scattering centers as 2D (range and cross-range) point features like before with the XPATCH generated data. As only two articulations are available for each object (T72 tank – serial number a64 and ZSU anti-aircraft gun – number d08), we could not apply the body and turret model building approach directly, because we require two data sets for the models and need a third for test. Instead, we use baseline (turret straight) sets for models and articulated (turret at 315°) sets for test data to obtain the recognition results for articulated object identification and body pose only (no turret pose).

In Tables 4 and 5, the columns with 'Point' show the results of this first step. All results are based on the top hypothesis which passed the dynamic matching threshold. The confusion matrix in Table 4 shows the correct identification rate for the T72 is 71.0% and for the ZSU 72.9% using the point features. The percent of correct body pose is given in Table 5, where the numbers inside the '( )' show the results when the body pose in the opposite direction (with a 180° difference) is counted as correct. The T72 body pose correctness using the point features are 50.8% and 54.0% for $\pm 10°$ and $\pm 20°$ accuracy, respectively. If we consider the opposite body pose recognition the correctness increases to 56.5% and 65.3%, respectively.

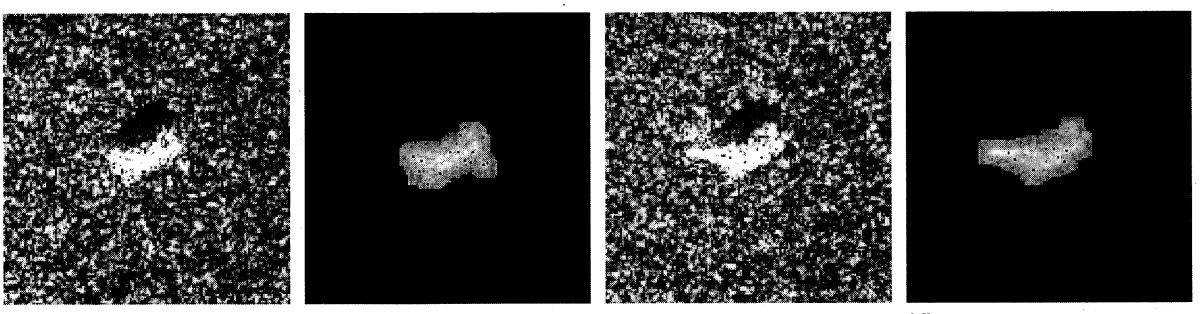

Fig. 8. MSTAR SAR images and ROIs (with peaks) for T72 tank #a64 at 56°: (a) T72: baseline image; (b) baseline ROI; (c) articulated image; (d) articulated ROI.

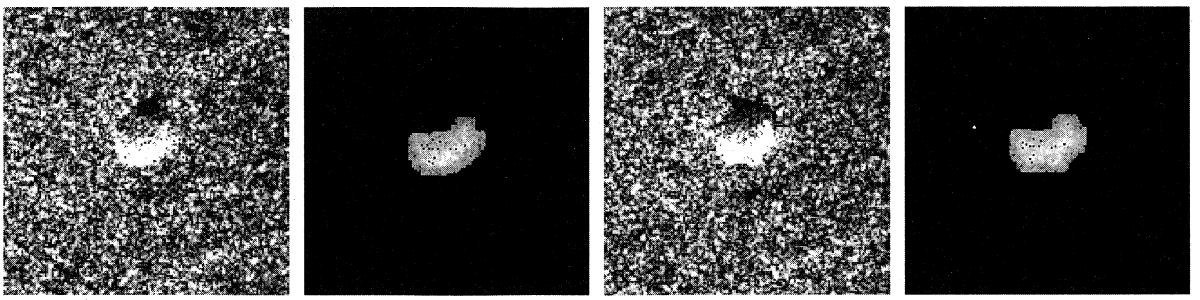

Fig. 9. MSTAR SAR images and ROIs (with peaks) for ZSU 23/4 #d08 at 66°: (a) ZSU: baseline image; (b) baseline ROI; (c) articulated image; (d) articulated ROI.

