

Recognition of Occluded Targets Using Stochastic Models

Bir Bhanu and Yingqiang Lin
Center for Research in Intelligent Systems
University of California, Riverside, CA 92521
Email: bhanu@cris.ucr.edu

Abstract

Recognition of occluded objects in synthetic aperture radar (SAR) images is a significant problem for automatic target recognition. In this paper, we present a hidden Markov modeling (HMM) based approach for recognizing objects in synthetic aperture radar (SAR) images. We identify the peculiar characteristics of SAR sensors and using these characteristics we develop feature based multiple models for a given SAR image of an object. The models exploiting the relative geometry of feature locations or the amplitude of SAR radar return are based on sequentialization of scattering centers extracted from SAR images. In order to improve performance we integrate these models synergistically using their probabilistic estimates for recognition of a particular target at a specific azimuth. Experimental results are presented using both synthetic and real SAR images.

Keywords: hidden Markov modeling, object recognition, multiple recognition models, rotation invariance, synthetic aperture radar images

1. Introduction

One of the critical problems for object recognition is that the recognition approach should be able to handle partial occlusion of object and spurious or noisy data [1]. In most of the object recognition approaches, the spatial arrangement of structural information of an object is the central part that offers the most important information. Under partial occlusion situations the recognition process must be able to work with only portions of the correct spatial information. Rigid template matching and shape-based recognition approaches depend on good prior segmentation results. But the structural primitive (e.g., line segments, point-

like features, etc.) extracted from occluded and noisy images may not have sufficient reliability, which will directly undermine the performance of those recognition approaches.

We want to suggest an object recognition mechanism that effectively makes use of all available structural information. Based on the nature of the problems caused by occlusion and noise, we view the spatial arrangement of structural information as a whole rather than view the spatial primitives individually. Because of its stochastic nature, a hidden Markov model (HMM) is quite suitable for characterizing patterns. Its nondeterministic model structure makes it capable of collecting useful information from distorted or partially unreliable patterns. Many successful applications of HMM in speech recognition [2, 3] and character recognition [4, 5] attest to its usefulness. Thus, it is potentially an effective tool to recognize objects with partial occlusion and noise.

However, the limit of traditional HMMs is that they are basically one dimensional models. So how to appropriately apply this approach to two dimensional image problems becomes the key. It has been largely an unsolved problem. In this paper we use the features based on the image formation process to encode the 2-D image into 1-D sequences. We use information from both the relative positions of the scattering centers and their relative magnitude in SAR images to address the fundamental issues of building object models and using them for robust recognition of objects in SAR images.

1.1 Overview of the approach

During an off-line phase, scattering centers are extracted from SAR images by finding local maxima of intensity. Both locations and magnitudes of these peak features are used in the approach. These features are viewed as *emitting patterns* of some hidden stochastic process. Multiple observation sequences based on

both the *relative* geometry and amplitude of SAR return signal are used to build the bank of stochastic models. These models provide robust recognition in the presence of occlusion and unstable features caused by speckle phenomena where some of the features may appear/disappear at random in an image. At the end of the off-line phase, hidden Markov recognition models for various objects and azimuths are obtained. Similar to the off-line phase, during the on-line phase features are extracted from SAR images and observation sequences based on these features are matched by the HMM forward process with the stored models obtained previously. Maximum likelihood decision is made on the classification results. Now the results obtained from multiple models are combined in a voting kind of approach that uses both the object, azimuth label and its probability of classification. This produces a rank ordered list of classifications of the test image and associated confidences.

1.2 Related work and our contribution

Fielding and Ruck [6] have used HMM models for spatio-temporal pattern recognition to classify moving objects in image sequences. Rao and Mersereau [7] have attempted to merge HMM and deformable template approaches for image segmentation. Template matching [8] and major axis based approaches [9] have been used to recognize and identify objects in SAR images, however, they are not suitable to recognize occluded objects. Kottler et al. propose a HMM-Based SAR ATR system [10]. They first segment the SAR image, and then extract features followed by Radon transforms. The feature sequences so obtained are input to HMMs.

The contributions of this paper are:

- Hidden Markov modeling approach commonly used for recognizing 1-D speech signals is applied in a novel manner to 2-D SAR images to solve the occluded object recognition problem.
- Unlike most of the work for model building in pattern recognition and computer vision, our recognition models using hidden Markov modeling concepts are based on the peculiar characteristics of SAR images where the number of models used for recognition is justified by the quantification of the azimuthal variance in SAR images.
- Multiple models derived from various observation sequences, based on both the relative geometry and signal amplitude (four based on geometry and one based on amplitude), are used to capture

the unique characteristics of patterns to recognize objects.

- Extensive amounts of data is used to test the approach for recognition of objects for various amounts of occlusion (10–50%) in both synthetic and real data.

