Automatic Model Construction for Object Recognition Using ISAR Images

Shujun Zhang and **Bir Bhanu**
College of Engineering
University of California
Riverside, CA 92521
{shujun,bhanu}@vislab.ucr.edu

Abstract

This paper discusses model construction for object recognition using ISAR images. A learning-from-examples approach is used to construct recognition models of the objects from their ISAR data. Given a set of ISAR data of an object of interest, structural features are extracted from the images. Statistical analysis and geometrical reasoning are then used to analyze the features to find spatial and statistical invariance so that a structural model of the object suitable for object recognition can be constructed. Results of experiments using the automatically constructed models in object recognition are presented.

1. Introduction

Object recognition from inverse synthetic aperture radar (ISAR) or synthetic aperture radar (SAR) imagery is an important aspect of current computer vision research [1], [2], [3], and [5]. The recent papers use template matching technique in which the template are manually designed for object recognition. As pointed out by Verly *et al.* [5], the designing of the templates is an "art instead of science." This approach is suitable only for recognizing a small number of classes of objects. A systematic model construction technique is highly desirable to develop an object recognition system capable of recognizing a large number of classes of object automatically. There is no published work on automatic model construction from ISAR/SAR images.

This paper presents our current research on automatic model construction for object recognition using ISAR/SAR images. A set of training images of an object, including sensor parameters with which the images are obtained, is used as the sole source of information about the object. The model construction scheme adapts a *learning-from-examples* technique - it learns information about an object by analyzing its image features extracted from a set of training data of the object and relations among the features. Based on the information summarized in the learning phase, the system represents this information in an appropriate data structure and generates a model of the object suitable for its recognition. This model is then used in a hierarchical search process to match image features from ISAR/SAR images for object recognition. The object and pose hypotheses are then subjected to further verification by a template matching algorithm to achieve the final object recognition.

2. The Data

Millimeter-wave high range resolution data is used in our research. The data was collected with the following parameters: nominal operating frequency: 35 GHz; stepped frequency waveform is used to achieve 1 ft. range resolution; depression angle: 5.5^o; azimuth increment about 1.04^o (quantized to 0.1^o); and phase Jump Every 2^o.

Currently 4 objects, Camaro, Dodge Van, Dodge Pickup and Bulldozer as shown in Fig. 1, are used in the system. For each of the objects, there are 351 images in both the training data and the testing data, respectively. The testing data are 0.2^o offset in azimuth angle from the corresponding training data. The depression angles are the same for both training and testing data.

To compress the dynamic range, a logarithmic filter is applied to original ISAR images to get log-magnitude images. Fig. 2 shows parts of the 351 log-magnitude polarimetric whiten filtered (PWF) ISAR images (see Verbout *et al.* [4] for details) of a Camaro, Dodge van, Dodge pickup and bulldozer. The size of each of the images is 28x18 pixels.

Although all the experiments reported in this paper use PWF ISAR images, the approach is equally applicable to single channel ISAR images, as well as to single channel

*This work was supported by ARPA grant MDA972-93-1-0010. The contents of the paper do not necessarily reflect the position or the policy of the U.S. Goverment.

1015-4651/96 $5.00 © 1996 IEEE
Proceedings of ICPR '96

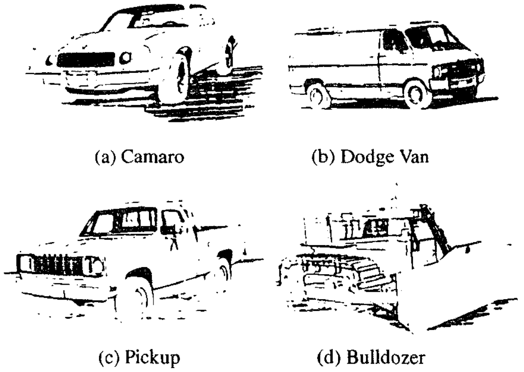

Figure 1. Vehicles used in the research.

