Multiple Stochastic Models for Recognition of Occluded Targets in SAR Images

Bir Bhanu and **Bing Tian**
College of Engineering
University of California, Riverside, California 92521
E-mail: bhanu@shivish.ucr.edu
URL: http://constitution.ucr.edu

Abstract

The correctness of results of structural object recognition approaches largely depends on the reliability of the features extracted from the image data. However, this cannot be satisfied in many practical situations where the applications require robust recognition during day/night under high clutter. Stochastic models provide some attractive features for pattern matching and recognition under partial occlusion and noise. In this paper, we present a hidden Markov modeling (HMM) based approach for recognizing objects in synthetic aperture radar (SAR) images. We develop multiple models for a given SAR image of an object and integrate these models synergistically using their probabilistic estimates for recognition. The models are based on sequentialization of scattering centers extracted from SAR images. Experimental results are presented using 99,000 training samples and 81,000 testing samples for 5 classes. We achieved better than 87% correct recognition performance when the objects are up to 30% occluded.

1 Introduction

One of the critical problems for object recognition is that the recognition process has to be able to handle partial occlusion of the object and spurious or noisy data. In most of the object recognition approaches, the spatial arrangement of structural information of the 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.

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

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 [1, 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 both from the relative positions of the scattering centers and their relative magnitude in SAR images [6]. In this paper we 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

Figure 1 provides an overview of the HMM based approach for recognition of occluded objects in SAR imagery. 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 relative amplitude of SAR return signal (obtained as a result of the physics of the SAR image forma-

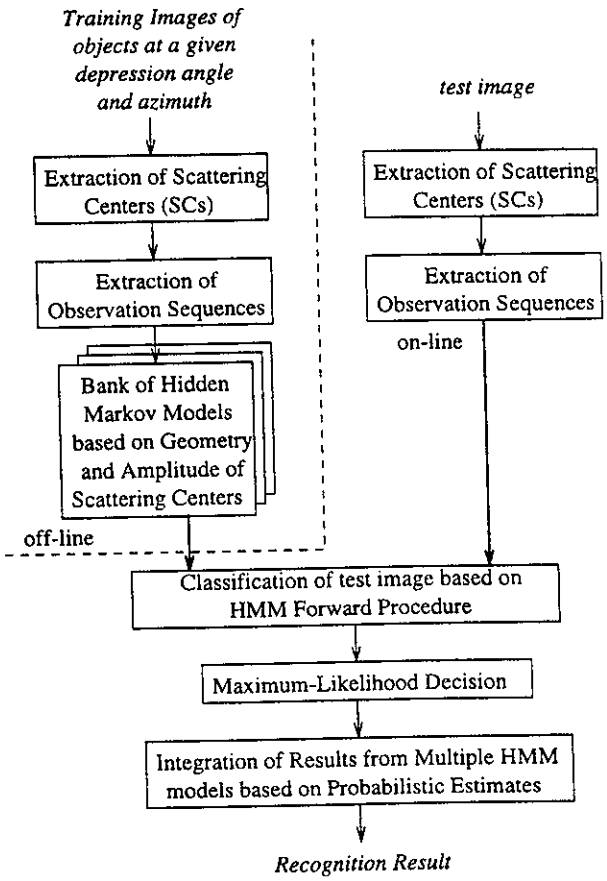


Figure 1: The HMM-Based approach for recognition of occluded objects.

tion process) are used to build the bank of stochastic models to provide robust recognition in the presence of severe occlusion and unstable features caused by scintillation 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

There is no published work on object recognition using HMM models. Fielding and Ruck [7] have used HMM models for spatio-temporal pattern recognition to classify moving objects in image sequences. Rao and Mersereau [8] have attempted to merge HMM and deformable template approaches for image segmentation. Template matching [9] and major axis based approaches [10] have been used to recognize and index objects in SAR images, however, they are not suitable to recognize occluded objects. Recently, invariant histogram in conjunction with template matching have also been used to recognize occluded objects in SAR images [11].