2 Hidden Markov Modeling Approach

It is well known that HMM can model speech signals well [2, 3]. It is a model used to describe a doubly stochastic process which has a set of states, a set of output symbols and a set of transitions. Each transition is from state to state and associated with it are a probability and an output symbol. The word ‘hidden’ means that although we observe an output symbol, we cannot determine which transition has actually taken place. At each time step t , the state of the HMM will change according to a transition probability distribution which depends on the previous state and an observation y_t is produced according to a probability distribution which depends on the current state.

Formally a HMM is defined as a triple $\lambda = (A, B, \pi)$, where a_{ij} is the probability that state i transits to state j , $b_{ij}(k)$ is the probability that we observe symbol k in a transition from state i to state j , and π_i is the probability of i being the initial state. Figure 1 shows an example of a N states HMM.

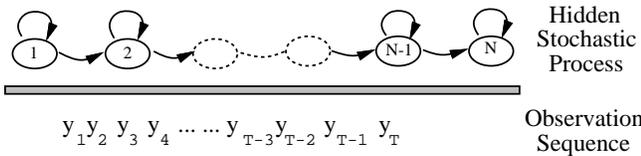
Recognition Problem – Forward Procedure: The HMM provides us a useful mechanism to solve the problems we face for robust object recognition. Given a model and a sequence of observations, the probability that the observed sequence was produced by the model can be computed by the forward procedure [11]. Suppose we have a HMM $\lambda = \{A, B, \pi\}$ and an observation sequence y_1^T . We define $\alpha_i(t)$ as the probability that the Markov process is in state i , having generated y_1^t .

$$\begin{aligned} \alpha_i(t) &= 0, \text{ when } t=0 \text{ and } i \text{ is not an initial state.} \\ \alpha_i(t) &= 1, \text{ when } t=0 \text{ and } i \text{ is an initial state.} \\ \alpha_i(t) &= \sum_j [\alpha_j(t-1)a_{ji}b_{ji}(y_t)], \text{ when } t > 0. \end{aligned} \quad (1)$$

The probability that the HMM stopped at the final state and generated y_1^T is $\alpha_{S_F}(T)$. The forward procedure is given below:

Let T be the length of an observation sequence and N is the number of states in the HMM.

1. Initialize $\alpha_i(0)$, where $i = 1, 2, \dots, N$.
2. Compute $\alpha_i(t)$ inductively (equation 1), where $t = 1, 2, \dots, T$. At each step, the previously com-



- N: the number of states.
M: the number of distinct observable symbols.
A: a_{ij} is the probability that state i will transit to state j .
B: $b_{ij}(k)$ is the probability that symbol k will be observed when there is a transition from state i to state j .
 π : π_i is the probability that state i is the initial state.

Figure 1. A N states forward-type HMM

puted $\alpha_i(t-1)$ is used. Repeat this process until t reaches T .

3. Output $\alpha_{S_F}(T)$, where $\alpha_{S_F}(T)$ is the probability that the HMM stopped at the final state and generated the observation sequence.

Usually, α becomes too small to be represented in computer after several iterations. We take the logarithm of the α value in the computation.

Training Problem – Baum-Welch Algorithm: To build a HMM is actually an optimization of the model parameters so that it can describe the observation better. This is a problem of training. The Baum-Welch re-estimation algorithm is used to calculate the maximum likelihood model. But before we use the Baum-Welch algorithm, we must introduce the counterpart of $\alpha_i(t) : \beta_i(t)$, which is the probability that the Markov process is in state i and will generate y_{t+1}^T .

$$\begin{aligned} \beta_i(t) &= 0, \text{ when } t=T \text{ and } i \text{ is not a final state.} \\ \beta_i(t) &= 1, \text{ when } t=T \text{ and } i \text{ is a final state.} \\ \beta_i(t) &= \sum_j [a_{ij} b_{ij}(y_{t+1}) \beta_j(t+1)], \text{ when } 0 \leq t < T. \end{aligned} \quad (2)$$

The probability of being in state i at time t and state j at time $t+1$ given observation sequence y_1^T and the model λ is defined as follows:

$$\begin{aligned} \gamma_{ij}(t) &= P(X_t = i, X_{t+1} = j | y_1^T) \\ &= \frac{\alpha_i(t-1) a_{ij} b_{ij}(y_t) \beta_j(t)}{\alpha_{S_F}(T)} \end{aligned} \quad (3)$$

Now the expected number of transitions from state i to state j given y_1^T at any time is simply $\sum_{t=1}^T \gamma_{ij}(t)$ and the expected number of transitions from state i to any state at any time is $\sum_{t=1}^T \sum_k \gamma_{ik}(t)$. Then, given some initial parameters, we could recompute $\overline{a_{ij}}$, the probability of taking the transition from state i to state

j as:

$$\overline{a_{ij}} = \frac{\sum_{t=1}^T \gamma_{ij}(t)}{\sum_{t=1}^T \sum_k \gamma_{ik}(t)} \quad (4)$$

Similarly, $\overline{b_{ij}(k)}$ can be re-estimated as the ratio between the frequency that symbol k is emitted and the frequency that any symbol is emitted:

$$\overline{b_{ij}(k)} = \frac{\sum_{t: y_t=k} \gamma_{ij}(t)}{\sum_{t=1}^T \gamma_{ij}(t)} \quad (5)$$

It can be proved that the above equations are guaranteed to increase $\alpha_{S_F}(T)$ until a critical point is reached, after which the re-estimate will remain the same. In practice, we set a threshold as the ending condition for re-estimation.

