Matching of Articulated Objects in SAR Images

Joon Soo Ahn and Bir Bhanu*
College of Engineering, University of California, Riverside, CA 92521-0425
E-Mail: {ahn, bhanu}@constitution.ucr.edu
URL: http://constitution.ucr.edu

Abstract

We present distance weighted geometric hashing-based indexing technique and a matching technique using body models and turret models for the automatic target recognition of articulated objects in synthetic aperture radar images. For each target, 360 body models and 360 turret models are built. These models are independent of the relative position between the body and turret. Four non-articulated targets (SCUD missile launcher, T-72 tank, M1 tank and T-80 tank) are used in the indexing stage to build the look-up table. In the matching stage, M1 tanks with turret rotated 30°, 60°, 90° relative to the body are used as data.

1 Introduction

Recognition of articulated objects in SAR images is a challenging problem. A simple approach may consider each of the articulated parts of an object as separate objects. However, such an approach is quite inefficient since it will require a large model database. We want to develop an efficient recognition approach that inherently models the articulated nature of an object such as a SCUD missile launcher or a tank with different positions of its turret.

Some of the representative work for target recognition using SAR images includes [2], [4] and [5]. These papers focus on template matching techniques in which the templates are manually designed. Recent work on the recognition of articulated objects in SAR images includes [1], [3]. In this paper we describe a geometric hashing-based indexing with weighted voting and a matching technique using body models and turret models. We have evaluated the performance of our initial approach using XPATCH data.

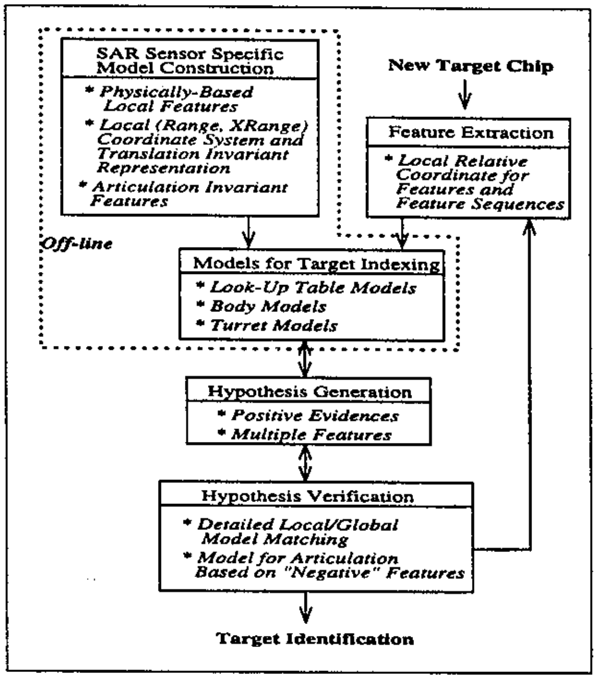

Figure 1: System for recognizing articulated targets in SAR images.

1.1 Approach

Figure 1 shows the system for recognizing articulated targets in SAR images. Our approach is based on local features and local reference coordinate system. The models for the look-up table are constructed by extracting the relative positions of the features from the non-articulated training data. The body and turret models are constructed by using three different articulation configurations of tank targets.

Detailed description of the geometric hashing technique, specifically designed for SAR, using the look-up table is given in [3]. Distance weighted geometric hashing-based indexing is an enhancement of the basic geometric hashing technique which increments each vote not by one, but by $max(|dx|, |dy|)$ where $|dx|$ and

*This work is supported by grant MDA972-93-1-0010. The contents and information do not reflect the position or policy of the U.S. Government.