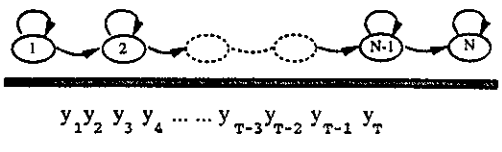
The original contributions of this paper are:

- 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.
- Multiple models derived from various observation sequences, based on both the geometry and signal amplitude are used to capture the unique characteristics of patterns to recognize objects.
- Unlike most of the work for model building in computer vision, our recognition models using hidden Markov modeling concept are based on the peculiar characteristics of SAR images where the number of models used for recognition is scientifically justified by the quantification of the azimuthal variance in SAR images.
- Extensive amounts of data (99,000 training samples and 81,000 testing samples obtained from 1800 images generated by the well known XPATCH SAR simulator [12] that uses 3-D CAD models of objects) is used to test the approach for recognition of objects for various amounts of occlusion (10−50%) and good recognition performance is obtained.

2 Hidden Markov Modeling Approach

It is well known that HMM can model speech signals well [1, 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



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 2: A N states forward-type HMM

is the probability of i being the initial state. Figure 2 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 [13]. 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.

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

The probability that the HMM stopped at the final state and generated y_1^T is $\alpha_{S_F}(T)$. After initialization of α, we compute it inductively. At each step the previously computed α is used, until the t reaches T. $\alpha_{S_F}(T)$ is the sum of probabilities of all paths of length T.

Usually, α will become 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.

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

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:

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

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

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

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:

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

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

3.1 Extraction of Scattering Centers

Scattering centers (location and magnitude) extracted from SAR images are used to train and test models for recognition. We consider a pixel as a scattering center if the magnitude of SAR return at this

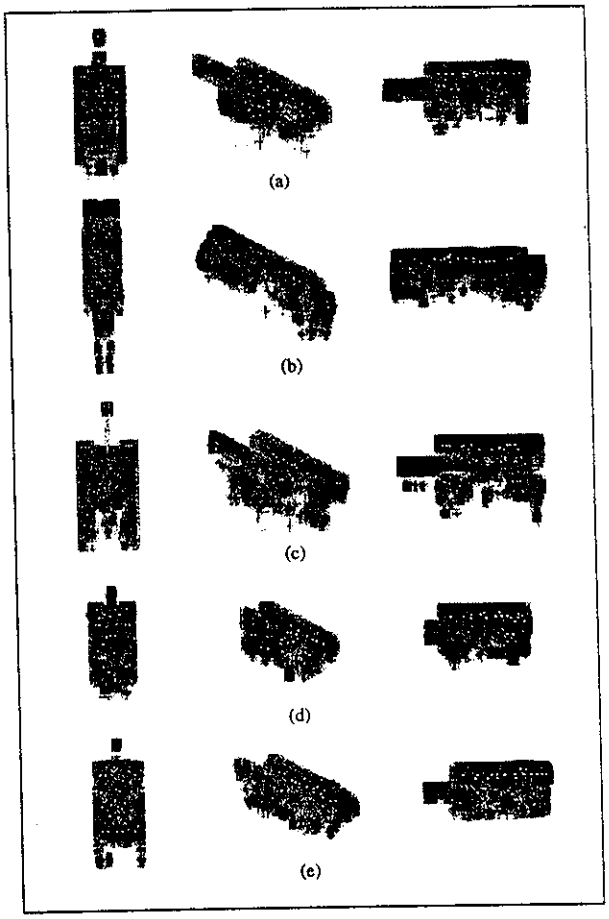


Figure 3: Examples of scattering centers (white dots) extracted from SAR images at azimuths $0°, 60°, 90°$. (a) Fred tank, (b) SCUD launcher with missile down, (c) T72 tank, (d) T80 tank, (e) M1a1 tank.

pixel is larger than all its eight neighbors. Figure 3 shows some examples of scattering centers extracted from SAR images (6″ resolution) of various objects at 15° depression angle and azimuths at 0°, 60°, and 90°.