So the whole process of training a HMM is as follows:

1. Initially, we have only an observation sequence y_1^T and blindly set (A, B, π) .
2. Use y_1^T and (A, B, π) to compute α and β (equations 1, 2).
3. Use α and β to compute γ (equation 3).
4. Use y_1^T , (A, B, π) , α , β and γ to compute A and B (equations 4, 5). Go to step 2.

A HMM is able to handle pattern distortions and the uncertainty of the locally observed signals, because of its nondeterministic nature. However, a HMM is primarily suited for sequential, one-dimensional patterns and it is not obvious that how a HMM can be used on 2-D patterns in object recognition. The basic ideas to apply a HMM for our purpose are (a) training the HMM λ by samples of SAR images of a certain object, and (b) recognizing an unknown object in a given SAR image. These two problems are addressed in the following. The key questions are what we shall use as observation data and how we get the observation sequences.

3 Hidden Markov Models for SAR Object Recognition

Scattering centers (location and magnitude) extracted from SAR images are used to train and test models for recognition. At six inch resolution, there exist a large number of peaks corresponding to scattering centers. We have selected peaks as features in this work since we wanted to evaluate the limits of our approach using six inch resolution XPA TCH and one foot resolution MSTAR data before more complicated

features are used. We consider a pixel as a scattering center if the magnitude of SAR return at this pixel is larger than all its eight neighbors.

Unlike the visible images, SAR images are extremely sensitive to slight changes in viewpoint (azimuth and depression angle) and are not affected by scale [12]. We evaluate [13] the characteristics of scattering centers to find out what kind of location invariance exists among scattering centers. We find that scattering centers for SAR images vary greatly with relatively small changes of azimuth angles. As a result, to recognize occluded objects we represent an object at a given depression angle by 360 azimuths taken in steps of 1° whenever possible. The squint angle, the angle between the flight path and radar beam, is known (90° here) and kept fixed for all the image data used in this paper.

3.1 Extraction of Observation Sequences

After the scattering centers are extracted, we need to encode the data into a 1-D sequence as the input to a recognition model based HMM process. It is one of the *key* factors which affects the performance of a HMM modeling approach for object recognition. There are many ways to choose observation sequences, but we want to use information from both the magnitude and the relative spatial location of the scattering centers extracted from a SAR image. Also the sequentialization method should not be affected by distortion, noise, or partial occlusion and should be able to represent the image efficiently. Based on the above considerations, we employ two approaches to obtain the sequences.

- Sequences based on amplitudes: $O_1 = \{Magnitude_1, Magnitude_2, \dots, Magnitude_n\}$, where $Magnitude_i$ is the amplitude of i th scattering center.
- Sequences based on *relative* geometrical relationships:

$$O_2 = \{d(1, 2), d(2, 3), \dots, d(n, 1)\} \text{ (length } n)$$

$$O_3 = \{d(1, 2), d(1, 3), \dots, d(1, n)\} \text{ (length } n - 1)$$

$$O_4 = \{d(2, 1), d(2, 3), \dots, d(2, n)\} \text{ (length } n - 1)$$

$$O_5 = \{d(3, 1), d(3, 2), \dots, d(3, n)\} \text{ (length } n - 1),$$
 where $d(i, j)$ is the Euclidean distance between scattering centers i and j .

Sequence O_1 is obtained by sorting the scattering centers by their magnitude. We label the scattering centers 1 through n in descending order. So in this approach, we do not use the location information and thus can avoid the instability caused by the error in localization of scattering centers. Sequences O_2 through O_5 are obtained based on the relative locations of the

scattering centers and magnitude of the scattering centers is *not used*. In experiments described in section 4, we consider a certain number of scattering centers (sorted in descending order of their magnitude). This is because we expect that the scattering centers with larger magnitude are relatively more stable than the weaker ones.

Since we use discrete HMMs, each element in the sequence should be converted to an observation symbol. It is like a label from 1 to K that represents the symbols which can be observed for a HMM. We use the K -means algorithm [14] to classify the magnitude values (or distance values) of all the scattering centers in the database into K classes (K is experimentally determined). Once we know to which class each of the elements of a sequence belongs, we label the element with the label of its class. Thus, for a given sequence, we obtain a sequence of observation symbols.

3.2 Training and Testing Phases

The procedure for building the model base is described as follows:

1. Loop (for a given depression angle) lines 2-4 for each object and each azimuth angle.
2. Generate images which simulate occlusion with scattering centers occluded from different directions (see Section 4).
3. Loop line 4 for each image generated by line 2.
4. Use Baum-Welch algorithm to re-estimate the HMM parameters. (Exit 3 - 4 loop when there is no further change in parameter values.)