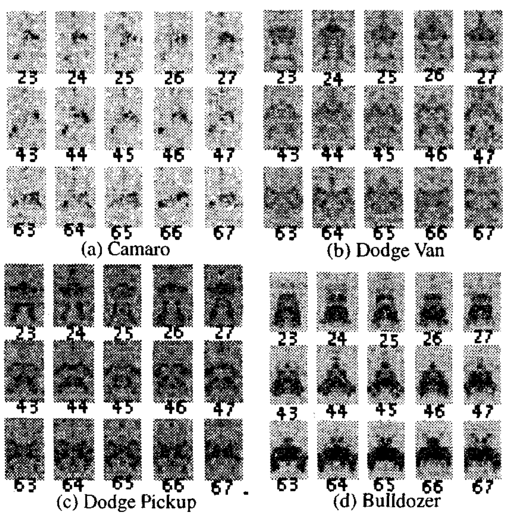

Figure 2. A set of PWF ISAR images of the objects used as training data for model construction in the system. The number under each frame of the images is the azimuth angle of the image.

SAR or PWF SAR images.

3. Model Construction

Fig. 3 summarizes this model construction scheme.

a. Model Features - Image features used are image scatterers (local peaks in the gray scale image) and their geometrical distribution. These scatterers are further grouped into more complex features by their magnitude, geometrical positions and geometrical relations with other scatterers via statistical analysis and geometrical reasoning. The following features are used: (1) **Global orientation** of all the scatterers in the signature of the object indicating the potential pose of the object; (2) **Vectors of persistent scatterers** characterizing the persistency of scatterers of the object over the changes in azimuth angles; (3) **Euclidean distance vector of persistent scatterers** characterizing the Euclidean distance from persistent scatterers to their centroids; and (4) **Local orientation vector of persistent scatterers** representing the angle of persistent scatterers with the image coordinates.

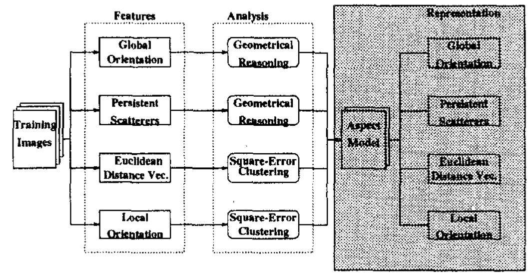

Figure 3. The Model Construction Scheme.

Global Orientation - The global orientation of scatterers in a signature of an object is used to provide useful information in predicting the pose of the potential object.

Persistent Scatterer Vectors - It has been found that although ISAR/SAR images are generally very sensitive to the changes of orientation of the object relative to the sensor and sensor parameters, some scatterers persist over a certain degree of azimuth at certain depression angles. Persisting is used to partition signatures of the object into a number of aspects so that signatures in the same aspect are similar. This reduces the storage needed for the model, as well as the search space in the matching phase. For each frame of ISAR images, a persistent scatterer vector

$$V = \{p_1, p_2, \ldots, p_m\}^T$$

is used as a feature in our model construction scheme.

Given two consecutive frames of images f_k and f_{k+1}, let s_{ij}^k and s_{ij}^{k+1} be the salient scatterer corresponding to $grid(i,j)$ in frame f_k and f_{k+1} respectively, and $p(s_{ij}^k)$ be the position of s_{ij}^k. A salient scatterer is defined as the one that persists over 1^o if

$$dis(p(s_{ij}^k), p(s_{ij}^{k+1})) < \epsilon$$

and

$$\sum_{s \in S} |dis(p(s_{ij}^k), p(s^k)) - dis(p(s_{ij}^{k+1}), p(s^{k+1}))| < \delta$$

where S is a set of neighboring salient scatters, $dis(p(s_{ij}^k), p(s_{ij}^{k+1}))$ is the Euclidean distance from point

$p(s_{ij}^k)$ to point $p(s_{ij}^{k+1})$, ϵ and δ are pre-defined threshold parameters.

Table 1 shows the average numbers of persistent scatterers over 3^o and 5^o of azimuth for 4 objects. On average, roughly 10 scatterers are found to be persistent over 1^o for each object, less than 2 scatters are found to be persistent over 10^o for all the objects.