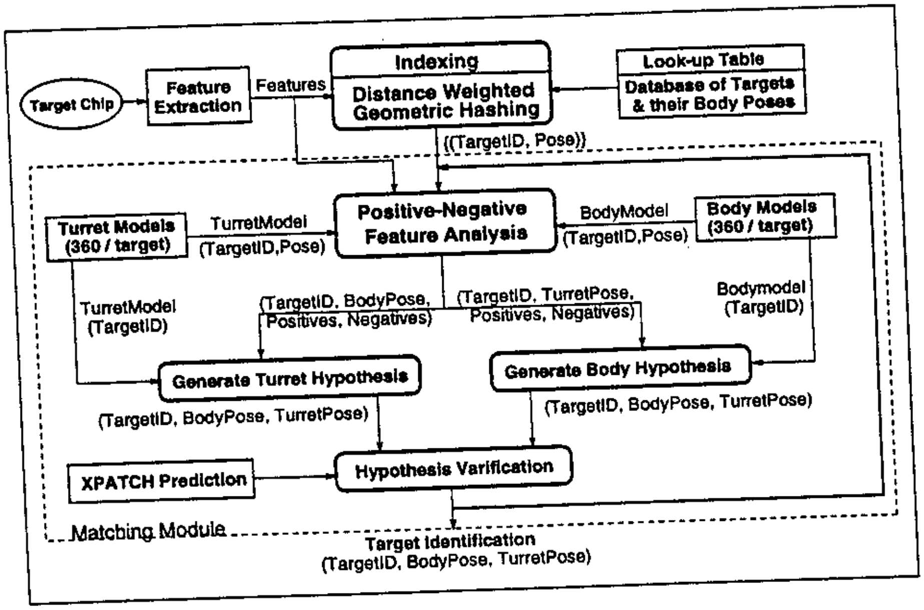

Figure 2: Indexing and matching components for recognizing articulated targets.

$|dy|$ represents the absolute values of the relative distances between two peak features in the direction of range and cross-range, respectively. This module generates set of hypotheses with target identification and pose which may be the body pose or the turret pose.

For the matching process, 360 body models and 360 turret models are built (one/degree azimuth) for each tank target and used to find the positive and negative point features from the data (see Figure 2). The positive features are used to generate hypotheses for target identification and body pose. Negative features are used to generate hypotheses for turret pose. As there might be multiple hypotheses from the indexing module, the matching module will loop for each of the hypotheses.

The basic assumption is that the positive points are from the body part. The negative features are produced as a result of articulation and interaction between the body and the turret. For some targets like the M1 tank, the turret part is so large that the indexed pose may be the turret pose. To resolve this problem, the **Positive-Negative Feature Analysis** stage uses body models and turret models to detect the part. This stage uses the body model with specific target type (ID) and body pose to detect the positive features. If the number of positive points are larger than a fraction of the number of the specific body model points, then we generate the turret hypothesis. If the number of positive points are less than a fraction of the number of the body model points, we use the turret model for the target ID and turret pose and go to the next step to generate the body hypothesis.

2 Off-line Model Building

2.1 Extraction of scattering centers to build non-articulated model base

We employed a simple method of detecting local maxima. The method is based on comparing the pixel value with its immediate eight neighbors. If the current pixel value is greater than all the other immediate eight neighbors, then it is a local maximum.

2.2 Building non-articulated model base

We extracted the top fifty local maxima from the images of SCUD missile launcher with missile down, T72 tank, M1 tank and T80 tank with turrets straight to the bodies. The top fifty local maxima are then sorted in descending order of their magnitudes of SAR return signals.

```
build_non-articulated_model_base()
{  N = number of local maxima
   for (Object = 1 to NO_OF_OBJECTS){
       for (Angle = 0 to 359){
           model = get_model_image(Object, Angle)
           peaks = extract_local_maxima(model, N)
           save(Object, Angle, peaks)
}}}
```

2.3 Building body models

Figure 3 shows T-72 tanks with turret 0°, 60°, and 90° rotated relative to the body whose pose is 283°. The fourth figure shows the body model of T-72 tank

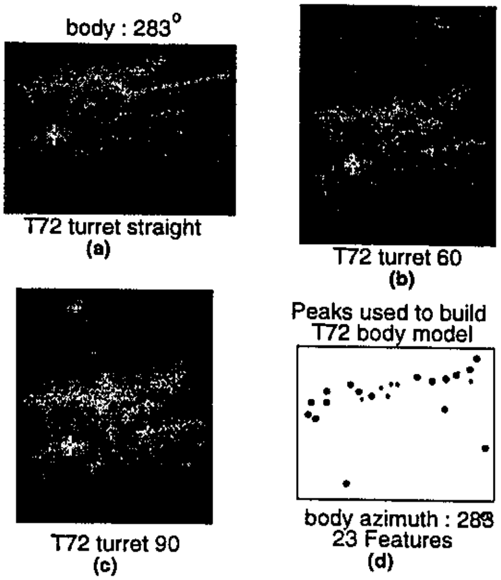

Figure 3: T-72 body model with azimuth 283°. (a) body azimuth = 283°, turret straight. (b) body azimuth = 283°, turret 60°. (c) body azimuth = 283°, turret 90°. (d) body model at 283°.