The recognition procedure is described as follows:

1. Loop lines 2-3 for all the testing observation sequences.
2. Loop line 3 for all the models in the model base.
3. Feed the observation sequence into the model, $(A, B, \Pi)_{(M_i^*, a_j^*)}$, Use Forward algorithm to compute the probability that this sequence is produced by this model.
4. The model with maximum probability of an observation sequence is selected as the best match.

4 Experiments on XPATCH data

4.1 XPATCH SAR Data

We use XPATCHSAR simulator [15] to generate the data to perform controlled experiments. In our

research, we selected six inch resolution data because it is possible to collect real data at this resolution and also super resolution techniques [16] exist to achieve such data. Also we wanted to see how well we can solve the occluded target recognition problem at this resolution. We generate one set of SAR images of 5 objects (Fred tank, SCUD missile launcher, T72 tank, T80 tank and M1a1 tank) at 15° depression angle, and 90° squint angle (fixed), at each of the azimuth angles from 0° to 359° . We extract the 20 scattering centers (local maxima) with largest magnitudes. In the experiments, since we want to test the performance of our approach under partial occlusion and spurious data, we simulate realistic occlusion situations and generate images for training and testing.

Simulating occlusion: There does not exist an acceptable model for occlusion for automatic object recognition. We consider the occlusion to occur possibly from 9 different directions as shown in Figure 2. Scattering centers being occluded are not available. Moreover, we add some spurious data into the image. For instance, 20 scattering centers are shown in each image of Figure 2. They are obtained by removing 4 scattering centers (20% occlusion) from the center of one object or from one particular direction (simulated occlusion) and adding 4 spurious scattering centers into the image. The spurious scattering centers are added based on the following rules:

- The location of the scattering center is generated as a pair of random numbers.
- The magnitude of the scattering center depends on a random number r between 1 and 50. We use the magnitude of the r th brightest scattering center as the magnitude of the spurious scattering center.

Training Data: Based on the method of simulating occlusion described above, we generate 90 images from the original image (10 samples for each of 9 directions) at 5% occlusion and another 90 images at 10% occlusion. Including the original image, we have 181 images per object per azimuth angle to train multiple HMM models. Thus, we have a total of 99,000 (5 objects, 360 azimuths, 55 occluded images) samples for training.

Testing Data: We generate one image with o scattering centers occluded ($o = 2, 4, 6, 8$ or 10) from direction d ($d = 0, 1, \dots, 8$) per azimuth angle per object. So there are 1800 images (5 objects \times 360 degrees) generated for testing of occlusion with o scattering centers occluded from direction d . Thus, we have a total of 81,000 (5 objects, 360 azimuths, 5 different occlusions 10% – 50%, and 9 directions) samples for testing.

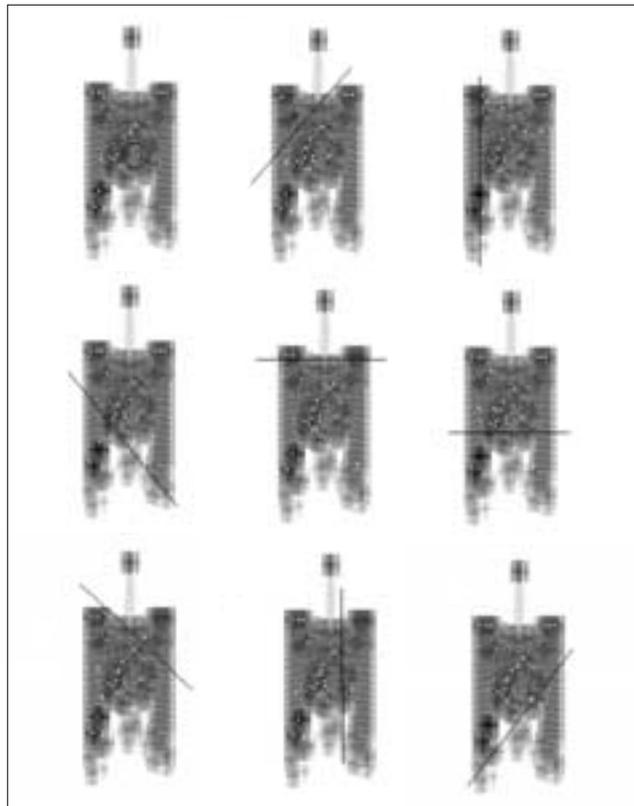


Figure 2. Scattering centers of T72 tank at azimuth 0° , part of scattering centers are occluded from a particular direction (0-8, left to right, top to bottom).

4.1.1 Training— Building Bank of HMM Models for Recognition

We performed experiments to choose the optimum of number of states and number of symbols of the HMM. We use data from five azimuth angles of five objects (Fred tank, SCUD missile launcher, T80 tank, T72 tank, and M1a1 tank). We find that with the increase in the number of states and symbols, recognition performance increases. Considering both the recognition performance and the computation cost, we choose 8 states and 32 symbols as the optimal number of states and symbols for our HMM models. Figure 3 illustrates example parameters of a 5 state, 4 symbol HMM.