3.2 Rotation Variance of Scattering Centers and Representation of 3-D Objects

Unlike the visible images, SAR images are extremely sensitive to slight changes in viewpoint (azimuth and depression angle) and are not affected by scale [14]. We evaluate the characteristics of scattering centers to find out what kind of location invariance exists among scattering centers. Figure 4(a) shows the rotation invariance for T72 tank. The data is obtained by rotating the image at azimuth $i°$ (for a fixed depression angle) by $x°$ (x from 1 to 10), and comparing the rotated image with the image of $(i+x)°$ to see how many scattering centers do not change their location. Since the object chip is 256×256

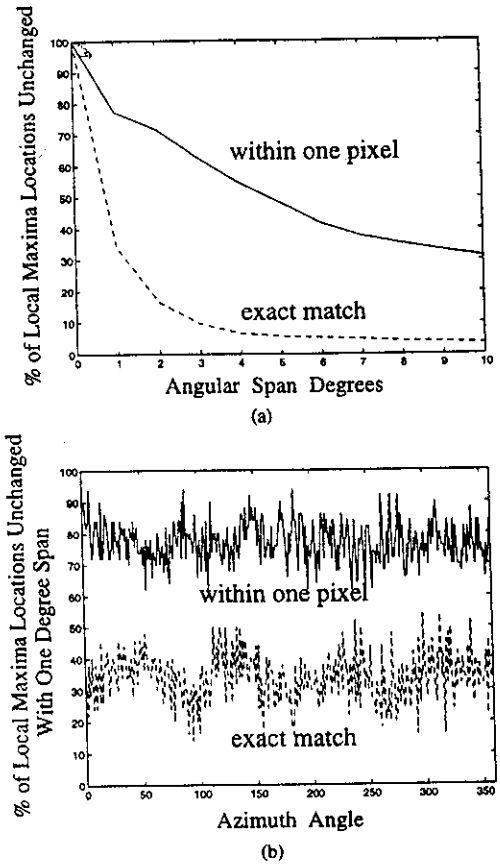


Figure 4: (a)T72 tank Rotational Invariance.(b)T72 tank Rotational Invariance With 1° Angular Span.

pixels, we rotate the image with respect to the center point $(127.5, 127.5)$. The distance measurement criteria "exact match" and "within one pixel" are defined in the following:

$$\begin{cases} \textbf{x}_\textbf{r} \textbf{ exactly matches x:} \\ \text{if } MAX(|x-x_r|, |y-y_r|) < \frac{1}{2} \text{ pixel} \\ \textbf{x}_\textbf{r} \textbf{ and x are within one pixel:} \\ \text{if } MAX(|x-x_r|, |y-y_r|) < 1\frac{1}{2} \text{ pixel} \end{cases}$$

Figure 4(a) shows the average result for images at all the 360 azimuth angles. The top 50 scattering centers are used for each image. Figure 4(b) gives the percentage of scattering center locations unchanged vs. azimuth angle with 1° angular span for the exact match and within one pixel match. These results show that scattering centers for SAR images vary greatly with relatively small changes of azimuth angles. As a result we represent an object at a given depression angle by 360 azimuths taken in steps of 1°.

3.3 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

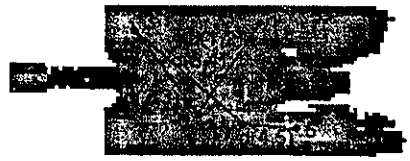


Figure 5: Example of an observation sequence superimposed on an image of T72 tank.

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 represent the image efficiently.

Based on the above considerations, we employ two approaches to obtain the sequences.

- Sequences based on relative amplitudes: $O_1 = \{Magnitude_1, Magnitude_2, ..., Magnitude_n\}$
- Sequences based on geometrical relationship:
  $O_2 = \{d(1,2), d(2,3), ..., d(n,1)\}$ $(length\ n)$
  $O_3 = \{d(1,2), d(1,3), ..., d(1,n)\}$ $(length\ n-1)$
  $O_4 = \{d(2,1), d(2,3), ..., d(2,n)\}$ $(length\ n-1)$
  $O_5 = \{d(3,1), d(3,2), ..., d(3,n)\}$ $(length\ n-1)$

where $Magnitude_i$ is the amplitude of ith scattering center and $d(i,j)$ is the euclidean distance between scattering centers i and j. Figure 5 gives an example to illustrate how we get the sequences. 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. In experiments described in section 4, we only consider the top 20 scattering centers. 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 [15] to classify the magnitude values (or distance values) of all the scattering centers in the database into K classes. 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, the sequence of magnitude values (or distance values) now is changed to a label between 1 to K which represents how different scattering centers fall into the different groups and finally, for a given sequence, we obtain a sequence of observation symbols.