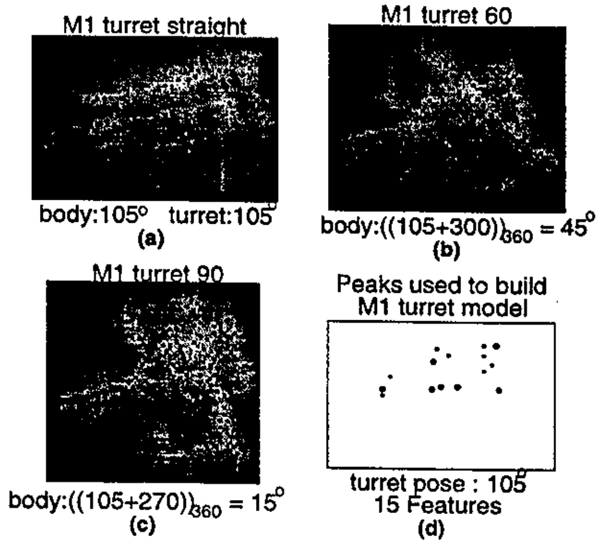

Figure 4: M-1 turret model with azimuth 105°. (a) turret pose = 105°, body azimuth = 105°. (b) turret pose = 105°, body azimuth = 55°. (c) turret pose = 105°, body azimuth = 15°. (d) turret model at 105°.

with the body azimuth at 283°. To build this model, first, conjunction operation between two images with different turret poses has been performed. This operation generates three sets of point features (0° & 60°, 0° & 90°, 60° & 90°). Then, union operation on these three sets of point features give a set of point features which represent the body model of T-72 at 283°. The conjunction operation represents the best matching between two sets of point features. The union operation represents the union of two sets of point features where one set of point features are translated appropriately to have the best matching between them. The result in 3(d) shows 23 point features. There are 17 large dots in the model, which represents the point features for the best matching among the three sets of original point features.

2.4 Building turret models

Figure 4 shows M-1 tanks with turret 0°, 60°, and 90° rotated relative to the body whose poses are 105°, 45°, and 15° respectively. The fourth figure shows the turret model of M-1 tank with the turret azimuth at 105°. To build this model, first, conjunction operation between two figures with different turret pose has been performed. This operation generates three sets of point features. Then, union operation on these three sets of point features give a set of point features which represent the turret model of M-1 at 105°. The result in Figure 4(d) shows 15 point features. There are 7 large dots in the model, which represents the point features for the best matching among the three sets of original point features.

3 On-line Indexing and Matching

3.1 Distance Weighted Geometric Hashing-Based Indexing

The paper by Jones & Bhanu [3] describes the geometric hashing technique in detail. We have enhanced the indexing module by incrementing the variable *vote* by $max(|dx|, |dy|)$ instead of incrementing it by one. This new weighted voting scheme is different from the original non-weighted voting scheme in employing the relative distance between two points as the weighting factor. This approach improves the indexing results as shown in Figure 5.

3.2 Matching

Following algorithm generates set of hypotheses and finds the best correspondence between data and the set of hypotheses.

```
Exact_Matching(data_image)
{  data = extract_local_maxima(data_image, N)
   candidates = Weighted_Geometric_Hashing(data)
   for (each model in top K candidates){
      Positive_Negative(data_points, model_points)
   }
   Sort the hypotheses in descending order of
      the positive points
   Select model with the most matching points
}
```

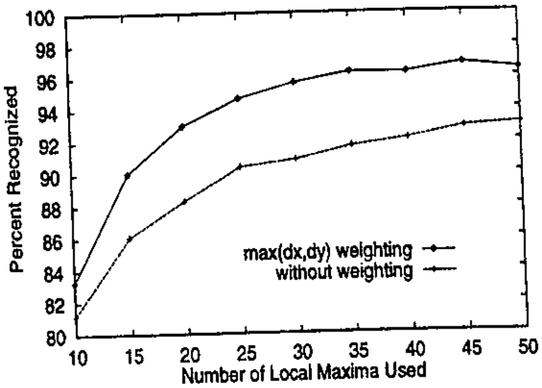

Figure 5: Comparison of the two indexing schemes $max(|dx|, |dy|)$ weighting scheme and no-weighting scheme for the top hypotheses.