Using the algorithm in Section 3, we built recognition models. For a selected sequence type we have 1800 ($= 360$ azimuths \times 5 object classes) HMM models. Since we have defined five kinds of observation sequences for each image (O_1, O_2, O_3, O_4, O_5), we get

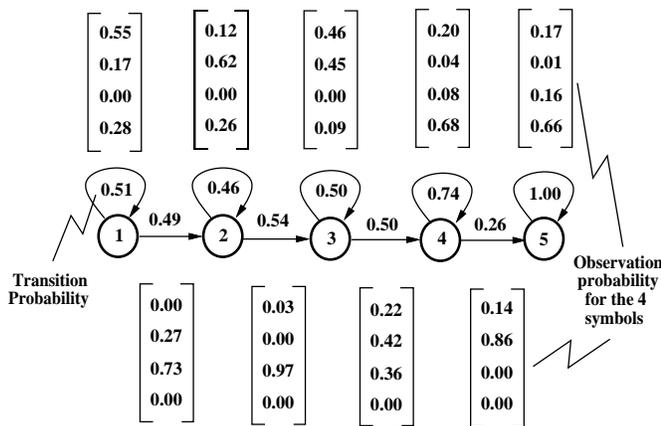


Figure 3. An example: parameters of a 5 states, 4 symbols HMM. The number on edges represents the transition probability, and the vector associated with each transition represents $b_{ij}(k)$. In our case, we use HMM with 8 states, 32 symbols

models based on each kind of observation sequence.

4.1.2 Testing Results

During testing phase, for a given observation sequence type each of the 81,000 testing images is tested against all models (1800 models: 5 objects, each has 360 models for each azimuth angle). If the model with the maximum probability is the model which produced the sequence, we count it as one correct recognition (object type and its pose). Otherwise, we count it as one incorrect recognition. After we get the results on images with scattering centers occluded from all 9 directions, we average these results and associate this recognition performance with the selected model for a given percent of occlusions.

Figure 4 shows the testing results for each of the five kinds of sequences: O_1, O_2, \dots, O_5 (section 3.1). The top curve, a dotted line, is the percentage that the test case object and pose is among the top ten recognition results, and the lower curve, in solid line, indicates the percentage that the recognition result with the highest probability is the same as the test case object and pose.

4.1.3 Integration of results from multiple sequences

Since not all models based on various sequences for a particular object and azimuth will provide optimal

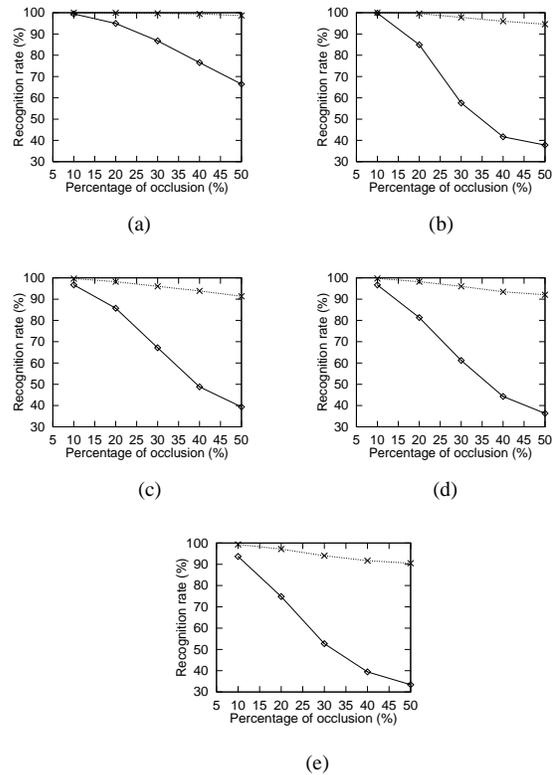


Figure 4. Recognition rate vs. percentage of occlusion for HMM models based on (a) O_1 , (b) O_2 , (c) O_3 , (d) O_4 , and (e) O_5 .

recognition performance under occlusion, noise, etc., we improve the recognition performance by combining the results obtained from all five kinds of models. Before discussing the approach for integration, we ask the question that if one testing image cannot be recognized correctly by model based on a particular sequence, say O_2 , can it be recognized correctly by models based on other kinds of sequences? The answer to this question is yes as the following results demonstrate.

The results of using models based on O_1 to O_5 are shown in Table 1. This table shows how many incorrect recognitions, made by using models based on sequence O_2 , can be correctly recognized (“captured”) by models based on other sequences. We define the “upper bound” as the highest possible recognition performance that can be achieved using the 5 kinds of models considering *only* the top candidate for recognition from each of the models. The total number of errors corresponding to “upper bound” are shown in the 7th column of the table.

We draw two curves (Figure 5(a)) to show the possible “upper bound” and “lower bound” of recognition

Table 1. Testing results for occluded object recognition using of 81,000 testing cases. Results based on integration of O_1 to O_5 .