3.4 Off-line Training Phase

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 4.1).
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.)

3.5 On-line Recognition Phase

The recognition procedure is described as follows:

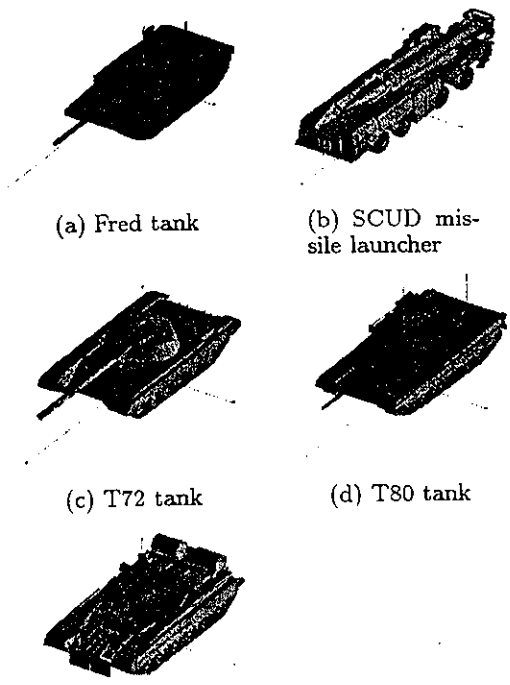
1. Loop lines 2-3 for all the testing observation sequences.
2. Loop line 3 for all the models in the model base.
3. Feed the observation sequence into the model, $(A, B, \Pi)_{(M_i^*, a_j^*)}$, Use Forward algorithm to compute the probability that this sequence is produced by this model.
4. The model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

4 Experiments

4.1 Data

Using the well known XPATCH [12] SAR simulator, we generate one set of SAR images of 5 objects (Fred tank, SCUD missile launcher, T72 tank, T80 tank and M1a1 tank, shown in Figure 6.) at 15° depression angle, 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: We consider the occlusion to occur possibly from 9 different directions as shown in Figure 7. 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 7. They are obtained by removing 4 scattering centers 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:



(a) Fred tank (b) SCUD missile launcher

(c) T72 tank (d) T80 tank

(e) M1a1 tank

Figure 6: Targets.

- 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. If r is between 1 and 20, we use the magnitude of the rth brightest image scattering center as the magnitude of the spurious center. Otherwise, we choose the magnitude of the 21st brightest scattering center if it was not already assigned to another spurious center. If it was already chosen, we will select the magnitude of the first unused scattering center (the 22nd, the 23rd, and so on).

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. Then 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 ($d = 0, 1, ..., 8$) 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.

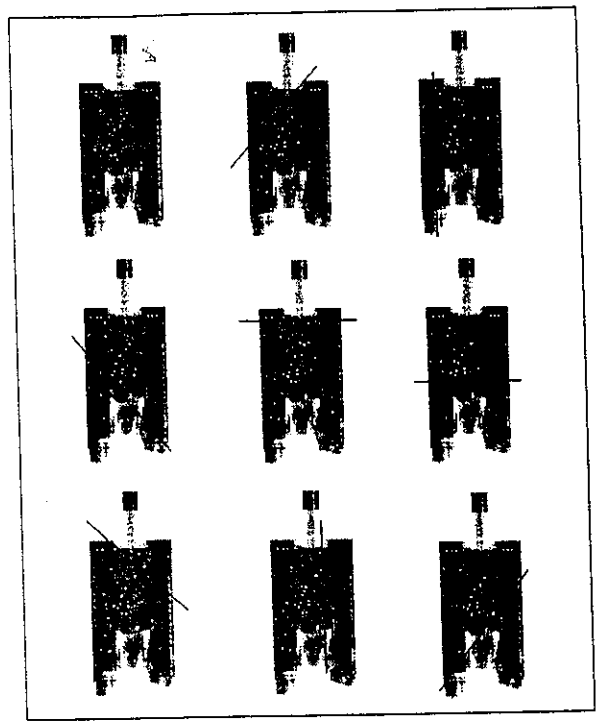


Figure 7: Scattering centers of T72 tank at azimuth 0°, part of scattering centers are occluded from a particular direction (0-8, left to right, top to bottom).