Given two sets of points, 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 peculiar characteristics of the SAR sensor.

Algorithm given in Figure 6 shows how to find the positive and negative points. The best case time complexity is $O(MD)$ and the worst case time complexity is $O(MD^2)$ where M and D represents the number of model and data point features respectively.

4 Experiment Results

4.1 Building Models

In model building, we use four non-articulated objects, SCUD missile launcher with missile down, T72 tank, M1 tank and T80 tank with the turret straight to the body. For each non-articulated object, we generate 360 images (for each degree in azimuth) for a given depression angle of 15°. From each image, we extracted the top 50 scattering centers with their signal returns and locations as point features of the model. So, the total number of models in the model database is 1440 (4 non-articulated objects * 360).

4.2 Generating testing data

For the testing data, we used M-1 tanks with three articulated turret positions, 30°, 60° and 90° rotated relative to the tank body. For each articulated position, we generated 360 images (one for each degree in azimuth) for a given depression angle of 15°. From each image, we extracted the top 50 scattering centers with their signal returns as point features of the data. So, the total number of data in the experiment is 1080 (3 articulated objects * 360).

```
GIVEN : • 2D model points and Model Coordinate System(MCS)
        • 2D data points and Data Coordinate System (DCS)
FIND :  • maxcount : the maximum number of correspondences
          between model and data points.
        • correspond : the maximum set whose elements are pairs
          of locations for the correspondence between model and data.
        • positives : the data points in the correspond
        • negatives : the data points which are not in the correspond

procedure Positive_Negative (model_points, data_points)
  place model points on 2D array,A, using MCS
  initialize maxcount to 0, correspond to empty set
  for (each model point M)
    for (each data point D)
      initialize count to 0, corr to empty set
      compute the offset between M and D
      for (each data point D1)
        apply the offset to D1 (convert it from DCS to MCS)
        if (there is a model point M1 at the same location in A)
          increment count and append (D1,M1) to the corr
      if (count > maxcount)
        update maxcount with count and
        update correspond with corr
    if (maxcount >= the number of remaining M) return
```

Figure 6: Algorithm for the Positive-Negative Feature Analysis.

4.3 Example of positives from turret

Figure 7 shows an example of the positive points in the data compared to the non-articulated model which is recognized by the turret pose instead of the body azimuth. Note that the positive feature (Figure 7(c)) matches better with the turret model (Figure 7(e)) then with the body model (Figure 7(d)).

4.4 Discussion

Figure 8 shows the results for indexing. The *enhanced indexing with weight* $max(|dx|, |dy|)$ curve shows the target ID and body pose up to the 40th position in the list of hypotheses sorted in descending order of the vote. The cumulative percentage accuracy up to 40th hypotheses is 92.87%.

The *enhanced indexing with positive-negative* curve shows the cumulative percentage of the correct target ID and body pose up to the 40th position in the list of new hypotheses sorted in descending order of the number of positive features from the **positive-negative feature analysis** without the correction of the confusion between body and turret. This curve shows that the **positive-negative feature analysis** brings the correct answer to the top of the hypotheses list if the correct answer is among the top 40 hypotheses of the indexing result.

The *body detection by positive-negative feature analysis* curve shows the cumulative percentage of the correct target ID and body pose after the correction of the confusion between body and turret using the body models and turret models. Based on the top 40 an-