Per cent. occlusion	Errors with model O_2	Errors Captured by models				Errors using models O_1 to O_5	% Correct R ecognition (“upper bound”)	% Based on Inte gration R ecognition	% Based on Inte gration Indexing
		O_1	O_3	O_4	O_5				
10%	4	0	1	0	1	2	100.0	99.9	99.9
20%	271	19	53	74	101	121	99.6	98.9	99.6
30%	763	111	294	339	418	144	98.6	93.4	97.6
40%	1050	265	580	629	675	79	95.6	79.4	91.8
50%	1119	397	726	755	784	37	91.8	62.2	83.3
<i>Average Recognition Rate</i>							97.1	86.8	94.4

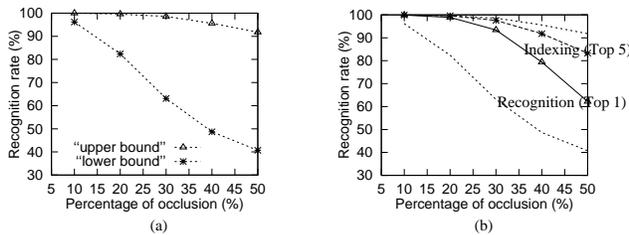


Figure 5. (a) “Upper” and “lower” bound of recognition rate vs. percentage of occlusion. (b) Performance of integrated models: using integrated models O_1 to O_5 . The results for recognition (Top 1) and indexing (Top 5) candidates are superimposed on the figure shown in (a).

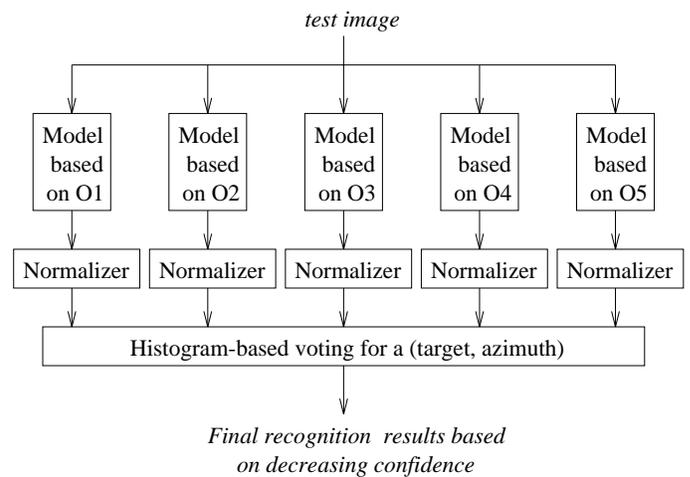


Figure 6. Integration of results by histogram-based method.

rate we can achieve based on the 5 kinds of models. The curve on the top is obtained by considering all 5 kinds of models, if one of them can correctly recognize the test data, we count it as a correct recognition. The “lower bound” or the bottom curve is the worst recognition result out of the five models.

We have developed a histogram-like method (shown in Figure 6) to integrate the results from models based on 5 different sequences. The algorithmic steps are:

1. For each test image, we collect the ten highest possibilities in the testing results corresponding to each of the sequences O_1, O_2, \dots, O_5 .
2. A normalization is done to the ten probabilistic estimates corresponding to each of the sequences. So we have 50 normalized numbers for each test image.

3. We draw a histogram with probability vs. object and pose (here we combine object and pose as one parameter). This is because each number corresponds to an object and a pose. The number is the probability that the test image is the image of that object at that pose.
4. If the object associated with the highest probability in the histogram is the same as the groundtruth, we count it as one correct recognition.

The second curve from the bottom in Figure 5(b) is the result for recognition. The corresponding confusion matrix for various amounts of occlusion is shown

Table 2. Confusion Matrix for 5 objects classes at varying amounts of occlusion (10% – 50%).

	% Occlusion	<i>Fred</i>	<i>SCUD</i>	<i>T72</i>	<i>T80</i>	<i>M1a1</i>
<i>Fred</i>	10	100.0	0.0	0.0	0.0	0.0
	20	99.2	0.0	0.1	0.4	0.3
	30	95.9	0.2	0.6	1.9	1.4
	40	87.1	0.7	2.8	5.5	3.9
	50	73.2	1.6	7.1	12.1	6.0
<i>SCUD</i>	10	0.0	100.0	0.0	0.0	0.0
	20	0.0	99.7	0.2	0.1	0.0
	30	0.9	97.3	1.2	0.4	0.3
	40	3.1	88.8	4.9	1.9	1.3
	50	5.6	77.9	11.9	2.7	1.9
<i>T72</i>	10	0.0	0.0	100.0	0.0	0.0
	20	0.4	0.2	99.2	0.1	0.2
	30	2.4	0.5	95.3	1.1	0.6
	40	9.1	2.1	82.5	3.8	2.4
	50	16.8	5.2	65.9	6.8	5.4
<i>T80</i>	10	0.0	0.0	0.0	100.0	0.0
	20	1.2	0.0	0.1	98.6	0.1
	30	6.9	0.0	0.6	91.1	1.4
	40	21.5	0.1	1.6	72.6	4.2
	50	37.4	0.8	3.1	50.9	7.8
<i>M1a1</i>	10	0.0	0.0	0.0	0.0	100.0
	20	1.6	0.0	0.1	0.3	98.0
	30	8.5	0.2	0.7	2.9	87.8
	40	22.5	0.8	2.0	8.5	66.1
	50	36.9	1.1	5.2	13.8	42.9