Table 1. Average numbers of persistent scatterers of the training images used in the research.

Object	Images	Over 3^o	Over 5^o
Camaro	351	5.8	2.1
Dodge Van	351	6.2	1.8
Dodge Pickup	351	5.2	1.3
Bulldozer	351	5.7	2.9

Persistent scatterers are extracted by using multiple frames of images during this model construction stage. They are used to best characterize the invariance of the ISAR images of an object. However, this feature itself cannot be used in the object recognition stage, when only a single frame of image is available. A model-based persistent scatterer extraction approach will be introduced in Section 4 for robust model matching during recognition phase.

Euclidean Distance Vector of Persistent Scatterers - Geometrical relations between pairs of and among triples of scatterers provide powerful and reliable constraints for the matching process. Given m persistent scatterers,

$$P = \{p_1, p_2, \ldots, p_m\}^T$$

in which $p_i =< x_i, y_i >$, let p_c be the centroid of the persistent scatterers (defined as the mean values of $p_1, p_2, \ldots, p_m$). The distances from each of the persistent scatterers to the centroid p_c forms a *Euclidean distance vector of persistent scatterers*,

$$V_d = \{d_1, d_2, \ldots, d_m\}^T$$

in which

$$d_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}$$

is the Euclidean distance between p_i and p_c.

Euclidean Distance Vector of Persistent Scatterers -

Orientation of persistent scatterers - It plays an important role in characterizing the object. Local orientation vector of persistent scatterers of an image is defined.

$$V_\alpha = \{\theta_1, \theta_2, \ldots, \theta_m\}^T$$

where θ_i is the angle between $\overline{p_c p_i}$ and the *x-axis*.

b. Extracting Invariance - Statistical analysis and geometrical reasoning are applied to all of the features extracted from the training data to find geometrical and statistical invariance for model construction.

Global Orientation - It has been found that with the change of azimuth of images, the global orientation of the signature changes gradually. 351 images for each of the four objects are grouped into 36 aspects based on their global orientations. This number of aspects is selected based on our experiments. It is a trade-off among accuracy of the model, time required for object recognition and storage requirements. For example, for a bulldozer,

$$\mathcal{A}_1^O = \{1^o, 2^o, 3^o, 4^o, 180^o, 181^o, 182^o, 183^o, 184^o\}$$

indicates that images of a bulldozer with the azimuth angles $1^o, 2^o, 3^o, 4^o, 180^o, 181^o, 182^o, 183^o, 184^o$ have the same global orientation and, therefore, grouped into the same aspect.

Persistent Scatterers Vector - Persistent scatterer vector of 351 training images

$$V_P = \{V_P^1, V_P^2, \ldots, V_P^{351}\}$$

of each object are grouped into 36 aspects based on their similarities. Two vectors are deemed similar if

$$d(V_P^s, V_P^t) = \sum_{i=1}^{N} dis(V_P^s(i), V_P^t(i)) < \epsilon_P$$

where N is the number of persistent scatterers in the vector, ϵ_P is a pre-set threshold. The mean vector of each cluster of aspects is used to characterize the cluster of the aspect in the matching process.

Euclidean Distance Vector of Persistent Scatterers - A square-error clustering approach is used to group training images into aspects based on the Euclidean distance vector of persistent scatterers. Training images of an object are grouped into 36 aspects

$$\{\mathcal{A}_E^1, \mathcal{A}_E^2, \ldots, \mathcal{A}_E^{36}\}$$

based on the Euclidean distance vector of persistent scatterers.

Local Orientation Vector of Persistent Scatterers - Similar square-error clustering technique is used to group training images into 36 clusters of aspects based on the local orientation vector of persistent scatterers.

c. Model Representation - The model of an object is represented as 4 sets of clusters of aspects

$$\mathcal{A}_O, \mathcal{A}_P, \mathcal{A}_E \text{ and } \mathcal{A}_L$$

generated by the clustering process discussed earlier, where $\mathcal{A}_O$ is the clusters of aspects generated by global orientation, $\mathcal{A}_P$ by persistent scatterer vector, $\mathcal{A}_E$ by Euclidean distance vector of persistent scatterers, and $\mathcal{A}_L$ by local orientation vector of persistent scatterers respectively. Each of the aspects is represented by the name of the component images and their centroid attributes.