4.2 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 5 azimuth angles of five objects (Fred tank, SCUD missile launcher, T80 tank, T72 tank, and M1a1 tank). The results are shown in Table 1.

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 8 illustrates example parameters of a 5 state, 4 symbol HMM.

We have 1800 ($= 360 \times 5$) HMM models. Further, since we have defined five kinds of observation sequences for each image (O_1, O_2, O_3, O_4, O_5), we get models based on each kind of observation sequence.

Table 1: Recognition rate of HMM with different number of states and symbols.
N - # of states.
M - # of symbols.
R - Recognition rate % (top answer is correct).
I - Indexing rate % (correct answer is in the top 5).

		id only		*id with pose*	
N	M	R	I	R	I
4	8	76.1	96.5	62.6	79.9
4	16	89.6	98.4	85.4	93.1
4	24	95.1	99.3	91.8	97.3
4	32	96.6	99.9	94.8	99.0
4	64	99.7	100.0	99.6	100.0
5	8	80.1	97.4	67.3	84.0
5	16	91.9	98.6	86.7	93.7
5	24	96.6	99.7	94.6	98.6
5	32	97.8	99.8	96.7	99.3
5	64	99.9	100.0	99.9	100.0
6	8	82.5	96.9	71.7	84.8
6	16	93.8	99.5	90.1	96.7
6	24	98.5	99.8	97	99.7
6	32	98.9	100.0	98.5	99.9
6	64	100.0	100.0	100.0	100.0
8	8	84.3	97.6	77.4	87.6
8	16	96.4	99.8	94.6	98.3
8	24	99.4	100.0	99	99.9
8	32	99.8	100.0	99.8	100.0
8	64	100.0	100.0	100.0	100.0
10	8	100.0	100.0	100.0	100.0
10	16	98.3	99.9	97.3	99.6
10	24	99.9	100.0	99.9	99.9
10	32	100.0	100.0	99.9	100.0
10	64	100.0	100.0	100.0	100.0

4.3 Testing Results

During testing phase, 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. Otherwise, we count it as one incorrect recognition. After we get the results of scattering centers occluded from all 9 directions, we average the result and associate this recognition performance with the model.

Figure 9 shows the testing results for each of the five kinds of sequences: $O_1, O_2, ..., O_5$ (section 3.3). 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.

Integration of results from multiple sequences: Since not all models based on various sequences for a particular object and azimuth will provide optimal 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 models based on a particular sequence, say O_1, will it be recognized correctly by models based on other kinds of sequences?

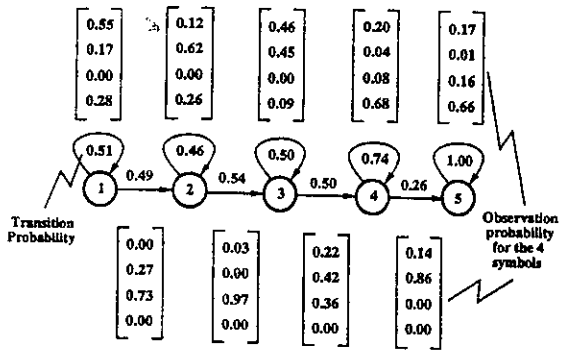


Figure 8: An example: parameters of a 5 states, 4 symbols HMM. The number on edges represents the transition probability, and the vector associated with each transition represents $b_{ij}(k)$. In our case, we use HMM with 8 states, 32 symbols

From the testing results, we obtained Table 2 which 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 draw two curves (Figure 10(a)) to show the possible "upper bound" and "lower bound" of recognition rate we can achieve based on the 5 kinds of models. We *define* the "upper bound" as the highest possible recognition performance that can be achieved using the 5 models (O_1 to O_5) considering *only* the top candidate for recognition from each of the 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 total number of errors corresponding to "upper bound" are shown in the 7th column of the Table. 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 11 to integrate the results from models based on 5 different sequences.