Table 4
Confusion matrix for point, area, and magnitude matching

Feature	T72			ZSU		
	Point	Area	Mags.	Point	Area	Mags.
T72	71.0	87.1	90.3	29.0	12.9	9.7
ZSU	27.1	5.1	4.2	72.9	94.9	95.8

Table 5
Percent of correct body pose: $\pm 10°$ ($180° \pm 10°$) and $\pm 20°$ ($180° \pm 20°$)

Feature	$\pm 10°$ ($180° \pm 10°$)			$\pm 20°$ ($180° \pm 20°$)		
	Point	Area	Mags.	Point	Area	Mags.
T72	50.8 (56.5)	82.3 (83.1)	85.5 (86.3)	54.0 (65.3)	85.5 (87.1)	88.7 (90.3)
ZSU	27.1 (38.1)	65.3 (80.5)	65.3 (81.4)	29.7 (45.8)	70.3 (90.6)	70.3 (91.5)

For the second step, once we find the best data-to-model transformation for each hypothesis at the first step, we apply this transformation to the data ROI, and find the overlapping area, A_o. We then calculate the ratio $R = (2 \times A_o)/(A_d + A_m)$, where A_d and A_m are area of ROIs for the data and model, respectively. Columns with 'Area' in Tables 4 and 5 show the results of the second step. This second step improves the overall identification rate from 72.0% to 91.0%. After the second step the body pose correctness (within $\pm 10°$ and $\pm 20°$) improves to 82.3% and 85.5% for T72, 65.3% and 70.3% for the ZSU. If we consider the opposite body pose recognition ($180 \pm 10°$), the correctness increases much more (to 80.5% and 90.6%) for the ZSU case, which means that for the ZSU there is more confusion with the body pose in the opposite direction.

The third step considers the magnitudes of the corresponding k scattering centers from the first step. If, in the second step, the difference of R between the first and second top hypotheses is less than a (2.5%) threshold, we compute $S = (1/n)(\sum_{k=1}^{n} | S_{d,k} - S_{m,k} |)$, where $S_{d,k}$ is the magnitude of the kth scattering center from the data image and $S_{m,k}$ is the model counterpart, and select the hypothesis with the smaller value of S as the answer. Columns with 'Mags' show the results of this third step.

Note that the indexing module (detailed in (Jones and Bhanu, 1999)), is the critical time element in the end-to-end system. (There the models are implemented as look-up tables and the results are generated using an efficient search for positive evidence in the test images.) This paper is concerned with the verification of the few hypo-

theses that are generated by the upstream indexing module. The processing time is less than one second on a SUN Ultra 1 workstation.

6. Conclusions

We have developed an end-to-end system for recognition of articulated objects and we have demonstrated the performance of our approach with extensive experiments, both under occluded conditions using XPATCH data and with real SAR images from the MSTAR data. We have explicitly modeled and demonstrated the ability to recognize the pose of both parts of articulated objects. While the current work has been primarily directed toward simple point features, in the future we plan to use complex features obtained from real SAR data.

Acknowledgements

This work was supported in part by DARPA/AFOSR grant F49620-97-1-0184. The contents of the information do not necessarily reflect the position or the policy of the U.S. Government.

References

Andersh, D.J., Lee, S.W., Ling, H., Yu, C.L., 1994. Xpatch: A high frequency electromagnetic scattering prediction code using shooting and bouncing ray. In: Proc. Ground Target Modeling and Validation Conf., pp. 498–507.

Beinglass, A., Wolfson, H.J., 1991. Articulated object recognition, or: How to generalize the generalized Hough transform. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 461–466.

Bhanu, B., Jones, G., 2000. Recognizing MSTAR target variations and articulations. Opt. Eng. 39 (3), 712–723.

Grimson, W.E.L., Huttenlocher, D.P., 1991. On the verification of hypothesized matches in model-based recognition. IEEE Trans. Pattern Anal. Machine Intell. 61 (3), 1201–1213.

Hel-Or, Y., Werman, M., 1994. Recognition and localization of articulated objects. In: Proc. IEEE Motion of Non-Rigid and Articulated Objects, 1994.

Jones, G., Bhanu, B., 1999. Recognition of articulated and occluded objects. IEEE Trans. Pattern Anal. Machine Intell. 21 (7), 603–613.

Jones, G., Bhanu, B., 2001. Recognizing articulated targets in SAR images. Pattern Recognition 34 (2), 469–485.

Khoros, 1998. Pro v2.2 User's Guide. Addison Wesley, Longman, New York.

Novak, L.M., Owirka, G.J., Netishen, C.M., 1994. Radar target identification using spatial matched filters. Pattern Recognition 27 (4), 607–617.

Ross, T., Worell, S., Velten, V., Mossing, J., Bryant, M., 1998. Standard SAR ATR experiments using the MSTAR public release data set. In: Proc. SPIE Algorithms for Synthetic Aperture Radar Imagery V, Vol. 3370, pp. 566–573.