4. Model Validation

It is necessary to validate the model generated in Section 3 before it can be fully used in constructing a complete object recognition system. In our work, this model validation is achieved by applying the model generated from the above statistical and geometrical analysis to a number of sets of testing data.

Matching the Global Orientation - To recognize an object from a testing image, scatterers are first extracted from the original image, using dynamic thresholding as in Sections 3 as well as the global orientation of the signature in the image. This global orientation is then matched against the cluster of global orientations of all aspects in the model. For testing image I, if its global orientation matches with cluster global orientation of aspect $\mathcal{A}_O^j$, poses in the aspect $s_{i_1}^O, s_{i_2}^O, ..., s_{i_{n_i}}^O$ are deemed as the possible poses of the object. All the following matching processes are based on these potential poses.

Extraction of Persistent Scatterer Vector - Persistent scatterers are used as features because we believe that they best characterize the objects with the change in azimuth angles in ISAR images. However, this persistency cannot be detected from a single image. In the model construction phase, persistent scatterer vector of an image is extracted from 2 images, the current frame and the previous frame. In the object recognition phase, this contextual information is no longer available to help the extraction of persistent scatterers. A model-based approach is used to extract the persistent scatterers from a single image.

Instead of using contextual information (scatterers in multiple frames of images), information in the models of the object is used in finding persistent scatterer vectors. For the given image I and its potential poses

$$\{s_{i_1}^O, s_{i_2}^O, \ldots, s_{i_{n_i}}^O\}$$

generated by the global orientation matching, let $\mathcal{A}_P^{i_1}, \mathcal{A}_P^{i_2}, \ldots, \mathcal{A}_P^{i_j}$ be the aspects which contains the potential poses. The mean vector of persistent scatterers of each aspect is used in extracting the persistent scatterers vector from the current image to match the corresponding aspect. As discussed in the previous section, for each grid, pixels with the k highest magnitude are selected as the candidates for the persistent scatterers. Pixels which are the closest to the corresponding points in the mean vector of a candidate aspect are selected as the persistent scatterers. It should be noted that for different aspects, the persistent scatterer vector of the image could be different.

Matching Scores for Euclidean Distance Vector - Persistent scatterer vectors obtained for the testing image are evaluated by the Euclidean distance and local orientation measures. The Euclidean distance vector of the image is matched with the characterization vectors of every possible aspect generated by the previous search. Given 2 Euclidean distance vectors of persistent scatterers (one generated from the testing image, the other the mean vector of one aspect) V_d^1 and V_d^2, a matching score between the 2 vectors is defined as

$$\mathcal{M}_d(V_d^1, V_d^2) = \frac{V_d^1 \bullet V_d^2}{||V_d^1||||V_d^2||}$$

Currently, if the matching score is greater than 0.75 it is deemed a good match.

Matching Scores for Local Orientation Vector - Similarly, the local orientation vector of persistent scatterers of the testing image are matched with the characterization vector of each of the candidate aspects. The matching score between two local orientation vectors of persistent scatterers V_α^1 and V_α^2 is defined as

$$\mathcal{M}_\alpha(V_\alpha^1, V_\alpha^2) = \frac{V_\alpha^1 \bullet V_\alpha^2}{||V_\alpha^1||||V_\alpha^2||}$$

Voting - The confidence of a hypothesis h is voted by the 2 matching scores generated by matching Euclidean distance vector and local orientation vector. The final confidence of a hypothesis is voted as

$$confidence(h) = \frac{\mathcal{M}_\alpha + \mathcal{M}_d}{2}$$

Hypotheses generated by the above search are sorted so that the first 5 hypotheses with the highest confidence score (greater than 0.75) are accepted for further evaluation. Object in the testing image is not recognized if there is no hypothesis whose confidence value is greater than 0.75.