1. For each test image, we collect the ten highest possibilities in the testing results corresponding to each of the sequences $O_1, O_2, ..., O_5$.
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.

Table 2: Testing results for occluded object recognition using of 81,000 testing cases.

Percent. occlusion	*Errors with model O_2*	Errors *Captured by models*				*Errors using models O_1 to O_5*	*% Correct Recognition ("upper bound")*	*% Based on Integration Recognition*	*% Based on Integration 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
Average Recognition Rate							97.1	86.8	94.4

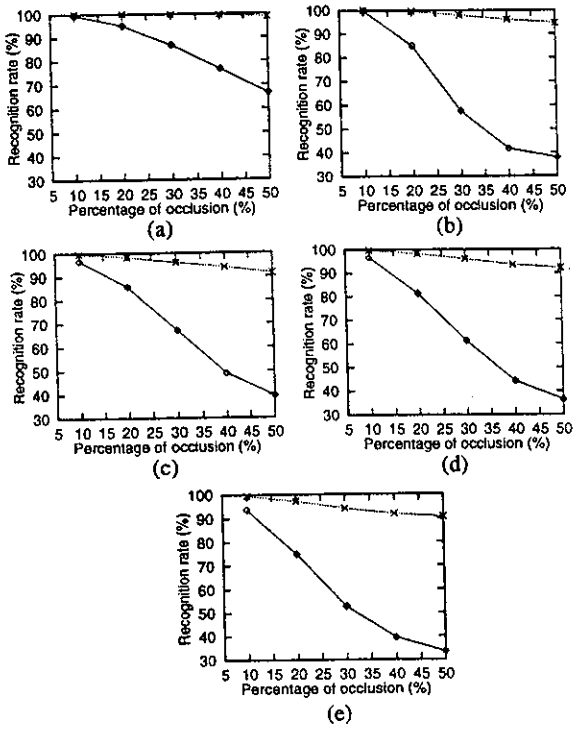


Figure 9: Recognition rate vs. percentage of occlusion for HMM models based on (a) O_1, (b) O_2, (c) O_3, (d) O_4, and (e) O_5.

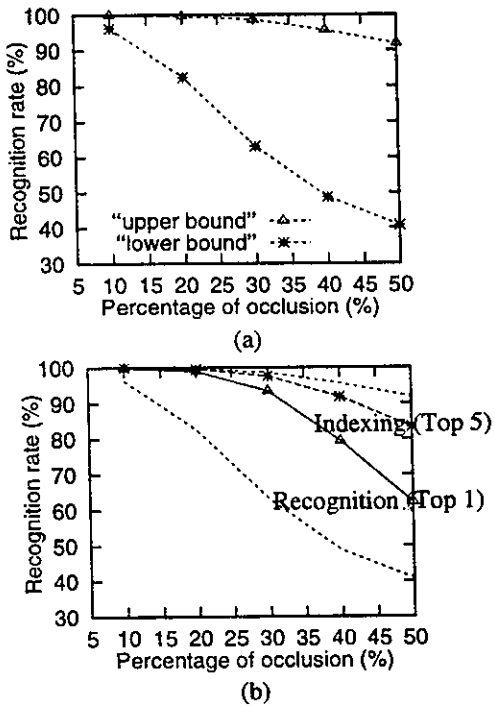


Figure 10: (a) "Upper" and "lower" bound of recognition rate vs. percentage of occlusion. (b)Performance of integrated models: using integrated models O_1 to O_5. The results for recognition (Top 1) and indexing (Top 5) candidates are superimposed on the figure shown in (a).

3. We draw a histogram with probability vs. object and pose (here we combine object and pose as one parameter). This is because 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 ground truth, we count it as one correct recognition.

The second curve from the bottom in Figure 10(b) is the result. The corresponding confusion matrix for various amounts of occlusion is shown in Table 3. 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 10(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 in Figure 10(a) into Figure 10(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.