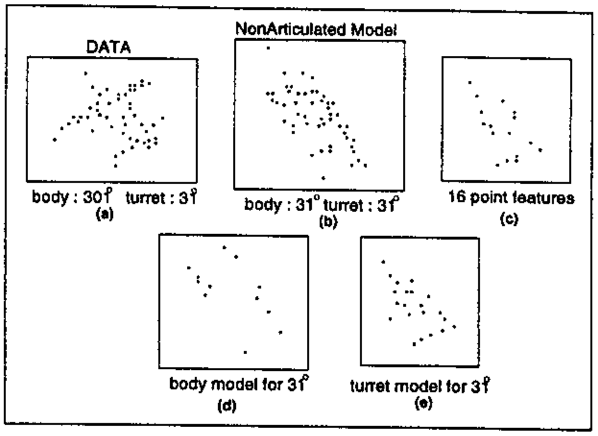

Figure 7: Example of positives from the turret of M-1 tank. (a) Data: body azimuth 301° and turret pose 31°. (b) Non-articulated model hypothesis generated by indexing: body azimuth 31° and turret pose 31°. (c) The positive features of the data detected by the non-articulated model. (d) M-1 body model: azimuth 31°. (e) M-1 turret model: pose 31°.

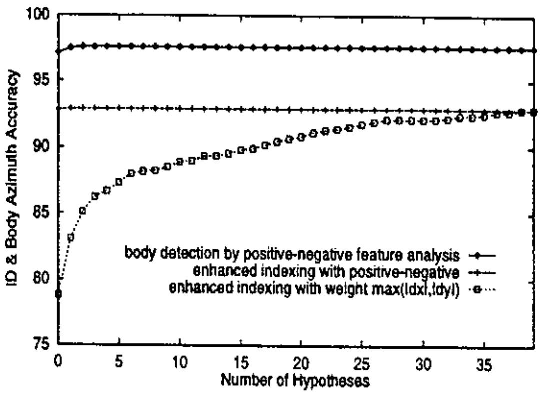

Figure 8: Results for Indexing

swer, the correct target ID and body pose increased to 97.59%, which increases the correct target ID and body pose by 4.74%.

Figure 9 shows the results based on the top hypothesis. In the column of **ID, Body, Turret**, 0, 1 and X represents incorrect, correct and don't care. The column **Exact** shows the result that testing data is recognized at exact pose. The column **Within 5** shows the result that testing data is recognized within +/− 5°.

5 Conclusions and Future Work

In this paper, we have presented the initial research for matching. The goal is to develop physically-based approaches having multiple representations (variety of feature types) for matching to recognize articulated

	ID	Body	Turret	Exact	Within ±5°
	0	X	X	0.56(%)	0.56(%)
	1	0	0	1.94(%)	1.39(%)
	1	0	1	0.37(%)	0.19(%)
	1	1	0	46.20(%)	21.02(%)
	1	1	1	50.93(%)	72.69(%)
Exact	99.44(%)	97.13(%)	51.30(%)		
Within ±5°	99.44(%)	98.98(%)	72.87(%)		

Figure 9: Recognition results. These results are based on the top hypothesis only.

targets in SAR images.

We are developing integrated matching technique and analysis for verifying hypothesis (target ID, body pose, turret pose) using articulation variants, positive/negative features, XPATCH prediction, surface reflector type and relative geometry of parts of articulated objects (e.g. M-1 / T-72). We are investigating a Bayesian probabilistic approach to combine the above known information in an integrated manner.

References

[1] B. Bhanu, G. Jones, J. Ahn, M. Li, and J. Yi. Recognition of Articulated Objects in SAR Images. In *Proc. ARPA Image Understanding Workshop*, Palm Springs, California, February 13-16 1996.

[2] D. E. Dudgeon, R. J. Lacoss, C. H. Lazott, and J. G. Verly. Use of persistent scatterers for model-based recognition. In *Proceedings of SPIE Conference on Algorithms for Synthetic Aperture Radar Imagery*, volume 2230, pages 356–368, Orlando, FL, April 1994.

[3] G. Jones and B. Bhanu. Invariant features for the recognition of articulated and occluded objects in SAR images. In *Proc. ARPA Image Understanding Workshop*, New Orleans, LA, 1997. May 13-15.

[4] L. M. Novak, G. J. Owirka, and C. M. Netishen. Performance of a high-resolution polarimetric SAR automatic target recognition system. *The Lincoln Laboratory Journal*, 6(1):11–24, 1993.

[5] A. M. Waxman, M. Seibert, A. M. Bernardon, and D. A. Fay. Neural systems for automatic target learning and recognition. *The Lincoln Laboratory Journal*, 6(1):77–116, 1993.