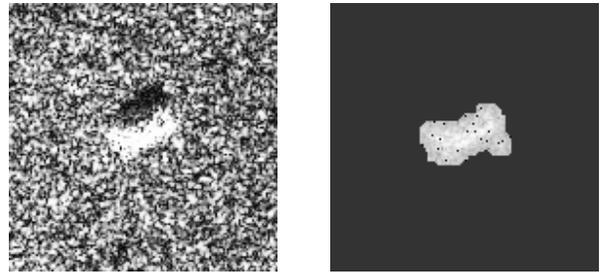
in Table 2. On the average, we find 80.35% correct recognition performance when the objects are occluded from 10 – 50%. The second curve from the top in Figure 5(b) is obtained by counting a correct indexing result when the ground truth is in the objects associated with the highest 5 probabilities in the histogram. For the purpose of comparison, we have also superimposed the curves of Figure 5(a) into Figure 5(b) with “lower/upper” bounds. Considering the correct indexing answer in the top 5 responses, the average performance is 93.3% for 5 objects occluded from 10% – 50%. Thus, our method of integration produces good results in comparison to “upper bound” which is 95.3% for 5 objects for 10% – 50% occlusion.

4.2 Real SAR Data

The methods used here are the same as those used in the previous subsection. The only difference is that here we experiment on real data instead of XPA TCH data. We use MSTAR public real SAR images (at one foot resolution and depression angle 15°) of 2 objects (T72 tank with serial number #a64, shown in Figure

7, and ZSU tank with serial number #d08). Ideally, we can have 360 object models for each azimuth for each object. However, we don’t have 360 SAR images for each object in the MSTAR data set. For the T72 tank, there are 288 images available for different azimuths. Also for the ZSU tank, 288 images are available. Thus, each object consists of 288 azimuths which we call object models. Each object model consists of HMM models based on observation sequences (O_1 to O_5). We extract 30 scattering centers with largest magnitudes from each SAR image.

We consider the occlusion to occur possibly from 9 different directions (center, 4 sides and 4 corners of the image). Scattering centers being occluded are not available. Moreover, we add back into the image at random locations a number of spurious scattering centers, equal to the number of occluded scatterers, of random magnitude. The random magnitude could be equal to the magnitude of any of the top 30 scatterers. For example, for 30% occlusion, we remove 9 scattering centers from the center of one object or from one particular direction and add randomly 9 spurious scattering centers back into the image. We compute the observation sequences based on the scattering centers available after the occlusion process has taken place.



(a) SAR image of a T72 tank

(b) Features extracted for T72 tank

Figure 7. Real SAR images and region of interests (ROIs) (with peaks shown as black dots superimposed on the ROI) for T72 tank #a64.

4.2.1 Training and Testing Data

Training Data: We generate 91 training sequences of each type (O_1 to O_5) from each SAR image. The first one is obtained from the original SAR image without occlusion. Then we occlude the SAR image from 9 directions. For each direction, the occlusion level is 5% and 10%. For each occlusion level, we extract 5 training observation sequences. So 91 sequences are generated from each image. We have two objects and 288 SAR images of each object, thus the number of

Table 3. Confusion matrix for various occlusions using MSTAR data

T arget Type	Occlusion Level			
	20%		30%	
	T72	ZSU	T72	ZSU
T72	280(97.2%)	8(2.8%)	233(80.9%)	55(19.1%)
ZSU	5(1.7%)	283(98.3%)	64(22.2%)	224(77.8%)
T arget Type	Occlusion Level			
	40%		50%	
	T72	ZSU	T72	ZSU
T72	200(69.4%)	88(30.6%)	164(56.9%)	124(43.1%)
ZSU	98(34.0%)	190(66.0%)	117(40.6%)	171(59.4%)

Table 4. Recognition results for various occlusions using MSTAR data

T arget Type	Occlusion Level					
	20%			30%		
	correct	error	rejection	correct	error	rejection
T72	279(96.9%)	7(2.4%)	2(0.7%)	230(79.9%)	47(16.3%)	11(3.8%)
ZSU	281(97.6%)	4(1.4%)	3(1.0%)	216(75.0%)	59(20.5%)	13(4.5%)
T arget Type	Occlusion Level					
	40%			50%		
	correct	error	rejection	correct	error	rejection
T72	188(65.3%)	75(26.0%)	25(8.7%)	156(54.2%)	110(38.2%)	22(7.6%)
ZSU	168(58.3%)	91(31.6%)	29(10.1%)	164(56.9%)	102(35.5%)	22(7.6%)

training sequences of each type (O_1 to O_5) is 52,416. Since there are 5 kinds of observation sequences, the total number of sequences is 262,080.