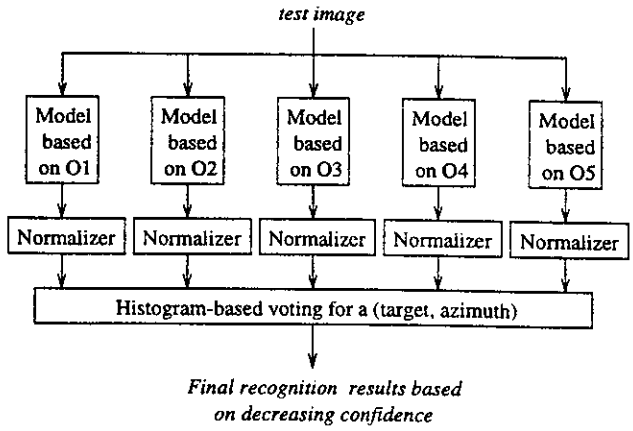


Figure 11: Integration of results by histogram-based method.

5 Conclusions and Future Work

We have presented a novel 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. Using 99,000 training samples and 81,000 testing samples, we find 86.76% average correct recognition performance on 5 classes of objects with 10% – 50% occlusion. 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.

We have also done some initial experiments for articulated object recognition using HMM approach. We have three sets of data: the original images for the objects (T72 tank, T80 tank, and M1a1 tank), the images for the objects with turret at 60 degree articulation, and the images for the objects with turret at 90 degree articulation. We compared the observation sequences O_1 extracted from the three sets of images. Figure 12 shows the analysis graph for T72 tank. Figure 12 (a1), (b1), and (c1) are obtained by counting the number of observation symbols in observation sequence of one image which are the same as its corresponding one in observation sequence of another image. Figure 12 (a2), (b2), and (c2) are obtained by counting the sum of differences between observation symbols in observation sequence of one image and its corresponding one in observation sequence of another image.

We used two sets out of three sets of images as training data to train the HMM models, and tested the HMM models on the other set. Table 4 shows the results. These experimental results are obtained by using observation sequence O_1 only, the experiments using other sequences O_2 through O_5 will be done in the future.

Table 3: Confusion Matrix for 5 objects classes at varying amounts of occlusion (10% – 50%).

	% Occlusion	*Fred*	*SCUD*	*T72*	*T80*	*M1a1*
Fred	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
SCUD	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
T72	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
T80	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
M1a1	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

References

[1] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proc. of the IEEE*, Vol. 77(2), pp. 257-285, 1989.

[2] L. R. Rabiner and B. H. Juang, "An introduction to Hidden Markov Models", *IEEE ASSP Magazine*, Vol. 3(1), pp. 4-16, 1986.

[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 Technical Journal*, Vol. 64(6), pp. 1211-1233, 1985.

[4] O. E. Agazzi and S. S. Kuo, "Pseudo Two-Dimensional Hidden Markov Models for Document Recognition", *AT & T Technical Journal*, Vol. 72(5), pp. 60-72, 1993.

[5] O. E. Agazzi and S. S. Kuo, "Hidden Markov Model based optical character recognition in the presence of deterministic transformations", *Pattern Recognition*, Vol. 26(12), pp. 1813-1826, November, 1993.

[6] B. Bhanu, G. Jones, J. Ahn, M. Li and J. Yi, "Recognition of Articulated Objects in SAR Images", *In Proceedings DARPA Image Understanding Workshop*, pp. 1237-1250, 1996.