Hypothesis Verification - Hypotheses generated by the above matching process are subject to further verification. A template matching (similar to the approach used by Verly *et al.* [5]) can be used for this purpose. We do not use this template matching for the results reported in this paper.

5. Experiments

A set of testing data, with a total number of 1404 PWF ISAR images of 4 objects have been used for the experiments. The testing data are about 0.2^o offset in azimuth from their corresponding training data. The depression angle of the testing images is the same as the training images used in model construction. The set of testing images of each object is used to test the effectiveness of each of the 4 models obtained after training. Table 2 shows the confusion matrix of the experiments using persistent scatterer as features. For each image, at most 5 hypotheses are generated. To simplify the comparison, here only the hypothesis with the highest voting score is used. The voting scores of all the accepted hypotheses are greater than 0.75. All the parameters are initially fine tuned for recognizing bulldozers. These parameters are then used to recognize the other three objects without modification.

Table 2. Confusion matrix of recognizing 1404 ISAR images of 4 objects using the automatically constructed models.

Object	Camaro	Van	Pickup	Bdozer	None
Camaro	78	6	2	5	9
Van	9	69	4	8	10
Pickup	5	8	74	7	6
Bulldozer	2	3	5	84	6

Fig. 4 shows the accuracy of the hypotheses by the differences between the recognized azimuth angle and the ground-truth azimuth of the image. It shows the percentage of the hypotheses (H) whose recognized azimuth angle is β degrees ($\beta = 1^o, 2^o, 3^O, 4^o, 5^o$ or greater than 5^o) away from the ground-truth azimuth angles. Here only correct hypotheses (correct identification of the object) are shown. Hypotheses that fall within 5^o are deemed as good hypotheses. It is noted that there is only a small portion of the hypotheses whose recognized azimuth angles are more than 5^o away from the actual azimuth angle.

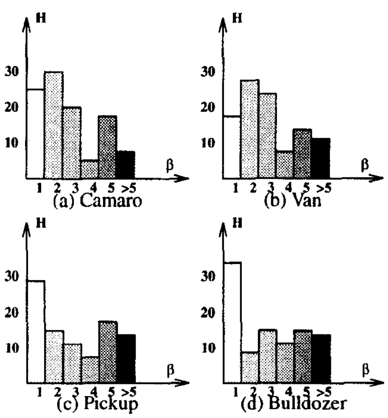

Figure 4. Accuracy of the hypotheses generated in the object recognition process.

To illustrate the effectiveness of the model based persistent scatterer extraction approach, we also applied the object recognition system using local maxima as feature vectors. The result is shown in Table 3.

Table 3. Confusion matrix of recognizing 1404 ISAR images of 4 objects using local maxima instead of persistent scatterers

Object	Camaro	Van	Pickup	Bdozer	None
Camaro	58	7	6	9	20
Van	9	48	13	7	23
Pickup	12	7	61	2	18
Bdozer	3	6	13	59	19

6. Conclusions

This paper presented a prototype of an automatic model construction scheme for object recognition systems using ISAR images. The major contributions of the paper are:

- A framework of a learning-from-examples approach for automatically constructing recognition model of various objects using ISAR images.
- Experimental results that show our approach using ISAR images is promising for model construction.

References

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

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

[3] K. Sato, K. Ikeuchi, and T. Kanade. Model based recognition of specular objects using sensor models. *CVGIP: Image Understanding*, 55(2):155–169, March 1992.

[4] S. M. Verbout, C. M. Netishen, and L. M. Novak. Polarimetric techniques for enhancing SAR imagery. In *SPIE Proceedings: Synthetic Aperture Radar*, volume 1630, pages 141–173, January 1992.

[5] J. G. Verly, R. L. Delanoy, and C. H. Lazott. Principles and evaluation of an automatic target recognition system for synthetic aperture radar imagery based on the use of functional templates. In *SPIE Proceedings: Automatic Target Recognition III*, volume 1660, Orlando, FL, April 1993.

[6] Donald R. Wehner. *Hight resolution radar.* Artech House, 1987.