Testing Data: From each SAR image, we generate 36 testing sequences of each type (O_1 to O_5). We occlude the SAR image from 9 directions. For each direction, the occlusion level is from 20% to 50% with 10% increment. That is, the numbers of occluded scattering centers are 6, 9, 12, and 15 respectively. Thus, we have total 20736 testing sequences of each type. When testing, we only use the sequences which are obtained when the occlusion was from direction 7, which is the direction from the right side of the image. So, for each occlusion level, we have 576 testing sequences. Since there are 5 kinds of observation sequences, the total number of sequences for each occlusion level is 2880.

4.2.2 Experiment Results

Tables 3 and 4 show the experimental results on MSTAR real data. These results are obtained by integrating the results from 5 different type of sequences O_1 , O_2 , O_3 , O_4 , and O_5 .

Table 3 shows the results from recognizing 20% to 50% occluded T72 and ZSU tank. The confusion matrix shows how many of them are correctly identified and how many are incorrectly recognized. In this experiment, we are concerned only with the identity of the object. The test object is the type of the model with maximum probability.

Table 4 shows the results similar to Table 3. The difference is that here we use “probability ratio threshold” instead of considering only the maximum probability. Only when the ratio between the maximum and second maximum (other object type) probabilities is greater than the probability ratio threshold (1.01 used

here) we accept the recognition result. Otherwise, it is rejected and the test data is labeled as unknown object. The above results show that recognition results are somewhat satisfactory, especially when the occlusion is below 40%.

5 Conclusions

Recognition of occluded object has been a significant problem for automatic target recognition. In this paper, we have presented a conceptual approach for the recognition of occluded objects in SAR images. The approach uses multiple HMM based models for various observation sequences that are chosen based on the SAR image formation and account for both the geometry and magnitude of SAR image features. We have shown the results on both XPATCH and real SAR data. The number of observation sequences and the number of features are design parameters which can be optimized by following the approach presented in the paper.

Acknowledgements: This research is supported in part by grant F49620-97-1-0184. The contents and information do not necessarily reflect the positions and policies of the U.S. Government.

References

- [1] B. Bhanu, D. E. Dudgeon, E. G. Zelnio, A. Rosenfeld, D. Casasent, and I. S. Reed. Introduction to the special issue on automatic target detection and recognition. *IEEE Trans. Image Processing* 6(1):1-6, January 1997.
- [2] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of the IEEE* 77(2):257-285, Feb 1989.
- [3] L. R. Rabiner, B. H. Juang, S. E. Levinson, and M. M. Sondhi. Recognition of isolated digits using hidden Markov models with continuous mixture densities. *AT&T Technic aJournal*, 64(6):1211-1233, July-August 1985.
- [4] O. E. Agazzi and S. S. Kuo. Hidden markov model based optical character recognition in the presence of deterministic transformations. *Pattern Recognition*, 26(12):1813-1826, November 1993.
- [5] O. E. Agazzi and S. S. Kuo. Pseudo two-dimensional hidden Markov models for document recognition. *AT&T Technic aJournal*, 72(5):60-72, 1993.
- [6] K. H. Fielding and D. W. Ruck. Spatio-temporal pattern recognition using hidden Markov models. *IEEE Trans. Aerospace and Electronic Systems*, 31(4):1292-1300, 1995.
- [7] R. R. Rao and R. M. Mersereau. On merging hidden Markov models with deformable templates. In *Proc. International Conf. on Image Processing* pages 556-559, 1995.
- [8] 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.
- [9] J. H. Yi, B. Bhanu, and M. Li. Target indexing in SAR images using scattering centers and Hausdorff distance. *Pattern Recognition Letters* 17:1191-1198, September 1996.
- [10] D. P. Kottle, P. D. Fiore, K. L. Brown, and J. K. Fwu. A design for HMM-based SAR ATR. In *SPIE Conf. on Algorithms for Synthetic Aperture Radar Imagery V*, pages 541-551, April 1998.
- [11] K. F. Lee. *Automatic Speech Recognition - The development of the SPHINX system*. Kluwer Academic Publishers, Boston, 1989.
- [12] J. P. Fitch. *Synthetic Aperture Radar*. New York: Springer-Verlag, 1988.
- [13] G. Jones III and B. Bhanu. Recognition of articulated and occluded objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21(7):603-613, July 1999.
- [14] J. T. Tou and R. C. Gonzalez. *Pattern Recognition Principles*. Addison-Wesley, London, 1974.
- [15] D. J. Andersh, S. W. Lee, H. Ling, and C. L. Yu. XPATCH: A high frequency electromagnetic scattering prediction code using shooting and bouncing ray. In *Proc. Ground Target Modeling and Validation Conference*, pages 498-507, August 1994.
- [16] L. Novak, G. Benitz, G. Owirka, and L. Besette. ATR performance using enhanced resolution SAR. In *SPIE Conf. on Algorithms for Synthetic Aperture Radar Imagery*, April 10 - 12 1996.