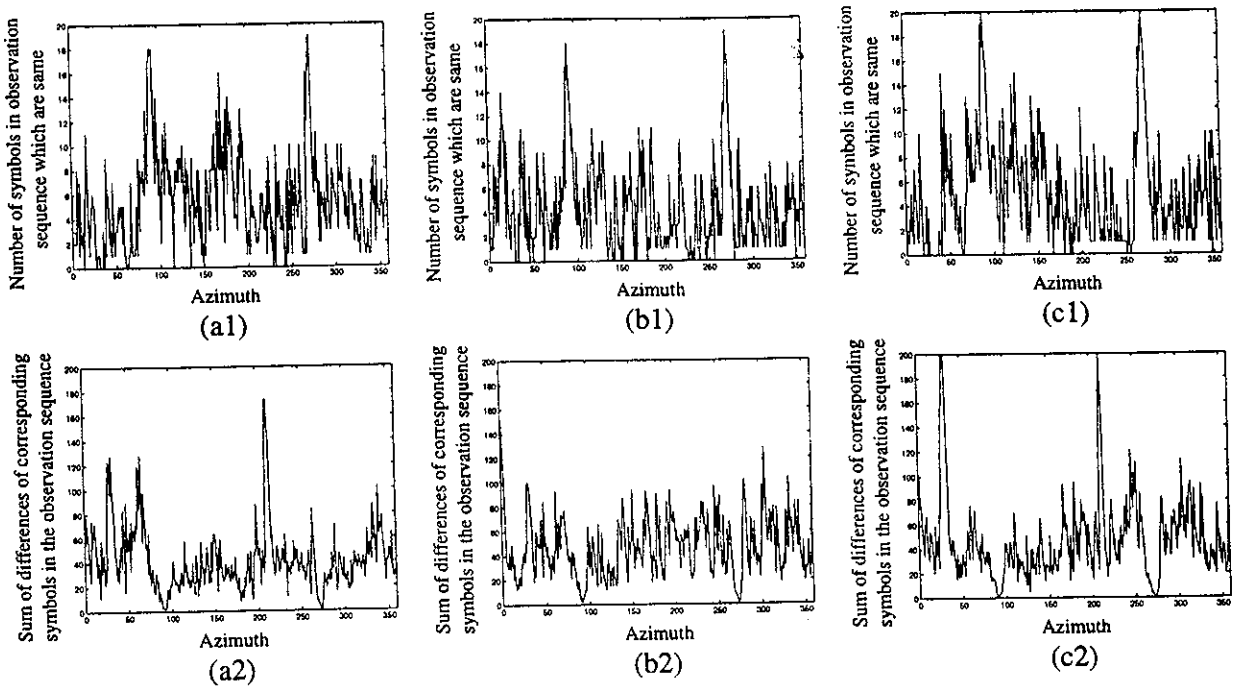


Figure 12: Comparison of observation sequence O_1 extracted from three sets of images for T72 tank. (a1, a2) 0° vs. 60°, (b1, b2) 0° vs. 90°, (c1, c2) 60° vs. 90°.

Table 4: Indexing results for articulated objects (id only).

Training T72	Testing T72	Indexing (top 5) performance
Turret at 0° and 60° 720 images	Turret at 90° 360 images	95.2
Turret at 0° and 90° 720 images	Turret at 60° 360 images	94.0
Turret at 60° and 90° 720 images	Turret at 0° 360 images	94.6
Average Performance		94.6

[7] K. H. Fielding and D. W. Ruck, "Spatio-Temporal Pattern Recognition Using Hidden Markov Models", *IEEE Trans. on AES*, Vol. 31(4), pp. 1292-1300, 1995.

[8] R. R. Rao and R. M. Mersereau, "On Merging Hidden Markov Models with Deformable Templates", *ICIP Proc.*, pp. 556-559, 1995.

[9] 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*, Vol. 6(1), pp. 11-24, 1993.

[10] J.H. Yi, B. Bhanu and M. Li, "Target indexing in SAR images using scattering centers and Hausdorff distance", *Pattern Recognition Letters*, Vol. 17, pp. 1191-1198, September 1996.

[11] K. Ikeuchi, T. Shakunaga, M. D. Wheeler, and T. Yamazaki, "Invariant Histograms and Deformable Template Matching for SAR Target Recognition", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 100-105, 1996.

[12] 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", *Proceedings of Ground Target Modeling and Validation Conference*, pp. 498-507, 1994.

[13] K. F. Lee, *Automatic Speech Recognition - The Development of the SPHINX System*, Kluwer Academic Publishers, 1989.

[14] J. P. Fitch, *Synthetic Aperture Radar*, Springer-Verlag, 1988.

[15] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*, Addison-Wesley, 1974.